The cover time of a biased random walk on a random cubic graph

Colin Cooper*   Alan Frieze†   Tony Johansson‡

January 3, 2018

Abstract

We study a random walk that prefers to use unvisited edges in the context of random cubic graphs. We establish asymptotically correct estimates for the vertex and edge cover times, these being $\approx n \log n$ and $\approx \frac{3}{2} n \log n$ respectively.

1   Introduction

Our aim in this paper is to analyse a variation on a simple random walk that may tend to speed up the cover time of a connected graph. This variation is just one of several possible approaches which include (i) non-bactracking walks, see Alon, Benjamini, Lubetzky and Sodin [3], (ii) walks that prefer unused edges, see Berenbrink, Cooper and Friedetzky [4] or (iii) walks that a biassed toward low degree vertices, see Cooper, Frieze and Petti [7] or any number of other ideas. In this paper we study idea (ii) in the context of random cubic graphs, partially solving a problem left from [4].

1.1   Unvisited Edge Process

The papers [4], [11] describe a modified random walk $X = (X(t), \ t \geq 0)$ on a graph G, which uses unvisited edges when available at the currently occupied vertex. If there are *unvisited edges* incident with the current vertex, the walk picks one u.a.r. and make a transition along

this edge. If there are no unvisited edges incident with the current vertex, the walk moves to a random neighbour.

In [4] this walk was called an *unvisited edge process* (or edge-process), and in [11], a *greedy random walk*. For random d-regular graphs where $d = 2k$ (d even), it was shown in [4] that the edge-process has vertex cover time $\Theta(n)$, which is best possible up to a constant. The paper also gives an upper bound of $O(n\omega)$ for the edge cover time. The ω term comes from the w.h.p. presence of small cycles (of length at most ω). In [6], the constant for the vertex cover time was shown to be $d/2$.

Theorem 1. *Let X be an unvisited edge-process on a random d-regular graph, d even. For $d \geq 4$, the following holds w.h.p. The vertex cover time of the edge-process is $T_{cov}^V(G) \sim dn/2$.*

The paper [4] included the experimental data shown in Figure 1 for the performance of red-blue walks on odd degree regular graphs. Namely, for $d = 3$ the cover time is $\Theta(n \log n)$ and decreases rapidly with increasing d. For d even, the experiments confirm the cover time result of Theorem 1 that $T_{\text{COV}}^V(G) \sim dn/2$.
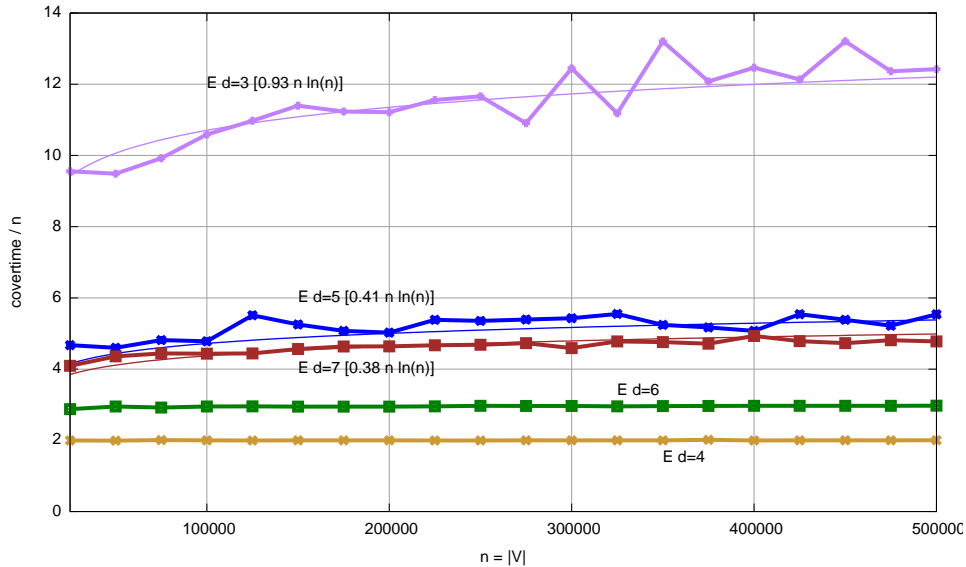

Figure 1: Normalised cover time of the unvisited edge process on random d-regular graphs as function of $n = |V|$. Figure from [4]

1.2 Our results

Let $G = (V, E)$ be a connected 3-regular (multi)graph on an even number n of vertices. Consider the following random walk process, called a *biased random walk*. It is an edge colored version of the previously described unvisited edge process. Initially color all edges red, and pick a starting vertex v_0. At any time, if the walk occupies a vertex incident to at least one red edge, then the walk traverses one of those red edges chosen uniformly at

2

random, and re-colors it blue. If no such edge is available, the walk traverses a blue edge chosen uniformly at random. For $s \in \{1, \ldots, n\}$ let $C_V(s)$ denote the number of steps taken by the walk until it has visited s vertices, and similarly let $C_E(t)$ denote the number of steps taken to visit $t \in \{1, \ldots, 3n/2\}$ edges.

We will let G be a random graph, and we use $\mathbb{E}_G(X)$ to denote the expectation of X with the underlying graph G fixed.

Theorem 2. *Let s, t be fixed such that $n - n \log^{-1} n \le s \le n$ and $(1 - \log^{-2} n)\frac{3n}{2} \le t \le 3n/2$. Let $\varepsilon > 0$ also be fixed. Suppose G is chosen uniformly at random from the set of 3-regular graphs on n vertices. Then with high probability, G is connected and*

$$\mathbb{E}_G(C_V(s)) = (1 \pm \varepsilon)n \log \left(\frac{n}{n - s + 1} \right) + o(n \log n), \tag{1}$$

$$\mathbb{E}_G(C_E(t)) = \left(\frac{3}{2} \pm \varepsilon \right) n \log \left(\frac{3n}{3n - 2t + 1} \right) + o(n \log n). \tag{2}$$

Here $a = b \pm c$ is taken to mean $a \in [b - c, b + c]$. Note in particular that this shows that the expected vertex and edge cover times are asymptotically $n \log n$ and $\frac{3}{2}n \log n$ with high probability, respectively. The same statement is true with the word "graphs" replaced by "configuration multigraphs". Thus we have the following corollary.

Corollary 1. *W.h.p. the vertex cover time $T_{cov}^V(G)$ of G is asymptotically equal to $n \log n$ and the edge cover time $T_{cov}^E(G)$ is asymptotically equal to $\frac{3}{2}n \log n$.*

It is of interest to compare the result of Corollary 1 with other versions of random walk. Cooper and Frieze [5] showed that w.h.p. the vertex cover time of a random d-regular graph on n vertices is asymptotically equal to $\frac{r-1}{r-2}n \log n$. The argument there also shows that the edge cover time of a random d-regular graph on n vertices is asymptotically equal to $\frac{r(r-1)}{2(r-2)}n \log n$. For $r = 3$ these values are $2n \log n$ and $3n \log n$ respectively and are to be compared with $n \log n$ and $\frac{3}{2}n \log n$. For a non-bactracking random walk, Cooper and Frieze [6] show that the vertex and edge cover times are asymptotically $n \log n$ and $\frac{r}{2}n \log n$ respectively. Interestingly, these values coincide with the results in Corollary 1.

2 Outline proof of Theorem 2

We will choose the multigraph G according to the configuration model. Each vertex v of G is associated with a set $\mathcal{P}(v)$ of 3 configuration points. We set $\mathcal{P} = \cup_v \mathcal{P}(v)$ and generate G by choosing a pairing μ of $\mathcal{P}$ uniformly at random. The pairing μ is exposed along with the biased random walk.

Starting at a uniformly random configuration point $x_1 \in \mathcal{P}$, we define $W_0 = (x_1)$. Given a walk $W_k = (x_1, x_2, \ldots, x_{2k+1})$, the walk proceeds as follows. Set $x_{2k+2} = \mu(x_{2k+1})$, thus exposing the value of x_{2k+1} if not previously exposed. If x_{2k+2} belongs to a vertex v which

is incident to some red edge (other than (x_{2k+1}, x_{2k+2}) which is now recoloured blue), the walk chooses one of the red edges uniformly at random, setting x_{2k+3} to be the corresponding configuration point. Otherwise, x_{2k+3} is chosen uniformly at random from $\mathcal{P}(v)$. Set $W_{k+1} = (x_1, \ldots, x_{2k+3})$. We will refer to x_1 and x_{2k+1} (and the vertices to which they belong) as the *tail* and *head* of W_k, respectively. We will also refer to $\{x_1, x_2, \ldots, x_{2k+1}\}$ as the points of $\mathcal{P}$ that have been *visited*.

Define partial edge and vertex cover times

$$C_E(t) = \min\{k : W_k \text{ spans } t \text{ edges}\}, \tag{3}$$
$$C_V(t) = \min\{k : W_k \text{ spans } t \text{ vertices}\}. \tag{4}$$

We will mainly be concerned with the partial edge cover time, and write $C(t) = C_E(t)$ from this point on.

For $t \in \{1, 2, \ldots, \frac{3n}{2}\}$ we define a subsequence of walks by

$$W(t) = W_{C(t)-1} = (x_1, x_2, \ldots, x_{2k+1}) \tag{5}$$

where k is the smallest integer such that $|\{x_1, x_2, \ldots, x_{2k+1}\}| = 2t - 1$. In other words, $W(t)$ denotes the walk up to the point when $2t - 1$ of the members of $\mathcal{P}$ have been visited. Thus throughout the paper:

- Time t is measured by the number of edges t that have been visited at least once.

- The parameter $\delta = \delta(t)$ is given by the equation

$$t = (1 - \delta)\frac{3n}{2}. \tag{6}$$

  $\delta(t)$ is important as a measure of how close we are to the edge cover time.

- the walk length k is measured by the number of steps taken so far. Equation (5) relates t and k.

A 3-regular graph G chosen u.a.r. is connected w.h.p. and we will implicitly condition on this in what follows. The bulk of the paper will be spent proving the following lemma.

Lemma 1. *For any fixed $\varepsilon > 0$ and $(1 - \log^{-2} n)\frac{3n}{2} \leq t \leq \frac{3n}{2}$,*

$$\mathbb{E}\left(C(t)\right) = \left(\frac{3}{2} \pm \varepsilon\right) n \log\left(\frac{3n}{3n - 2t + 1}\right) + o(n \log n) \tag{7}$$

for n large enough. Furthermore, for $n - \frac{n}{\log n} \leq s \leq n$,

$$(1 - \varepsilon) n \log\left(\frac{n}{n - s + 1}\right) \leq \mathbb{E}\left(C_V(s)\right) \leq (1 + \varepsilon) n \log\left(\frac{n}{n - s + 1}\right). \tag{8}$$

4

Expectations in Lemma 1 are taken over the full probability space. In particular, if $\mathcal{G}$ denotes the set of graphs,

$$\frac{3}{2}n\log\left(\frac{3n}{3n-2t+1}\right) \approx \mathbb{E}\left(C(t)\right) = \frac{1}{|\mathcal{G}|}\sum_{G\in\mathcal{G}}\mathbb{E}_G\left(C(t)\right).$$

In Section 7.6 we strengthen Lemma 1 to stating that almost every G satisfies $\mathbb{E}_G\left(C(t)\right) \approx \frac{3}{2}n\log(3n/(3n-2t+1))$ (and similarly for $C_V(s)$). Theorem 2 follows.

An essential part of the proof of Lemma 1 is a set of recurrences for the random variables $X_i(t)$, where $X_i(t)$ is the number of vertices incident with $i = 0, 1, 2, 3$ untraversed edges at *time* t, $t = 1, 2, ..., 3n/2$. We will argue that for most of the process, it takes approximately $3n/(3n-2t)$ steps of the walk to increase time by one. As the process finishes at time $3n/2$ we see that the edge cover time should be approximately

$$\sum_{t=1}^{3n/2}\frac{3n}{3n-2t+1} \approx \frac{3}{2}n\log n,$$

as claimed in Corollary 1.

Now the recurrence for $X_3(t)$ has a solution that implies that $X_3(t) \approx n\delta^{3/2}$, where δ is as in (6). Given this, we would expect $X_3(t)$ to be zero when δ is smaller than $n^{-2/3}$ or equivalently, when $3n/2 - t$ is less than $n^{1/3}$. Thus we would expect that vertex cover time to be

$$\sum_{t=1}^{3n/2-n^{1/3}}\frac{3n}{3n-2t+1} \approx n\log n,$$

as claimed in Corollary 1.

We separate the proof of Lemma 1 into phases. Define

$$\delta_0 = \frac{1}{\log\log n}, \quad \delta_1 = \log^{-1/2}n, \quad \delta_2 = \log^{-2}n, \quad \delta_3 = n^{-2/3}\log^4 n \text{ and } \delta_4 = n^{-1}\log^{11}n$$

and set

$$t_i = (1-\delta_i)\frac{3n}{2} \text{ for } i = 0, 1, 2, 3, 4. \tag{9}$$

We do not hesitate to remind the reader of the meaning of these quantities.

The first phase, in which the first t_1 edges are discovered, will not contribute significantly to the cover time.

Lemma 2. *Let $\delta_1 = \log^{-1/2}n$ and $t_1 = (1-\delta_1)\frac{3n}{2}$. Then*

$$\mathbb{E}\left(C(t_1)\right) = o(n\log n).$$

Between times t_1 and t_4 we bound the time taken between discovering new edges. The proof, in Section 7, will be split into the ranges $t_1 \leq t \leq t_3$ and $t_3 \leq t \leq t_4$.

5

Lemma 3. *Let $\varepsilon > 0$. For $t_1 \leq t \leq t_4$ and n large enough,*

$$\mathbb{E}\left(C(t+1) - C(t)\right) = (3 \pm \varepsilon)\frac{n}{3n - 2t} + O(\log n).$$

Note that because $\frac{3n}{2} - t_1 = O(\delta_1 n)$, the $O(\log n)$ term only contributes an amount $O(n\delta_1 \log n) = o(n \log n)$ to the the edge cover time.

Finally, the following lemma shows that the final $\log^{11} n$ edges can be found in time $o(n \log n)$.

Lemma 4. *For $t > t_4$ and n large enough,*

$$\mathbb{E}\left(C(t) - C(t_4)\right) = o(n \log n).$$

We note now that Lemma 1 follows from Lemmas 2, 3 and 4.

3 Structural properties of random cubic graphs

Here we collect some properties of random cubic graphs.

Lemma 5. *Let G denote the random cubic graph on vertex set $[n]$, chosen according to the configuration model. Let ω tend to infinity arbitrarily slowly with n. Its value will always be small enough so that where necessary, it is dominated by other quantities that also go to infinity with n. Then with high probability,*

(i) *The second largest in absolute value of the eigenvalues of the transition matrix for a simple random walk on G is at most $2\sqrt{2}/3 + \varepsilon \leq .99$ for some ε.*

(ii) *G contains at most $\omega 3^\omega$ cycles of length at most ω,*

(iii) *The probability that G is simple is $\Omega(1)$.*

For the proof of (i) see Friedman [8] and for the proof of (iii) see Frieze and Karoński [9], Theorem 10.3. Property (ii) follows from the Markov inequality, given that the expected number of cycles of length $k \leq \omega$ can be bounded by $O(3^k)$.

Let $G(t)$ denote the random graph formed by the edges visited by $W(t)$. Let $X_i(t)$ denote the set of vertices incident to i red edges in $G(t)$ for $i = 0, 1, 2, 3$. Let $\overline{X}(t) = X_1(t) \cup X_2(t) \cup X_3(t)$. Let $G^*(t)$ denote the graph obtained from $G(t)$ by contracting the set $\overline{X}(t)$ into a single vertex, retaining all edges. Define $\lambda^*(t)$ to be the second largest eigenvalue of the transition matrix for a simple random walk on $G^*(t)$.

We note that [2, Corollary 3.27], if Γ is a graph obtained from G by contracting a set of vertices, retaining all edges, then $\lambda(\Gamma) \leq \lambda(G)$. This implies that $\lambda^*(t) = \lambda(G^*(t)) \leq \lambda(G) \leq 0.99$ for all t. Initially, for small t, we find that w.h.p. $G^*(t)$ consists of a single vertex. In this case there is no second eigenvalue and we take $\lambda^*(t) = 0$. This is in line with the fact that a random walk on a one vertex graph is always in the steady state.

4 Random walks and hitting times

We are interested in calculating $C(t + 1) - C(t)$, i.e. the time taken between discovering the tth and the $(t + 1)$th edge. Between the two discoveries, the biased random walk can be coupled to a simple random walk on the graph induced by $W(t)$ which ends as soon as it hits a vertex of $\overline{X}$. In this section we derive the hitting time of a certain type of expanding vertex set.

Consider a simple random walk on a cubic graph $G = (V, E)$ with eigenvalue gap $1 - \lambda > 0$. For a set S of vertices and a probability measure ρ on V, let $\mathbb{E}_\rho (H(S))$ denote the expected hitting time of the set S, when the initial vertex is chosen according to ρ. Let π denote the stationary distribution of the random walk, uniform in the case of a regular graph and proportional to degrees in general. Let $P_u^{(t)}(v)$ denote the probability that a simple random walk starting at u occupies vertex v at step t of the walk.

Lemma 6. *Suppose v is a vertex of a graph. Then the hitting time of v, starting from the stationary distribution π, is given by*

$$\mathbb{E}_\pi (H(v)) = \frac{Z_{vv}}{\pi_v}$$

where

$$Z_{vv} = \sum_{t \geq 0} (P_v^{(t)}(v) - \pi_v).$$

Lemma 6 can be found in [2] (Lemma 2.11), and can be applied to hitting times of sets by contracting a set of vertices to a single vertex. The following bound will be frequently used. Suppose G is a graph with eigenvalue gap $1 - \lambda(G)$, and S is a set of vertices in G. Then if G_S is the graph obtained by contracting S into a single vertex, retaining all edges, we have equal hitting times for S in G and G_S and

$$\mathbb{E}_\pi (H(S)) = \frac{n}{|S|} \sum_{t \geq 0} \left(P_S^{(t)}(S) - \pi_S \right) \leq \frac{n}{|S|} \sum_{t \geq 0} \lambda(G_S)^t = \frac{1}{1 - \lambda(G_S)} \frac{n}{|S|} \leq \frac{1}{1 - \lambda(G)} \frac{n}{|S|}.$$
(10)

Indeed, $|P_v^{(t)}(v) - \pi_v| \leq \lambda^t$ for any v, t in a graph with eigenvalue gap $1 - \lambda$, see for example Jerrum and Sinclair [10] and use $j = k$ in the middle of the proof there of Proposition 3.1. Also, $\lambda(\Gamma) \leq \lambda(G)$ for any Γ obtained from G by contracting a set of vertices (see [2, Corollary 3.27]).

In the following lemma we implicitly view G as a member of a sequence of graphs (G_n), and G having positive eigenvalue gap means that the second largest eigenvalue λ_n of G_n satisfies $\limsup \lambda_n < 1$.

Define $N_d(S)$ to be the set of vertices at distance exactly d from a vertex set S. The set S we consider induces $|S|/2$ edges all of which are 'far apart'.

Lemma 7. *Let G be a cubic graph on n vertices with positive eigenvalue gap. Suppose S is a set of vertices with $|S| \geq 2$ even such that $|S| = o(n)$ and*

$$|N_d(S)| = 2^d |S|$$

7

for all $1 \leq d \leq \omega$, where ω tend to infinity arbitrarily slowly with n. Then

$$\mathbb{E}_\pi(H(S)) \approx \frac{3n}{|S|}.$$

Proof. We first note that the set S contains exactly $|S|/2$ edges. Indeed, as $|N(S)| = 2|S|$ and the total degree of S is $3|S|$, S contains at most $|S|/2$ edges. As $|N_2(S)| = 4|S|$, each vertex of $N(S)$ must have exactly one edge to S, implying that S contains at least $|S|/2$ edges.

Consider the graph G_S obtained by contracting S into a single node s, retaining all edges. In the graph G_S, s has degree $3|S|$. Then s is a node with exactly $|S|/2$ self-loops, and is otherwise contained in no cycle of length at most ω, as $|N_d(S)| = 2^d|S|$ ensures that G_S is locally a tree up to distance ω from s. Since $\pi_s = |S|/n = o(1)$ we may choose ω tending to infinity with $\omega\pi_s = o(1)$. We have

$$Z_{ss} = \sum_{t \geq 0}(P_s^{(t)}(s) - \pi_s) = \left(\sum_{t=0}^{\omega} P_s^{(t)}(s)\right) - o(1) + \sum_{t > \omega}(P_s^{(t)}(s) - \pi_s).$$

Repeating the argument following (10),

$$\sum_{t > \omega}|P_s^{(t)}(s) - \pi_s| \leq \sum_{t > \omega}\lambda^t = O(\lambda^\omega) = o(1).$$

We now argue that

$$\sum_{t=0}^{\omega} P_s^{(t)}(s) = 3 + o(1).$$

It is argued in Cooper and Frieze [5], Lemma 7, that with no loops at vertex s, the expected number of returns to s within ω steps is $2 + o(1)$. With the loops, when at s, there is a $1/3$ chance of using a loop and so each visit to s yields $3/2$ expected returns; i.e. the 2 of [5] becomes $3 = 2 \times 3/2$. $\square$

We now expand Lemma 7 to a larger class of sets.

Definition 1. *Let $G = (V, E)$ be a cubic graph. A set $S \subseteq V$ is a root set of order ℓ if (i) $|S| \geq \ell^5$, (ii) the number of edges with both endpoints in S is between $|S|/2$ and $(1/2+\ell^{-3})|S|$, and (iii) there are at most $|S|/\ell^3$ paths of length at most ℓ between vertices of S that contain no edges between a pair of vertices in S.*

Root sets of large order may be thought of as sets that "almost satisfy the hypothesis of Lemma 7". The following lemma shows this definition is suitable for our purposes.

Lemma 8. *Let ω tend to infinity arbitrarily slowly with n. Suppose G is a cubic graph on n vertices with positive eigenvalue gap, containing at most ω^2 cycles of length at most ω. If S is a root set of order ω, then*

$$\mathbb{E}_\pi(H(S)) \approx \frac{3n}{|S|}.$$

8

Proof. Consider the contracted graph G_S, and let s denote the contracted node. Then s has degree $3|S|$, and s has at most $(1/2 + 2\omega^{-3})|S|$ self-loops. Apart from the self-loops, s lies on at most $|S|/\omega^3$ cycles of length at most ω, as any cycle of G_S containing s corresponds to a path between members of S in G.

Let $R = N(S)$, and note that $|R| = \Omega(|S|)$. Consider the graph Γ, defined as G_S induced on the set of vertices at distance $1, 2, \ldots, \omega$ from s. Note that s is not included in Γ. The graph Γ contains all of R, and as s lies on at most $|S|/\omega^3$ short cycles in G_S, the number of components in Γ containing more than one member of R is $O(|S|/\omega^3) = O(|R|/\omega^3)$. As G contains at most ω^2 short cycles, the number of components of Γ containing a cycle is at most $\omega^2 = O(|R|/\omega^3)$. This leaves $(1 - o(1))|R|$ connected components in Γ which are all complete binary trees of height ω, each rooted at a member of R and containing no other member of R. Let T denote the set of vertices on such components.

Arbitrarily choose $|S|/2$ of the self-loops of s in G_S, and designate them as *good*. Also say that an edge is good if it has both endpoints in $T \cup \{s\}$. All other edges are *bad*.

Consider a simple random walk $Z(\tau)$ of length ω on G_S, starting at s. Let $\mathcal{B}_\tau$ denote the event that $Z(\tau)$ traverses a bad edge to reach $Z(\tau + 1)$. Whenever the walk visits s, the probability that it chooses a bad edge is $O(\omega^{-3})$. If the walk is inside T, there are no bad edges to choose. So for any $\tau \geq 0$ we have

$$P_s^{(\tau)}(s) = \Pr\left\{ Z(\tau) = s \cap \bigcap_{r=0}^{\tau-1} \overline{\mathcal{B}_r} \right\} + \Pr\left\{ Z(\tau) = s \cap \bigcup_{r=0}^{\tau-1} \mathcal{B}_r \right\} = \Pr\left\{ Z(\tau) = s \cap \bigcap_{r=0}^{\tau-1} \overline{\mathcal{B}_r} \right\} + O(\omega^{-2}).$$

If $\mathcal{B}_r$ does not occur for any $r \leq \tau - 1$, then the walk $(Z(0), \ldots, Z(\tau - 1))$ can be viewed as the same Markov chain as considered in Lemma 7. So,

$$\sum_{\tau=0}^{\omega} P_s^{(\tau)}(s) = 3 + O(\omega^{-1}).$$

$\square$

We will argue that X_1 quickly makes up all but a $o(1)$ fraction of the vertices of $\overline{X}$. For the purposes of discussion, we regard all edges of $G(t)$ visited exactly once by the walk $W(t)$ as coloured green. In Section 5 we argue that the vertices of X_1 are, in a sense that will be made precise, uniformly distributed on those edges of $G(t)$ which are visited exactly once by $W(t)$. In Section 8 we prove that toward the end of the walk, the number of green edges is significantly larger than X_1, which implies that most vertices of X_1 will be separated by a large distance. As G contains few short cycles and X_1 makes up most of $\overline{X}$, this will imply that the number of vertices at distance k from $\overline{X}$ is about $2^k|\overline{X}|$. In Section 4 we show how this implies that the expected hitting time of $\overline{X}$ in the simple random walk is approximately $3n/|\overline{X}|$, which by concentration (Lemma 10) is about $3n/(3n - 2t)$. Thus the partial edge cover time $C(t)$, the time to visit t edges, will be given by

$$\mathbb{E}\left(C(t)\right) = \sum_{r=0}^{t-1} \mathbb{E}\left(C(r+1) - C(r)\right) \approx \sum_{r=0}^{t-1} \frac{3n}{3n - 2r} \approx \frac{3}{2} n \log\left(\frac{3n}{3n - 2t}\right).$$

In Section 5 we show how the vertices of X_1 are distributed within green edges. In Section 7 we use these results to calculate $\mathbb{E}\left(C(t+1) - C(t)\right)$ assuming certain concentration results, which are then proved in Section 8.

5   Random distribution of once-visited vertices

Eventually the biased random walk will spend the majority of its time at vertices in X_0, i.e. vertices with no red incident edges. To bound the cover time, we will bound the time taken to hit $X_1 \cup X_2$, which may be thought of as the boundary of X_0.

Let $W_k, k \geq 0$ denote the biased random walk after $2k + 1$ walk steps have been taken. Say that a fixed finite walk W is *feasible* if $\Pr\{W_k = W\} > 0$ for some $k \geq 0$, and fix a feasible walk W. Let t be the time associated with W as indicated in (5). Let Y denote the subset of vertices in $X_1(t)$ that were visited and left exactly once by W. Note that $|Y \triangle X_1| \leq 1$, as the tail v_0 and head v_k of the walk are the only vertices which may be in X_1 after being visited twice and then only when $v_0 = v_k$. Indeed, the first time a vertex v is visited, a feasible walk must enter and exit v via distinct edges. Color all vertices of Y green. We can write $Y = X_1(t) \setminus \{v_0\}$.

Given a feasible walk W, define a *green bridge* to be a part of the walk starting and ending in $V \setminus Y$, with any internal vertices being in Y. Note also that it is not necessary for a green bridge to contain any vertices of Y. Form the *contracted walk* $\langle W \rangle$ by replacing any green bridge by a single green edge, with the walk orientation intact. Let $[W]$ denote the pair of (contracted walk, set), $[W] = (\langle W \rangle, Y)$, noting that $\langle W \rangle$ contains no vertex of Y.

We define an equivalence relation on the set of feasible walks by saying that $W \sim W'$ if and only if $[W] = [W']$. See Figure 2. Thus the only way that W, W' differ is as to where the vertices in Y are placed on the green bridges.
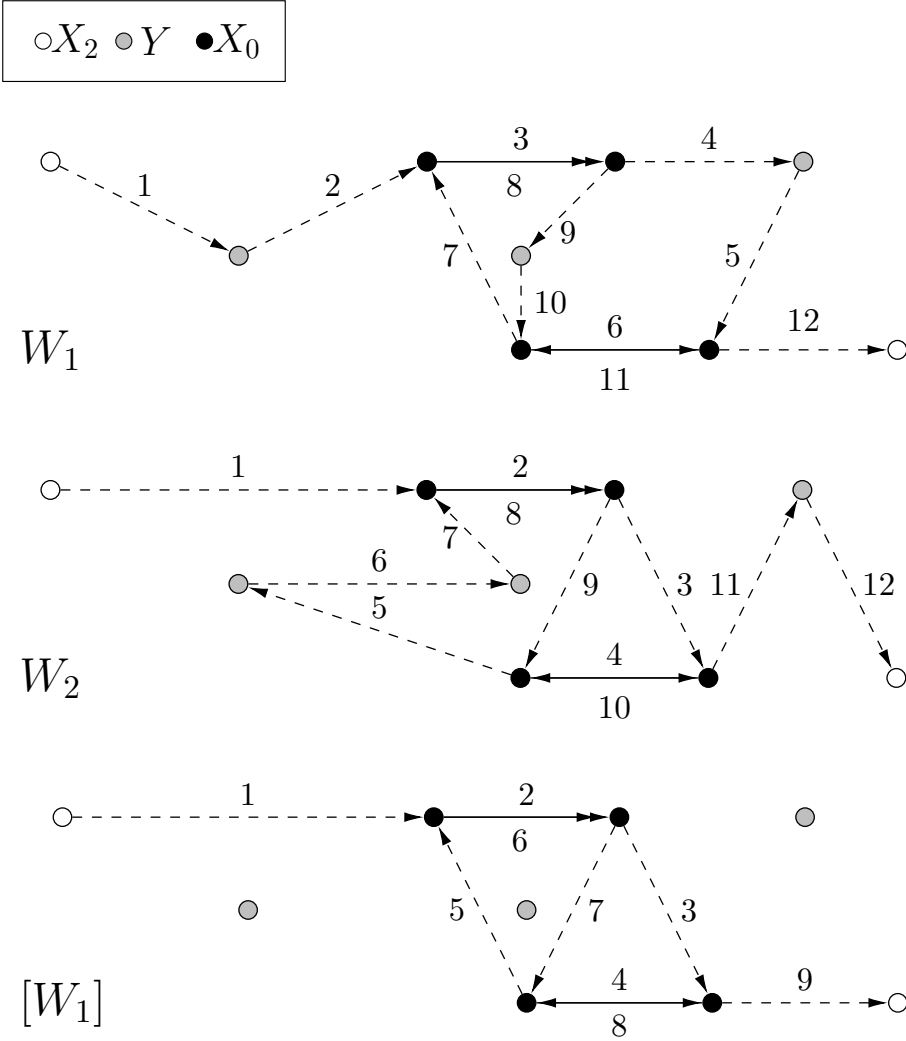
Figure 2: Two equivalent walks. Unvisited edges and vertices are not displayed, and edges visited exactly once are dashed. Lemma 9 shows that the walks are equiprobable.

Lemma 9. *Let $k > 0$ and suppose W is such that $Pr\{W_k = W\} > 0$. If $[W] = (\langle W \rangle, Y)$ and $\langle W \rangle$ contains ϕ green edges, then*

$$Pr\{W_k = W \mid [W_k] = [W]\} = \frac{1}{|[W]|} = \frac{1}{(\phi + |Y| - 1)_{|Y|}},$$

where $(a)_b = a(a-1)\cdots(a-b+1)$.

Proof. Let $W = (x_1, \ldots, x_{2k+1})$ be a feasible walk on vertices $(v_1, \ldots, v_{k+1})$. If $|\{x_1, \ldots, x_{2k+1}\}| = 2t + 1$ and $|\{v_1, \ldots, v_{k+1}\}| = r$ then for some s,

$$\Pr\{W_k = W\} = \frac{1}{3n} \frac{1}{2^{r-1}} \frac{1}{3^s} \prod_{j=1}^{t+1} \frac{1}{3n - 2j + 1}.$$

11

Explanation: Let $W_k = (y_1, \ldots, y_{2k+1})$ be the random walk. Firstly, $\Pr\{y_1 = x_1\} = 1/3n$. We reveal the matching μ along with the walk. If i is odd and $\mu(y_i)$ is not previously revealed, then $\Pr\{y_{i+1} = x_{i+1}\} = 1/(3n - 2j + 1)$ where $|\{x_1, \ldots, x_i\}| = 2j - 1$. If $x_{i+1} \in \mathcal{P}(v)$ for some v that has not been visited by W_k previously, then $\Pr\{y_{i+2} = x_{i+2}\} = 1/2$. All other steps in the random walk either have probability 1 or $1/3$. Here s is the number of times the walk leaves a vertex that is in X_0 at the time of leaving. The value of s is uniquely determined by $\langle W \rangle$.

Now suppose $v_i, 1 < i < k+1$ is a vertex which is visited by W exactly once, so that $v_i \in Y$ and x_{2i}, x_{2i+1} are visited only once. Let v_j be another vertex, not necessarily in Y, and let (x_{2j-1}, x_{2j}) be edge visited only once by W. Form the walk

$$W' = (x_1, x_2, \ldots, x_{2i-1}, x_{2i+2}, \ldots, x_{2j-1}, x_{2i}, x_{2i+1}, x_{2j}, \ldots, x_{2k-1}).$$

Call this a *transposition of W*. Then, as the number of steps of the different types is unchanged,

$$\Pr\{W_k = W'\} = \frac{1}{3n}\frac{1}{2^{r-1}}\frac{1}{3^s}\prod_{j=1}^{t+1}\frac{1}{3n - 2j + 1} = \Pr\{W_k = W\}. \tag{11}$$

Now, if $W \sim \widehat{W}$ for two feasible walks, then $\widehat{W}$ can be obtained from W by a sequence of transpositions, and $\Pr\{W_k = \widehat{W}\} = \Pr\{W_k = W\}$. It is important to observe here that a transposition can move a vertex of Y onto a green bridge that was previously empty of vertices in Y.

The walk $\langle W \rangle$ contains ϕ green edges. We form any member of $[W]$ by distributing the vertices in Y onto the green edges, assigning an internal order to each resulting green bridge. Let $Y = \{y_1, y_2, \ldots, y_m\}$. We place y_1 on one of ϕ green edges of $\langle W \rangle$, breaking the edge into two edges. There are then $\phi + 1$ choices for the placement of y_2, and so on. This implies

$$|[W]| = \phi(\phi+1)(\phi+2)\ldots(\phi+|Y|-1).$$

$\square$

If $\langle W(t) \rangle$ contains ϕ green edges $(e_1, e_2, \ldots, e_\phi)$, we let $(K_1, \ldots, K_\phi)$, $K_i \geq 1$, denote the lengths of the corresponding paths in $W(t)$. We remark that we can sample the vector $(K_1, \ldots, K_\phi)$ by a Pólya urn process. Initially placing ϕ balls of distinct colors in an urn, we repeat the following $|Y(t)|$ times: draw a ball uniformly at random, replace it in the urn and add another ball of the same color. The sizes of the resulting color classes, including the initial balls, are distributed as $(K_1, \ldots, K_\phi)$.

6 Recurrences for $\mathbb{E}(X_i(t))$

We discuss the growth rate of $|X_i(t)|, i \geq 0$. In an abuse of notation, we will use $X_i(t), i \geq 0$ to denote both the set and its size. The context should dispel any possible ambiguity.

Let $\mathcal{H}(t)$ denote the history of the process up until time t. We have $X_3(0) = n, X_i(0) = 0, i = 0, 1, 2$ and

$$\mathbb{E}\left(X_3(t+1) \mid \mathcal{H}(t)\right) = X_3(t) - \frac{3X_3(t)}{3n - 2t + 1} \tag{12}$$

$$\mathbb{E}\left(X_2(t+1) \mid \mathcal{H}(t)\right) = 1_{t=0} - \frac{2X_2(t)}{3n - 2t + 1} \tag{13}$$

$$\mathbb{E}\left(X_1(t+1) \mid \mathcal{H}(t)\right) = X_1(t) - \frac{2X_1(t)}{3n - 2t + 1} + \frac{3X_3(t)}{3n - 2t + 1} + \frac{2X_2(t)}{3n - 2t + 1} \tag{14}$$

$$\mathbb{E}\left(X_0(t+1) \mid \mathcal{H}(t)\right) = n - (X_1(t) + X_2(t) + X_3(t)). \tag{15}$$

We remark that $X_2 \leq 1$ and is almost irrelevant to the ensuing analysis. As justification for the above equations, consider X_1. There is a probability $\frac{3X_3(t)}{3n-2t+1}$ that the newly paired point is associated with a vertex in $X_3(t)$. In which case, X_1 increases by one. Now a step involves visiting two members of $\mathcal{P}$ and if the newly paired point is associated with a vertex $v \in X_1(t)$ then except in one exceptional case, the walk will move along an already visited edge $\{v, w\}$ where $w \in X_1(t)$ and then the second visited point will be in $\mathcal{P}(w)$. The exceptional case is when $w \in X_2(t)$, which explains the last term in (14). The other equations are explained similarly.

Thus,

$$\mathbb{E}\left(X_3(t)\right) = n \prod_{i=1}^{t} \left(1 - \frac{3}{3n - 2i + 1}\right) = n \exp\left\{-\sum_{l=1}^{\infty} \frac{3^l}{l 2^l} \sum_{i=1}^{t} \frac{1}{\left(\frac{3n}{2} - i + \frac{1}{2}\right)^l}\right\}$$

Now

$$\int_{x=0}^{t-1} \frac{dx}{\left(\frac{3n}{2} - x + \frac{1}{2}\right)^l} \leq S_l = \sum_{i=1}^{t} \frac{1}{\left(\frac{3n}{2} - i + \frac{1}{2}\right)^l} \leq \int_{x=0}^{t} \frac{dx}{\left(\frac{3n}{2} - x - \frac{1}{2}\right)^l}$$

which implies that if $\omega_1 = 3n - 2t \to \infty$ then

$$S_1 = \log\left(\frac{3n}{3n - 2t}\right) + O\left(\frac{1}{3n - 2t}\right) \text{ and } \sum_{l=2}^{\infty} S_l = O(\omega_1^{-1}).$$

It follows that

$$\mathbb{E}\left(X_3(t)\right) = n \left(\frac{3n - 2t}{3n}\right)^{3/2} \left(1 + O\left(\frac{1}{3n - 2t}\right)\right). \tag{16}$$

$$\mathbb{E}\left(X_1(t)\right) = 3n - 2t - \mathbb{E}\left(3X_3(t) + 2X_2(t)\right) \tag{17}$$

$$= (3n - 2t)\left(1 - \left(1 - \frac{2t}{3n}\right)^{1/2}\right) + O\left(\frac{3n - 2t}{n}\right)^{1/2}. \tag{18}$$

We also have some concentration around these values, as described in the following lemma which is proved in Section 8.

Lemma 10. *Let $0 < \varepsilon < 2/3$ and let ω tend to infinity arbitrarily slowly. Let $\delta = \delta(t) = (3n - 2t)/3n$.*

(i) If $\omega^{-1} \geq \delta \geq \omega n^{-2/3}$ then

$$Pr\left\{|X_3(t) - n\delta^{3/2}| \geq \frac{n\delta^{3/2}}{\omega^{1/2}}\right\} = o(1),$$

and if $\delta \leq \omega^{-1} n^{-2/3}$ then $|X_3(t)| = 0$ with high probability.

(ii) Let $\delta_1 = \log^{-1/2} n$ and let $t_1 = (1 - \delta_1)\frac{3n}{2}$.

$$Pr\{\exists\, 3n/4 \leq t \leq t_1 : X_1(t) < \delta n\} = o(1).$$

(iii) Let $\delta_4 = n^{-1} \log^{11} n$.

$$Pr\{\exists t_1 \leq t \leq t_4 : |X_1(t) - 3n\delta| \geq \omega^{-1}\delta n\} = o(1).$$

(iv) Set $\delta_3 = n^{-2/3} \log^4 n$.

$$Pr\{\exists t_1 \leq t \leq t_3 : \Phi(t) < (\delta_0\delta)^{1/2} n\} = o(1).$$

where $\Phi(t)$ the number of green edges in $W(t)$.

7   Calculating the cover time

7.1   Early stages

With t_1 as in Lemma 10, we show that $\mathbb{E}(C(t_1)) = o(n \log n)$. Suppose $W(t) = (x_1, x_2, \ldots, x_{2k-1})$ for some t and $k \geq 1$. If $x_{2k-1} \in \mathcal{P}(\overline{X}(t))$ then $x_{2k} = \mu(x_{2k-1})$ is uniformly random inside $\mathcal{P}(\overline{X}(t))$, and since $C(t+1) = C(t) + 1$ in the event of $x_{2k} \in \mathcal{P}(X_2 \cup X_3)$, we have

$$\mathbb{E}(C(t+1) - C(t)) \leq 1 + \mathbb{E}(C(t+1) - C(t) \mid x_{2k} \in \mathcal{P}(X_1)) \Pr\{x_{2k} \in \mathcal{P}(X_1)\}, \qquad (19)$$

We use the following theorem of Ajtai, Komlós and Szemerédi [1] to bound the expected change when $x_{2k} \in \mathcal{P}(X_1)$.

Theorem 3. *Let $G = (V, E)$ be an r-regular graph on n vertices, and suppose that each of the eigenvalues of the adjacency matrix with the exception of the first eigenvalue are at most λ_G (in absolute value). Let Z be a set of cn vertices of G. Then for every ℓ, the number of walks of length ℓ in G which avoid Z does not exceed $(1 - c)n((1 - c)r + c\lambda_G)^\ell$.*

The set Z of Theorem 3 is fixed. In our case the exit vertex u of the red walk is chosen randomly from $X_1(t)$. This follows from the way the red walk constructs the graph in the configuration model. The subsequent walk now begins at vertex u and continues until it hits a vertex of $Y_u = X_1(t) \setminus \{u\}$ (or more precisely $Y_u \cup X_2(t)$). Because the exit vertex u is random, the set $B_u = Y_u \cup X_2(t) \cup X_3(t)$ differs for each possible exit vertex $u \in X_1(t)$. To apply Theorem 3, we split $X_1(t)$ into two disjoint sets A, A' of (almost) equal size. For

14

$u \in A$, instead of considering the number of steps needed to hit B_u, we can upper bound this by the number of steps needed to hit $B' = A' \cup X_2 \cup X_3$.

Let $Z(\ell)$ be a simple random walk of length ℓ starting from a uniformly chosen vertex of A. Thus $Z(\ell)$ could be any of $|A|3^\ell$ uniformly chosen random walks. Let $c = |B'|/n$. The probability p_ℓ that a randomly chosen walk of length ℓ starting from A has avoided B' is at most

$$p_\ell \le \frac{1}{(|X_1(t)|/2)3^\ell}(1-c)n(3(1-c) + c\lambda_G)^\ell \le \frac{2(1-c)n}{|X_1(t)|}((1-c) + c\lambda)^\ell,$$

where $\lambda \le .99$ (see Lemma 5) is the absolute value of the second largest eigenvalue of the transition matrix of Z. Thus

$$\mathbb{E}_A\left(H(B')\right) \le \sum_{\ell \ge 1} p_\ell \le \frac{2(1-c)n}{|X_1(t)|}\frac{1}{c(1-\lambda)}. \tag{20}$$

As $|B'| = |X_1|/2 + |X_3|$, we have

$$\mathbb{E}\left(C(t+1) - C(t) \mid x_{2k} \in \mathcal{P}(X_1(t))\right) = O\left(\frac{(n - |X_3|)n}{|X_1|(|X_1| + |X_3|)}\right). \tag{21}$$

Phase I: $t \le 3n/4$.

We use the bound $|X_3(t)| \ge n - t \ge n/4$ and the fact that

$$\Pr\{x_{2k} \in \mathcal{P}(X_1(t))\} = \frac{|X_1(t)|}{3n - 2t - 1} = O\left(\frac{|X_1(t)|}{n}\right).$$

Then (19) and (21) imply

$$\mathbb{E}\left(C(t+1) - C(t)\right) \le 1 + O\left(\frac{tn}{|X_1(t)|n} \cdot \frac{|X_1(t)|}{n}\right) = 1 + O\left(\frac{t}{n}\right). \tag{22}$$

Summing over $t \le 3n/4$ gives $\mathbb{E}\left(C(3n/4)\right) = O(n)$.

Phase II: $3n/4 \le t \le (1 - \delta_1)3n/2$.

It follows from Lemma 10 (ii) that with high probability, for all $3n/4 \le t \le (1 - \delta_1)\frac{3n}{2}$, we have $|X_1(t)| \ge (3n - 2t)/3$. In particular, $|X_1|(|X_1| + |X_3|) = \Omega((3n - 2t)^2)$, and by (21),

$$\mathbb{E}\left(C\left((1 - \delta_1)\frac{3n}{2}\right) - C\left(\frac{3n}{4}\right)\right) \le \sum_{t=\frac{3n}{4}}^{(1-\delta_1)\frac{3n}{2}} O\left[\frac{(n - |X_3|)n}{|X_1|(|X_1| + |X_3|)}\right] \tag{23}$$

$$= O\left[\sum_{t=\frac{3n}{4}}^{(1-\delta_1)\frac{3n}{2}} \frac{n^2}{(3n - 2t)^2}\right] \tag{24}$$

$$= O\left(\frac{n}{\delta_1}\right) \tag{25}$$

$$= o(n \log n). \tag{26}$$

15

7.2 Later Stages

We will now use Lemmas 8 and 9, together with Definition 1 and Lemma 10.

For t with $\delta \leq \log^{-1/2} n$ we set $\omega = \omega(t) = \log(-\log \delta)$ and define the events (with $\overline{X}(t) = X_1(t) \cup X_2(t) \cup X_3(t)$)

$$\mathcal{A}(t) = \{|X_1(t) - 3n\delta| = O(\omega^{-1}\delta n)\}, \tag{27}$$
$$\mathcal{B}(t) = \{\overline{X}(t) \text{ is a root set of order } \omega\}. \tag{28}$$

and set $\mathcal{E}(t) = \mathcal{A}(t) \cap \mathcal{B}(t)$. As a consequence of Lemma 8, equation (20) and the fact that $\mathbb{E}\left(\overline{X}(t)\right) \approx 3n - 2t$, we have

$$\mathbb{E}\left(C(t+1) - C(t)\right) = (3 \pm \varepsilon)\frac{n}{3n - 2t}\Pr\{\mathcal{E}(t)\} + O\left(\frac{n}{3n - 2t}\right)\Pr\left\{\overline{\mathcal{E}(t)}\right\} + O(\log n). \tag{29}$$

Here the $O(\log n)$ and ε terms account for the number of steps needed to take for the random walk Markov chain to mix to within variation distance ε of the stationary distribution, at which time we apply Lemma 8. Here we rely on $\lambda^*(t) \leq 0.99$. In the event of $\overline{\mathcal{E}(t)}$ we use the fact that $\overline{X}(t) = \Omega(3n - 2t)$, which follows from Lemma 10(ii) and the hitting time bound $\frac{1}{1-\lambda}\frac{n}{\overline{X}(t)}$ (see (10)) to conclude that the hitting time is $O(n/(3n - 2t))$.

Lemma 10 implies that $\mathcal{A}(t)$ occurs with high probability for any fixed $t \geq 3n(1 - \log^{-1/2} n)/2$ and we will argue in Sections 7.3 and 7.4 that $\mathcal{B}(t)$ also occurs with high probability. Lemma 3 will follow. Lemma 4 is proved in Section 7.5 and Lemma 1 follows.

7.3 Expansion via concentration

As discussed above, we are interested in showing that the event $\mathcal{E}(t)$ occurs with high probability.

Lemma 11. *Fix t and let $\delta = (3n - 2t)/n$. If $\delta_1 = \log^{-1/2} n \geq \delta \geq \delta_3 = n^{-2/3}\log^4 n$ then,*

$$Pr\{\mathcal{E}(t)\} = 1 - o(1).$$

Proof. Fix some t, δ in the given range. Expose $[W(t)]$. Lemma 10 shows that with high probability, $\Phi(t), |X_1(t)|$ satisfy

$$\Phi(t) \geq (\delta_0\delta)^{1/2}n, \tag{30}$$
$$|X_1(t)| = 3\delta n + O(\omega^{-1}\delta n). \tag{31}$$

As already remarked, this shows that $\Pr\{\mathcal{A}(t)\} = 1 - o(1)$. We next bound $|X_3(t)|$. To do this, we will only use the fact that it is dominated by $|X_1(t)|$ throughout this phase: as $\delta \geq n^{-1/2}\log^4 n$ and $|X_1(t)| \approx 3\delta n$, by Lemma 10.

$$|X_3(t)| = \delta^{3/2}n + O(\omega^{-1}\delta^{3/2}n) = o(\delta n) = o(|X_1(t)|)$$

16

with high probability. We can now show that $\overline{X}(t) = X_1(t) \cup X_2(t) \cup X_3(t)$ is a root set of order ω with high probability. Here ω is chosen to satisfy (34) below.

Let E_t denote the set of t edges discovered by the walk, and E_t^c the set of (random) edges yet to be discovered. The number of edges inside $\overline{X}(t)$ is given by

$$e(\overline{X}(t)) = |E_t^c| + |E(X_1 \cup X_2) \cap E_t| \tag{32}$$

where $|E_t^c| = (X_1 + 2X_2 + 3X_3)/2$, so

$$|E_t^c| = \frac{|X_1|}{2} + O(\delta_1^{1/2}) = \frac{|X_1|}{2} + O(\omega^{-3})$$

for $\omega^3 \ll \delta_0^{-1/2}$.

We bound the number of paths of length at most ω between vertices of X_1 on edges of E_t, showing that the number is $O(|X_1|/\omega^3)$. Note that such paths include $E(X_1) \cap E_t$, so that the bound implies $|E(X_1) \cap E_t| = O(|X_1|/\omega^3)$.

Let $u, v \in X_1$. Suppose u is placed on some green edge f_1. There are at most 3^ω green edges at distance at most ω from f_1, so as v is placed in a random green edge,

$$\Pr\{d(u,v) \le \omega\} = O\left(\frac{3^\omega}{\Phi}\right) = O\left(\frac{3^\omega}{n(\delta_0\delta)^{1/2}}\right).$$

So the expected number of pairs $u, v \in X_1$ at distance at most ω is bounded by

$$\sum_{u,v \in X_1} \Pr\{d(u,v) \le \omega\} = O\left(\frac{|X_1|^2 3^\omega}{n(\delta_0\delta)^{1/2}}\right) = O(n\delta_0^{-1/2}\delta^{3/2}3^\omega) = o(|X_1|/\omega^3), \tag{33}$$

if we choose

$$\omega^3 3^\omega \ll (\delta_0/\delta)^{1/2}. \tag{34}$$

With high probability the number of paths is $O(|X_1|/\omega^3)$ by the Markov inequality. This shows that $\overline{X}(t)$ is a root set of order ω with high probability. $\qquad\square$

7.4 Maintaining expansion without concentration

Lemma 12. *For any fixed t such that $\delta = \delta(t)$ satisfies $\delta_4 = n^{-1}\log^{11} n \le \delta \le \delta_3 = n^{-2/3}\log^4 n$,*

$$Pr\{\mathcal{E}(t)\} = 1 - o(1).$$

Proof. By Lemma 10, at time $t_3 = (1 - \delta_3)\frac{3n}{2}$ the sizes of X_1, X_3 and Φ satisfy the following with high probabilty,

$$\Phi(t_3) \ge n(\delta_0\delta_3)^{1/2}, \tag{35}$$
$$|X_1(t_3)| \approx 3n\delta_3, \tag{36}$$
$$|X_3(t_3)| \approx n\delta_3^{3/2} = O(\log^6 n). \tag{37}$$

17

For $v \in X_1$ let $N_\ell^0(v)$ denote the number of vertices of X_0 at distance ℓ from v, using only edges of E_t. Let $X_1' \subseteq X_1$ denote the set of vertices $v \in X_1$ with $|N_\ell^0(v)| = 2^\ell$ for all $\ell \le \omega$, and let $X_1'' = X_1 \setminus X_1'$. We have $|X_1''(t_3)| = O(n\delta_3^{3/2}3^\omega) = O(\log^7 n)$ with high probability from (33). By Lemma 10 iii we have $X_1(t) \ge (1 - o(1))\log^{11} n$ for $\delta \ge \delta_4$. So for $t_3 \le t \le t_4$ we have $|X_1''(t)| \le |X_1''(t_3)| = O((\log n)^{11}/\omega^3) = O(|X_1(t)|/\omega^3)$ w.h.p. This shows that $\overline{X}(t)$ is a root set of order ω. $\qquad \square$

7.5   The final edges: Proof of Lemma 4

Recall that $\delta_4 = n^{-1}\log^{11} n$ and $t_4 = (1 - \delta_4)\frac{3n}{2}$. So far we have shown that the edge cover time claimed by Lemma 1 holds for all $t_1 \le t \le t_4$. We now show that

$$\mathbb{E}\left(C\left(\frac{3n}{2}\right) - C(t_4)\right) = o(n \log n).$$

Fix $t_4 \le t < \frac{3n}{2}$. We bound the hitting time of $\overline{X}(t) = X_1(t) \cup X_2(t) \cup X_3(t)$, which has size $3n - 2t$. We contract $\overline{X}(t)$ into a single node x and apply Lemma 6, using the bound

$$Z_{xx} \le \sum_{t \ge 0} |P_x^{(t)}(x) - \pi_x| \le \sum_{t \ge 0} \lambda^t = \frac{1}{1 - \lambda} = O(1).$$

As the random walk Markov chain mixes to within ε total variation distance of π in $O(\log n)$ steps, it follows that

$$\mathbb{E}\left(C(t + 1) - C(t)\right) \le O(\log n) + \frac{1}{1 - \lambda}\frac{n}{3n - 2t}.$$

So, as $3n/2 - t_4 = O(\log^{11} n)$,

$$\mathbb{E}\left(C\left(\frac{3n}{2}\right) - C(t_4)\right) \le O(\log^{12} n) + \frac{1}{1 - \lambda}\sum_{t=t_4}^{\frac{3n}{2}-1}\frac{n}{3n - 2t} = O\left(n \log\left(\log^{11} n\right)\right) = o(n \log n).$$

7.6   Strengthening to "with high probability"

So far, all expectations are taken over the full probability space of random graphs and random walks, simultaneously generated. In particular, the expected cover time is the average cover time of all cubic multigraphs. In this section we prove that almost all cubic multigraphs have the same cover time.

Let $\mathcal{G}$ denote the set of 3-regular (multi)graphs. For $G \in \mathcal{G}$ and a random variable X, write

$$\mathbb{E}_G(X) = \mathbb{E}(X \mid G),$$

so that, as $G \in \mathcal{G}$ is chosen uniformly at random,

$$\mathbb{E}(X) = \frac{1}{|\mathcal{G}|}\sum_{G \in \mathcal{G}}\mathbb{E}_G(X).$$

We prove the following lemma. Define $\delta_2 = \log^{-2} n$ and $t_2 = (1 - \delta_2)\frac{3n}{2}$.

18

Lemma 13. *Let $t > t_2$. If $G \in \mathcal{G}$ is chosen uniformly at random, then with high probability,*

$$\mathbb{E}_G\left(C(t)\right) = \left(\frac{3}{2} \pm \varepsilon\right) n \log\left(\frac{3n}{3n - 2t + 1}\right) + o(n \log n).$$

Proof of Lemma 13. Fix some $t > t_2$. We define the following subsets of $\mathcal{G}$, with $\varepsilon > 0$ arbitrary and ω to be defined shortly,

$$\mathcal{H} = \left\{ G \in \mathcal{G} : \left| \mathbb{E}_G\left(C(t) - C(t_2)\right) - \sum_{s=t_2}^{t-1} \frac{3n}{3n - 2s} \right| \geq \varepsilon n \log\left(\frac{3n}{3n - 2t + 1}\right) \right\}, \tag{38}$$

$$\mathcal{J} = \left\{ G \in \mathcal{G} : \mathbb{E}_G\left(C(t_2)\right) \geq \frac{n \log n}{\omega} \right\}, \tag{39}$$

$$\mathcal{K} = \left\{ G \in \mathcal{G} : \max_{t_1 \leq t \leq t_4} \frac{|X_1(t) - (3n - 2t)|}{3n - 2t} \geq \varepsilon \right\}, \tag{40}$$

$$\mathcal{L} = \left\{ G \in \mathcal{G} : \lambda(G) > 0.99 \right\}. \tag{41}$$

We will show that asymptotically, the union of these four sets has size $o(|\mathcal{G}|)$. In particular, almost all $G \in \mathcal{G}$ are in $\overline{\mathcal{H} \cup \mathcal{J} \cup \mathcal{L}}$, which implies that w.h.p. $\mathbb{E}_G\left(C(t)\right)$ has the desired value.

It follows from Lemma 5 that $|\mathcal{L}|/|\mathcal{G}| = o(1)$. Also, Lemma 15 proved below shows that $|\mathcal{K}|/|\mathcal{G}| = o(1)$.

In Section 7.1 (see (26)) we show that $\mathbb{E}\left(C(t_1)\right) = o(n \log n)$. From Lemma 11 and (29),

$$\mathbb{E}\left(C(t_2)\right) = \mathbb{E}\left(C(t_1)\right) + O\left(\sum_{t=t_1}^{t_2} \frac{3n}{3n - 2t}\right) = o(n \log n) + O\left(n \log\left(\frac{\delta_1}{\delta_2}\right)\right) = o(n \log n).$$

So $\mathbb{E}\left(C(t_2)\right) \leq \omega^{-2} n \log n$ for some ω tending to infinity. Then

$$\frac{n \log n}{\omega^2} \geq \mathbb{E}\left(C(t_2)\right) = \frac{1}{|\mathcal{G}|} \sum_{G \in \mathcal{G}} \mathbb{E}_G\left(C(t_2)\right) \geq \frac{|\mathcal{J}|}{|\mathcal{G}|} \frac{n \log n}{\omega},$$

which implies that $|\mathcal{J}|/|\mathcal{G}| \leq \omega^{-1} = o(1)$. Define $\mathcal{G}' = \mathcal{G} \setminus (\mathcal{J} \cup \mathcal{K} \cup \mathcal{L})$, and $\mathcal{H}' = \mathcal{H} \cap \mathcal{G}'$. We will show that $|\mathcal{H}'|/|\mathcal{G}'| = o(1)$, which will imply that $|\mathcal{H}|/|\mathcal{G}| = o(1)$.

Write $\Delta(s) = C(s+1) - C(s)$. Define for $G \in \mathcal{G}'$,

$$T_G = \left\{ t_1 \leq s < t : \left| \mathbb{E}_G\left(\Delta(s)\right) - \frac{3n}{3n - 2s} \right| \geq \frac{\varepsilon n}{3n - 2s} \right\}.$$

Define, for some ω tending to infinity with n,

$$\mathcal{F}' = \left\{ G \in \mathcal{G}' : \sum_{s \in T_G} \frac{n}{3n - 2s} \geq \frac{1}{\omega} n \log\left(\frac{3n}{3n - 2t + 1}\right) \right\}.$$

If $G \in \mathcal{G}'$ then $\mathbb{E}_G\left(\Delta(s)\right) \leq (1 - \lambda(G))^{-1} n/|X_1| \leq 100n/(3n - 2s)$ for all s, and so if $G \in \mathcal{G}' \setminus \mathcal{F}'$,

$$\left| \mathbb{E}_G\left(C(t) - C(t_1)\right) - \sum_{s=t_1}^{t-1} \frac{3n}{3n - 2s} \right| \leq \varepsilon \sum_{s \notin T_G} \frac{n}{3n - 2s} + \sum_{s \in T_G} \left| \mathbb{E}_G\left(\Delta(s)\right) - \frac{3n}{3n - 2s} \right| \quad (42)$$

$$\leq \frac{1}{2} \varepsilon n \log\left(\frac{n}{3n - 2t + 1} \right) + 103 \sum_{s \in T_G} \frac{n}{3n - 2s} \quad (43)$$

$$\leq \varepsilon n \log\left(\frac{3n}{3n - 2t + 1} \right). \quad (44)$$

In particular, $\mathcal{H}' \subseteq \mathcal{F}'$, and it remains to argue that $|\mathcal{F}'|/|\mathcal{G}'| = o(1)$.

For $t_1 \leq s < t$ let $\mathcal{F}'_s \subseteq \mathcal{G}'$ denote the set of graphs G with $s \in T_G$. Since $\Pr\{\mathcal{E}(s)\} = 1 - o(1)$ (Lemmas 11 and 12), almost all $G \in \mathcal{G}'$ are such that $\Pr\{\mathcal{E}(s) \mid G\} = 1 - o(1)$. For such a G we have (as in (29) with ε replaced by $\varepsilon/2$)

$$\mathbb{E}_G\left(\Delta(s)\right) = \left(3 \pm \frac{\varepsilon}{2}\right) \frac{n}{3n - 2s} \Pr\{\mathcal{E}(s) \mid G\} + O\left(\frac{1}{1 - \lambda} \frac{n}{3n - 2s}\right) \Pr\{\overline{\mathcal{E}(s)} \mid G\} + O(\log n) \quad (45)$$

$$= (3 \pm \varepsilon) \frac{n}{3n - 2s}, \quad (46)$$

and so $s \notin T_G$. Note that we have used $s \geq t_2$ here, in order to eliminate the $O(\log n)$ term. This shows that $|\mathcal{F}'_s|/|\mathcal{G}'| = o(1)$ for each $t_1 \leq s < t$. Let ω tending to infinity be such that $|\mathcal{F}'_s|/|\mathcal{G}'| \leq \omega^{-2}$ for all s. So,

$$\frac{1}{|\mathcal{G}|} \sum_{G \in \mathcal{G}} \sum_{s \in T_G} \frac{n}{3n - 2s} = \sum_s \frac{|\mathcal{F}_s|}{|\mathcal{G}|} \frac{n}{3n - 2s} \leq \frac{3}{2\omega^2} n \log\left(\frac{3n}{3n - 2t + 1} \right). \quad (47)$$

But by definition of $\mathcal{F}'$,

$$\frac{1}{|\mathcal{G}|} \sum_{G \in \mathcal{G}} \sum_{s \in T_G} \frac{n}{3n - 2s} \geq \frac{|\mathcal{F}'|}{|\mathcal{G}|} \frac{1}{\omega} n \log\left(\frac{3n}{3n - 2t + 1} \right)$$

and we conclude that $|\mathcal{F}'|/|\mathcal{G}| \leq \frac{3}{2} \omega^{-1} = o(1)$. This finishes the proof. $\qquad \square$

7.7 The vertex cover time

Using Lemma 10 (i) we can express the partial vertex cover time in terms of the partial edge cover time.

Lemma 14. *Let $n - n \log^{-1} n \leq s \leq n$ be fixed. Then with high probability, G is such that*

$$\mathbb{E}_G\left(C_V(s)\right) = (1 \pm \varepsilon) n \log\left(\frac{n}{n - s + 1} \right). \quad (48)$$

Proof. As $X_3(t)$ is the set of undiscovered vertices at time $C(t) - 1$, we can write $C_V(s) = C(\tau_s)$, where

$$\tau_s = \min\{t : |X_3(t)| = n - s\}.$$

Define $\delta_s = (1 - (s-1)/n)^{2/3}$. Let $\tau_s^- = (1 - \delta_s\omega)\frac{3n}{2}$ for some $\omega = O(\log\log n)$ tending to infinity with n. By Lemma 10, with high probability

$$|X_3(\tau_s^-)| \approx n(\delta_s\omega)^{3/2} = n\left(1 - \frac{s-1}{n}\right)\omega^{3/2} \approx (n-s)\omega^{3/2} \gg n - s,$$

so $\tau_s > \tau_s^-$ with high probability. By a similar calculation, if $\tau_s^+ = (1 - \delta_s\omega^{-1})\frac{3n}{2}$ then $\tau_s < \tau_s^+$ with high probability. Lemma 10 implies that $t_2 < \tau_s \le \frac{3n}{2}$ w.h.p. and so Section 7.6 implies that w.h.p., G is such that $\mathbb{E}_G(C(\tau_s^\pm)) \approx \frac{3}{2}n\log(3n/(3n - 2\tau_s^\pm + 1))$. So,

$$\mathbb{E}_G(C_V(s)) = \mathbb{E}_G(\mathbb{E}_G(C(\tau_s) \mid \tau_s)) \tag{49}$$

$$\le \mathbb{E}_G\left(\left(\frac{3}{2} + \varepsilon\right) n\log\left(\frac{3n}{3n - 2\tau_s + 1}\right)\right) \tag{50}$$

$$\le \Pr\{t \le \tau_s^+\}\left(\frac{3}{2} + \varepsilon\right) n\log\left(\frac{3n}{3n - 2\tau_s^+ + 1}\right) + \Pr\{t > \tau_s^+\}\left(\frac{3}{2} + \varepsilon\right) n\log n \tag{51}$$

$$= \left(\frac{3}{2} + \varepsilon\right) n\log\left(\frac{3n}{3n - 2\tau_s^+ + 1}\right) + o(n\log n) \tag{52}$$

Now, as $3n - 2\tau_s^+ = 3n\delta_s\omega$,

$$\log\left(\frac{3n}{3n - 2\tau_s^+ + 1}\right) = (1 + o(1))\log\left(\frac{1}{\delta_s\omega}\right) = \left(\frac{2}{3} + o(1)\right)\log\left(\frac{n}{n - s + 1}\right). \tag{53}$$

The lower bound for $\mathbb{E}(C_V(s))$ is found similarly. $\square$

Substituting $s = n$ into (48) gives us (1).

8   Concentration: Proof of Lemma 10

In this section we prove bounds for $X_1(t), X_3(t), \Phi(t)$, collected in Lemma 10. The following lemma contains a proof of part (i).

Lemma 15. *Let $\delta_3 = n^{-2/3}\log^4 n$, setting $t_3 = (1 - \delta_3)\frac{3n}{2}$. Then*

$$Pr\left\{\exists 1 \le t \le t_3 : X_3(t) > \frac{5}{4}n\delta^{3/2}\right\} = o(1). \tag{54}$$

Let ω tend to infinity with n. Then for any $\omega^{-1} \ge \delta \ge \omega n^{-2/3}$,

$$Pr\left\{|X_3(t) - n\delta^{3/2}| > \frac{n\delta^{3/2}}{\omega^{1/2}}\right\} = o(1),$$

and for any $\delta \le \omega^{-1}n^{-2/3}$,

$$Pr\{X_3(t) > 0\} = o(1).$$

Proof. For any set U of vertices with $|U| = k$, using calculations similar to (16),

$$\Pr\{U \subseteq X_3(t)\} = \left(1 - \frac{k}{n}\right) \prod_{j=0}^{t-1} \left(1 - \frac{3k}{3n - 2j - 1}\right) \tag{55}$$

$$\leq \left(\frac{3n - 2t}{3n}\right)^{3k/2} = \delta^{3k/2}. \tag{56}$$

Here the $1 - k/n$ factor accounts for the probability that the starting point of the walk is in $\overline{U}$. For any new edge (x, y) that is added, U can only be entered if the endpoint y which is chosen uniformly at random from all $3n - 2j - 1$ available configuration points is in U. We have, using the notation $(x)_k = x(x-1)\ldots(x-k+1)$,

$$\mathbb{E}(X_3(t)) = n\delta^{3/2}\left(1 + O\left(\frac{1}{\delta n}\right)\right), \quad \text{directly from (16).} \tag{57}$$

$$\mathbb{E}((X_3(t))_2) \leq n(n-1)\delta^3, \quad \text{from (56) with } k = 2. \tag{58}$$

$$\mathbb{E}((X_3(t))_k) \leq (n)_k \delta^{3k/2} \leq n^k \delta^{3k/2} \quad \text{from (56) in general.} \tag{59}$$

Firstly, it follows from (57) and (58) that we have $\text{Var}(X_3(t)) = O(n\delta^{3/2}) = O(\mathbb{E}(X_3(t)))$, and so Chebyshev's inequality shows that for any $\delta \geq \omega n^{-2/3}$,

$$\Pr\left\{|X_3(t) - n\delta^{3/2}| > \frac{n\delta^{3/2}}{\omega^{1/2}}\right\} = O\left(\frac{\mathbb{E}(X_3(t))\,\omega}{n^2\delta^3}\right) = \left(\frac{\omega}{\omega^2}\right) = o(1).$$

Secondly, the Markov inequality and (57) shows that for $\delta \leq \omega^{-1}n^{-2/3}$,

$$\Pr\{X_3(t) \geq 1\} \leq (1 + o(1))n\delta^{3/2} = o(1).$$

As $|X_3(t)|$ takes nonnegative integer values we can for any real $z > 1$, express the binomial theorem as

$$\mathbb{E}\left(z^{X_3(t)}\right) = \sum_{k \geq 0} \frac{\mathbb{E}((X_3(t))_k)(z-1)^k}{k!} \leq \sum_{k \geq 0} \frac{n^k \delta^{3k/2}(z-1)^k}{k!} = \exp\left\{(z-1)n\delta^{3/2}\right\},$$

So for any positive $\theta = o(1)$, the moment generating function of $X_3(t)$ satisfies

$$\mathbb{E}\left(e^{\theta X_3(t)}\right) \leq \exp\left\{\theta n\delta^{3/2}(1 + o(1))\right\}.$$

Let $\delta \geq \delta_3 = n^{-2/3}\log^4 n$. Then with $\theta = 2/\log n$,

$$\Pr\left\{X_3(t) > \frac{5}{4}n\delta^{3/2}\right\} \leq \frac{\mathbb{E}\left(e^{\theta X_3(t)}\right)}{\exp\left\{\frac{5}{4}\theta n\delta^{3/2}\right\}} \leq \exp\left\{-\frac{1}{4}\theta n\delta^{3/2}(1 - o(1))\right\} = o(n^{-1}).$$

Summing over $t = 1, 2, \ldots, t_3$, it follows that $X_3(t) \leq \frac{5}{4}n\delta^{3/2}$ for all $\delta \geq n^{-2/3}\log^4 n$ w.h.p.

$\square$

Lemma 15 implies bounds for $X_1(t)$ via $X_1(t) = 3n - 2t - 1 - 2X_2(t) - 3X_3(t)$. Firstly, for $\theta > 0$ with $\theta = o(1)$, as $X_2(t) \leq 1$,

$$\mathbb{E}\left(e^{-\theta X_1(t)}\right) \leq e^{-\theta(3n-2t+2)}\mathbb{E}\left(e^{3\theta X_3(t)}\right) \leq \exp\left\{-3\theta n\delta + 3\theta n\delta^{3/2}(1 + o(1))\right\}. \tag{60}$$

Secondly, $X_3(t) \leq \frac{5}{4}n\delta^{3/2}$ holding for all $3n/4 \leq t \leq t_3$ w.h.p. implies that w.h.p.,

$$X_1(t) \geq 3n\delta\left(1 - \frac{5}{4}\delta^{1/2}\right) - 3, \quad \frac{3n}{4} \leq t \leq t_3. \tag{61}$$

For $n^{-1}\log^{11} n = \delta_4 \leq \delta \leq \delta_3 = n^{-2/3}\log^4 n$ we have $X_3(t) \leq \frac{5}{4}n\delta_3^{3/2} = \log^6 n$ w.h.p., so w.h.p.

$$X_1(t) = 3n - 2t - 2X_2(t) - X_3(t) = 3n\delta\left(1 - O\left(\frac{\log^6 n}{\delta n}\right)\right) = 3n\delta(1 - o(1)). \tag{62}$$

Equations (61) and (62) together imply Lemma 10(ii) and (iii), assuming that $\omega = o(\log n)$.

Finally, we prove a lower bound for the number of green edges $\Phi(t)$.

Lemma 16.
$$Pr\left\{\exists t_1 \leq t \leq t_3 : \Phi(t) < (\delta\delta_0)^{1/2}n\right\} = o(1).$$

Note that as $\delta_0 = 1/\log\log n$, this implies that $\Phi(t) \gg n\delta$ whenever $\delta_1 \geq \delta \geq \delta_3$.

Proof. Fix some $t_1 \leq t \leq t_3$, and define $\delta = \delta(t)$ by $t = (1 - \delta)\frac{3n}{2}$ as usual. Recall that $Y(t) = X_1(t) \setminus \{v_0\}$ where v_0 is the tail of the walk $W(t)$, as defined in Section 5. Define events

$$\mathcal{E}(t) = \left\{Y(t) \geq \frac{3n - 2t}{2}\right\},$$

and let $\mathbf{1}_t$ denote the indicator variable for $\mathcal{E}(t)$. Note that by (60) with $\theta = 1/\log n$,

$$\Pr\left\{\overline{\mathcal{E}(t)}\right\} = \Pr\left\{e^{-\theta Y(t)} > e^{-\theta\frac{3n-2t}{2}}\right\} \leq \exp\left\{-\frac{3}{2}\theta n\delta(1 + o(1))\right\} \leq \eta$$

for $\delta \geq \delta_3$, where we define $\eta = \exp\left\{-n^{1/3}\log^3 n\right\}$.

Claim 1. *If $0 < \theta \leq n^{-2/3}\log^2 n$, then there exists an $\varepsilon = o(\log^{-1} n)$ such that for $t_0 \leq t \leq t_3$,*

$$\mathbb{E}\left(e^{-\theta(\Phi(t+1)-\Phi(t))}\mathbf{1}_t \mid [W(t)]\right) \leq \exp\left\{\frac{\theta\Phi(t)}{3n - 2t}(1 + \varepsilon)\right\}\mathbf{1}_t.$$

We now show how Claim 1 is used to provide a lower bound for $\Phi(t)$, before proving the claim. Firstly, as each vertex of $Y(t)$ is incident to exactly two green edges, we have $\Phi(t_0) \geq Y(t_0)$. So for $\theta > 0$ with $\theta = o(1)$, we have by (60),

$$\mathbb{E}\left(e^{-\theta\Phi(t_0)}\mathbf{1}_{t_0}\right) \leq \mathbb{E}\left(e^{-\theta(X_1(t_0)-1)}\right) \leq \exp\left\{-3\theta n\delta_0(1 - o(1))\right\}. \tag{63}$$

Set
$$f_t(\theta) = \mathbb{E}\left(e^{-\theta\Phi(t)}\mathbf{1}_t\right).$$

As $\theta, \Phi(t) > 0$, we have $e^{-\theta\Phi(t)} \leq 1$ and

$$f_t(\theta) = \mathbb{E}\left(e^{-\theta\Phi(t)}\mathbf{1}_{t-1}\right) + \mathbb{E}\left(e^{-\theta\Phi(t)}(\mathbf{1}_t - \mathbf{1}_{t-1})\right) \leq \mathbb{E}\left(e^{-\theta\Phi(t)}\mathbf{1}_{t-1}\right) + \Pr\{\mathcal{E}(t) \setminus \mathcal{E}(t-1)\}.$$

Note that $\Pr\{\mathcal{E}(t) \setminus \mathcal{E}(t-1)\} \leq \Pr\left\{\overline{\mathcal{E}(t-1)}\right\} \leq \eta$. Claim 1 implies that

$$f_{t+1}(\theta) = \mathbb{E}\left(e^{-\theta\Phi(t+1)}\mathbf{1}_t\right) + \eta \tag{64}$$

$$= \mathbb{E}\left(\mathbb{E}\left(e^{-\theta\Phi(t)}e^{-\theta(\Phi(t+1)-\Phi(t))}\mathbf{1}_t \mid [W(t)]\right)\right) + \eta \tag{65}$$

$$\leq \mathbb{E}\left(e^{-\theta\Phi(t)}\mathbb{E}\left(e^{-\theta(\Phi(t+1)-\Phi(t))}\mathbf{1}_t \mid [W(t)]\right)\right) + \eta \tag{66}$$

$$\leq \mathbb{E}\left(\exp\left\{-\theta\Phi(t) + \frac{\theta\Phi(t)}{3n-2t}(1+\varepsilon)\right\}\mathbf{1}_t\right) + \eta \tag{67}$$

$$= f_t\left(\theta - \frac{\theta(1+\varepsilon)}{3n-2t}\right) + \eta. \tag{68}$$

From the bound (63) it follows by induction that for $t > t_0$,

$$f_t(\theta) \leq \exp\left\{-3\theta n\delta_0(1-o(1))\prod_{s=t_0}^{t-1}\left(1 - \frac{1+\varepsilon}{3n-2s}\right)\right\} + (t-t_0)\eta.$$

With $\delta_0 = 1/\log\log n, \delta \leq \delta_1 = \log^{-1/2} n$ and $\varepsilon = o(\log^{-1} n)$, we have by calculations similar to those leading up to (16),

$$\prod_{s=t_0}^{t-1}\left(1 - \frac{1+\varepsilon}{3n-2s}\right) \approx \exp\left\{-\sum_{s=t_0}^{t-1}\frac{1}{3n-2s} - o\left(\frac{1}{\log n}\sum_{s=t_0}^{t-1}\frac{1}{3n-2s}\right)\right\} \approx \left(\frac{3n-2t}{3n-2t_0}\right)^{1/2},$$

as $\sum_{s=t_0}^{t-1} 1/(3n-2s) = O(\log n)$. This implies that for $\delta_3 \leq \delta \leq \delta_1$,

$$3n\delta_0(1-o(1))\prod_{s=t_0}^{t-1}\left(1 - \frac{1+\varepsilon}{3n-2s}\right) = 3n\delta_0\left(\frac{3n-2t}{3n-2t_0}\right)^{1/2}(1+o(1)) \tag{69}$$

$$\geq 2n(\delta\delta_0)^{1/2}, \tag{70}$$

so

$$f_t(\theta) \leq \exp\left\{-2\theta n(\delta\delta_0)^{1/2}\right\} + (t-t_0)\eta \tag{71}$$

Now, if

$$L(t) = n(\delta\delta_0)^{1/2}$$

then

$$\Pr\{\Phi(t) < L(t)\} \leq \Pr\{\Phi(t) < L(t), \mathcal{E}(t)\} + \Pr\left\{\overline{\mathcal{E}(t)}\right\} \tag{72}$$

and for $\theta > 0$, the bound $\mathbf{1}_{\{X>a\}} \le X/a$ for $X, a > 0$ implies

$$\Pr\{\Phi(t) < L(t), \mathcal{E}(t)\} = \mathbb{E}\left(\mathbf{1}_{\left\{e^{-\theta\Phi(t)} > e^{-\theta L(t)}\right\}} \mathbf{1}_t\right) \le \frac{\mathbb{E}\left(e^{-\theta\Phi(t)}\mathbf{1}_t\right)}{e^{-\theta L(t)}}. \tag{73}$$

Then for $\delta \le \delta_1 = \log^{-1/2} n$ and $\theta = n^{-2/3} \log^2 n$,

$$e^{\theta L(t)}\eta \le \exp\left\{\left(n^{-2/3}\log^2 n\right)\left(n(\delta_1\delta_0)^{1/2}\right) - n^{1/3}\log^3 n\right\} \tag{74}$$

$$\le \exp\left\{n^{1/3}\left(\frac{\log^{2-1/4} n}{\sqrt{\log\log n}} - \log^3 n\right)\right\} = o(n^{-2}), \tag{75}$$

and as $f_t(\theta) \le e^{-2\theta L(t)} + \frac{3n}{2}\eta$ by (71),

$$\Pr\{\Phi(t) < L(t)\} \le e^{\theta L(t)} f_t(\theta) + \eta \le e^{-\theta n(\delta\delta_0)^{1/2}} + o(n^{-1}).$$

We conclude that if $\theta = n^{-2/3} \log^2 n$,

$$\Pr\{\exists t_1 \le t \le t_3 : \Phi(t) < L(t)\} \le o(1) + \sum_{t=t_1}^{t_3}\left(\exp\left\{-\theta n(\delta\delta_0)^{1/2}\right\} + o(n^{-1})\right) \tag{76}$$

$$\le o(1) + O\left(n\exp\left\{-\frac{\log^{2+4/2} n}{\sqrt{\log\log n}}\right\}\right) \tag{77}$$

$$= o(1). \tag{78}$$

It remains to prove Claim 1.

Proof of Claim 1. We are interested in the distribution of $\Phi(t+1) - \Phi(t)$, conditioning on the contracted walk $[W(t)]$. Write $[W(t)] = (\langle W(t)\rangle, Y(t))$, where $Y(t) \subseteq X_1(t)$ is the set of vertices visited exactly once by $W(t)$, as defined in Section 5. Write $\langle W(t)\rangle = (x'_1, x'_2, \ldots, x'_{2s-1})$.

Reveal $x'_{2s} = \mu(x'_{2s-1})$. If $x'_{2s} \in \mathcal{P}(X_3(t) \cup X_2(t))$ then one green edge is added, and no other green edges are visited before the next edge is to be added, so $\Phi(t+1) = \Phi(t) + 1$.

On the other hand, if $x'_{2s} \in \mathcal{P}(X_1(t))$, then the walk will proceed until it finds another configuration point of $P(t)$. We are considering a random walk on $\langle W(t)\rangle$ starting at x'_{2s}. Let v_0 denote the tail vertex of the walk. Then $\langle W(t)\rangle$ induces a graph $G(t)$ on $X_0 \cup \{v_0\}$ with blue and green edges, in which each vertex of X_0 has degree 3 while v_0 has degree $1, 2$ or 3. Each time the walk on $G(t)$ traverses a green edge, we reveal the length of the corresponding green bridge in $W(t)$. The walk ends either when (i) a green edge is traversed and revealed to contain a vertex of $Y(t)$, or (ii) the vertex v_0 is reached, assuming $v_0 \in X_1 \cup X_2$.

Suppose there are $\Phi(t)$ green edges in $W(t)$, and $\phi = \Phi(t) - Y(t)$ green edges in $[W(t)]$. Let $e_1, \ldots, e_\phi$ denote the green edges of $[W(t)]$. By Lemma 9, the number of edges in the corresponding green bridges in $W(t)$ is a vector $(K_1, \ldots, K_\phi)$, uniformly drawn from all vectors with $K_i \ge 1$ and $\sum K_i = \Phi$. For $\ell \in \{1, \ldots, \phi\}$ we have

$$\Pr\{K_i = 1 \text{ for } i = 1, 2, \ldots, \ell\} = \prod_{i=1}^{\ell} \frac{\binom{\Phi-i-1}{\phi-i-1}}{\binom{\Phi-i}{\phi-i}} = \prod_{i=1}^{\ell} \frac{\phi-i}{\Phi-i} = \prod_{i=1}^{\ell}\left(1 - \frac{Y(t)}{\Phi(t)-i}\right) \le \left(1 - \frac{Y(t)}{\Phi(t)}\right)^{\ell}.$$

This shows that the number of green edges traversed before finding v_0 or a vertex of $Y(t)$ is stochastically dominated by a geometric random variable with success probability $Y(t)/\Phi(t)$.

This shows that in distribution, conditioning on the walk $[W(t)]$,

$$\Phi(t+1) - \Phi(t) \stackrel{d}{=} 1 - B\left(\frac{X_1(t)}{3n - 2t - 1}\right) R_t$$

where $B(p)$ denotes a Bernoulli random variable with success probability p, R_t is stochastically dominated by a geometric random variable with success probability $Y(t)/\Phi(t)$, and the two random variables in the right-hand side are independent. We have

$$\mathbb{E}\left(e^{-\theta(\Phi(t+1)-\Phi(t))} \mid [W(t)]\right) = e^{-\theta}\left(1 - \frac{X_1(t)}{3n - 2t - 1} + \frac{X_1(t)}{3n - 2t - 1}\mathbb{E}\left(e^{\theta R_t} \mid [W(t)]\right)\right). \quad (79)$$

As $x \mapsto e^{\theta x}$ is increasing, we can couple R_t to a geometric random variable Z_t with success probability $Y(t)/\Phi(t)$ so that

$$\mathbb{E}\left(e^{\theta R_t} \mid [W(t)]\right) \leq \mathbb{E}\left(e^{\theta Z_t} \mid [W(t)]\right).$$

As $[W(t)]$ is such that $\mathcal{E}(t)$ holds, i.e. $Y(t) \geq (3n - 2t)/2 = \Omega(n^{1/3}\log^4 n)$,

$$|(1 - e^{-\theta})\mathbb{E}(Z_t)| \leq (\theta + O(\theta^2))\frac{\Phi(t)}{Y(t)} \leq \frac{\log^2 n}{n^{2/3}}\frac{3n/2}{\Omega(n^{1/3}\log^4 n)} = o\left(\frac{1}{\log n}\right).$$

So, as Z_t is geometrically distributed,

$$\mathbb{E}\left(e^{\theta Z_t}\right) = \frac{\frac{Y(t)}{\Phi(t)}e^{\theta}}{1 - \left(1 - \frac{Y(t)}{\Phi(t)}\right)e^{\theta}} = \frac{1}{1 - \frac{\Phi(t)}{Y(t)}(1 - e^{-\theta})} = 1 - \frac{\Phi(t)}{Y(t)}(1 - e^{-\theta}) + O\left(\frac{\Phi(t)^2}{Y(t)^2}(1 - e^{-\theta})^2\right)$$

$$(80)$$

$$= 1 + \frac{\theta\Phi(t)}{Y(t)} + O\left(\frac{\theta^2\Phi(t)^2}{Y(t)^2}\right) \quad (81)$$

$$\leq 1 + \frac{\theta\Phi(t)}{Y(t)}(1 + \varepsilon) \quad (82)$$

for some $\varepsilon = o(\log^{-1} n)$. So if $[W(t)]$ is a class of walks with $Y(t) \geq (3n - 2t)/2$ then

$$\mathbb{E}\left(e^{-\theta(\Phi(t+1)-\Phi(t))}\mathbf{1}_t \mid [W(t)]\right) \leq e^{-\theta}\left(1 - \frac{X_1(t)}{3n - 2t - 1} + \frac{X_1(t)}{3n - 2t - 1}\left(1 + \frac{\theta\Phi(t)}{Y(t)} + O\left(\frac{\theta^2\Phi(t)^2}{Y(t)^2}\right)\right)\right)$$

$$(83)$$

$$\leq 1 + \frac{\theta\Phi(t)}{3n - 2t} + O\left(\frac{\theta^2\Phi(t)^2}{(3n - 2t)^2}\right) \quad (84)$$

$$\leq \exp\left\{\frac{\theta\Phi(t)}{3n - 2t}(1 + \varepsilon)\right\} \quad (85)$$

where $\varepsilon = o(\log^{-1} n)$. $\qquad\square$

This finishes the proof of Lemma 16. $\qquad\square$

References

[1] M. Ajtai, J. Komlós and E. Szemerédi, Deterministic simulation in LOGSPACE, *Proceedings of STOC 1987*, 132-140.

[2] D. Aldous and J. Fill, *Reversible Markov Chains and Random Walks on Graphs*, http://stat-www.berkeley.edu/pub/users/aldous/RWG/book.html

[3] N. Alon, I. Benjamini, E. Lubetzky, and S. Sodin. Non-backtracking random walks mix faster, *Communications in Contemporary Mathematics*, 9 (2007) 585–603.

[4] P. Berenbrink, C. Cooper and T. Friedetzky, Random walks which prefer unvisited edges: exploring high girth even degree expanders in linear time, *Random Structures and Algorithms* **46(1)** (2013), pages 36–54.

[5] C. Cooper and A. M. Frieze, The cover time of random regular graphs, *SIAM Journal on Discrete Mathematics*, 18 (2005) 728-740.

[6] C. Cooper and A. Frieze, Vacant sets and vacant nets: Component structures induced by a random walk, *SIAM Journal on Discrete Mathematics* (2016)

[7] C. Cooper, A.M. Frieze and S. Petti, The cover time of a biased random walk on $G_{n,p}$, Proceedings of ANALCO 2018.

[8] J. Friedman, A proof of Alon's second eigenvalue conjecture and related problems, Memoirs of the American Mathematical Society, 2008.

[9] A.M. Frieze and M.Karoński, Introduction to Random Graphs, Cambridge University Press, 2015.

[10] M.R. Jerrum and A. Sinclair, Approximate Counting, Uniform Generation and Rapidly Mixing Markov Chains, Information & Computation 82 (1989) 93-133.

[11] T. Orenshtein and I. Shinkar. Greedy random walk. *Combinatorics, Probability and Computing*, 23, (2014), 269-289.