The Cover Time of a Biased Random Walk on a Random Cubic Graph

Colin Cooper[1]
Department of Informatics, King's College, University of London, London WC2R 2LS, UK
colin.cooper@kcl.ac.uk

Alan Frieze[2]
Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA 15213, USA
alan@random.math.cmu.edu

Tony Johansson[3]
Department of Mathematics, Uppsala University, 751 05 Uppsala, Sweden
tony.johansson@math.uu.se

Abstract

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

2012 ACM Subject Classification Theory of computation → Random walks and Markov chains

Keywords and phrases Random walk, random regular graph, cover time

Digital Object Identifier 10.4230/LIPIcs.AofA.2018.16

Related Version https://arxiv.org/abs/1801.00760

1 Introduction

Our aim in this paper is to analyse a variation on the simple random walk that may tend to speed up the cover time of a connected graph. A simple random walk on a graph is a walk which repeatedly moves from its currently occupied vertex v to one of its neighbours, chosen uniformly at random. The vertex cover time $T^V_{\text{COV}}(G)$ of a simple random walk on a graph G is the expected number of steps needed to visit each vertex of G, defined as the maximum over all starting vertices. Feige [9, 10] showed that for any graph G on n vertices,

$$(1 - o(1))n \log n \leq T^V_{\text{COV}}(G) \leq (1 + o(1))\frac{4}{27}n^3.$$

When G is chosen uniformly at random from the set of d-regular graphs, Cooper and Frieze [6] showed that w.h.p.[4] G is such that $T^V_{\text{COV}}(G)$ is asymptotically equal to $\frac{d(d-1)}{2(d-2)} n \log n$.

In recent years, variations of the simple random walk have been introduced with the aim of achieving faster cover times. In this paper we do this by choosing to walk along unvisited edges whenever possible. This variation is just one of several possible approaches which include non-backtracking walks, see Alon, Benjamini, Lubetzky and Sodin [3], or walks that

---

[1] Research supported in part by EPSRC grants EP/J006300/1 and EP/M005038/1.
[2] Research supported in part by NSF grant DMS0753472.
[3] Research supported in part by the Knut and Alice Wallenberg Foundation.
[4] An event $\mathcal{E}$ is said to hold *with high probability* (w.h.p.) if $\Pr\{\mathcal{E}\} \to 1$ as $n \to \infty$.

29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2018).
Editors: James Allen Fill and Mark Daniel Ward; Article No. 16; pp. 16:1–16:12
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

are biased toward low degree vertices, see Cooper, Frieze and Petti [8], or any number of other ideas.

The papers [4], [13] describe the random walk model considered here, 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.[5] and makes 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 [13], a *greedy random walk*. We use the name *biased random walk* for the same process. For random d-regular graphs where $d = 2k$ (d even), it was shown in [4] that the biased random walk 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 factor comes from the fact that cycles of length at most ω exist w.h.p. In [7], the constant for the vertex cover time was shown to be $d/2$.

▶ **Theorem 1.** *Let $d \geq 4$ be even and suppose G is chosen u.a.r. from the set of d-regular graphs. W.h.p., G is such that the vertex cover time of the biased random walk is*[6] *$T^V_{cov}(G) \approx dn/2$.*

This is faster than any of the other random graph models mentioned here by a factor of $\log n$, and the biased random walk generally performs well on even-degree graphs. Orenshtein and Shinkar [13, Lemma 2.9] showed that in an even-degree graph, the biased random walk has cover time at most that of the simple random walk plus the number of edges in the graph. Briefly, this is because there are at most two vertices incident to an odd number of unvisited edges at any time. In the random setting this means that the most likely scenario is that traversing an unvisited (random) edge will bring the walk to a vertex incident to at least one more unvisited edge, and the walk will find a large number of unvisited edges in succession. This is no longer true in odd-degree graphs. The paper [4] included experimental data for the performance of red-blue walks on odd degree regular graphs. Namely, for $d = 3$ the vertex cover time is $\Theta(n \log n)$ and decreases rapidly with increasing d.

Random walks have applications in networks where each vertex only has local information, e.g. each vertex knows only of its immediate neighbours. For example, random walks provide efficient routing algorithms in Wireless Sensor Networks [15]. The vertex cover time measures the expected number of steps needed to spread information to each vertex of the network. A drawback of biased random walks in general applications is that it requires $O(|E|)$ additional memory usage, but in networks with independently acting agents, the additional memory for each agent is $O(\Delta)$ where Δ denotes the maximum degree of the network.

1.1 Our results

Let $G = (V, E)$ be a connected cubic (i.e. 3-regular) (multi)graph on an even number n of vertices. Consider the following random walk process, called a *biased random walk*. 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 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.

---

[5] We use u.a.r. for *uniformly at random*.
[6] We say that $a_n \approx b_n$ if $\lim a_n/b_n = 1$.

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. Note that a cubic graph on n vertices contains exactly $3n/2$ edges.

▶ **Theorem 2.** *Let s, t be 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 cubic graphs on n vertices. Then w.h.p., 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 $\frac{3}{2} n \log n$ w.h.p., respectively. The same statement is true with the word "graphs" replaced by "configuration multigraphs" (defined in Section 3). Thus, taking $s = n$ and $t = 3n/2$ we have the following corollary.

▶ **Corollary 3.** *Suppose G is chosen uniformly at random from the set of cubic graphs on n vertices. W.h.p., G is such that the vertex cover time $T^V_{COV}(G)$ of G is asymptotically equal to $n \log n$ and the edge cover time $T^E_{COV}(G)$ is asymptotically equal to $\frac{3}{2} n \log n$.*

Cooper and Frieze [6] showed that w.h.p. the vertex cover time for a simple random walk on a random d-regular graph on n vertices is asymptotically equal to $\frac{d-1}{d-2} n \log n$. The argument there also shows that the edge cover time of a random d-regular graph on n vertices is asymptotically equal to $\frac{d(d-1)}{2(d-2)} n \log n$. For $d = 3$ 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-backtracking random walk, Cooper and Frieze [7] show that the vertex and edge cover times are asymptotically $n \log n$ and $\frac{3}{2} n \log n$ respectively. Interestingly, these values coincide with the results in Corollary 3.

1.2 Outlook

Our proof relies on the fact that the set of vertices incident to exactly one unvisited edge coincides with the set of vertices visited exactly once by the biased random walk, modulo the head and tail of the walk. This is no longer true when $d \ge 5$, and additional analysis would be required to extend the method to larger degrees. We expect the walk to behave similarly for higher degrees and conjecture that Corollary 3 generalizes to $T^V_{\text{COV}}(G_d) \approx \frac{1}{d-2} n \log n$ and $T^E_{\text{COV}}(G_d) \approx \frac{d}{2(d-2)} n \log n$ for the random d-regular graph G_d, for any odd $d \ge 3$.

For fixed graphs, the behaviour of the greedy random walk is not well understood. See [13] for a list of open problems, including questions regarding transience and recurrence on infinite lattices.

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. See Section 3 for more details on the configuration model.

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 $\mu(x_{2k+1})$ 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 cubic graph G chosen u.a.r. is connected w.h.p. (this follows from Lemma 8 (i) below) and we will implicitly condition on this in what follows. The bulk of the paper will be spent proving the following lemma.

▶ **Lemma 4.** *For any fixed $\varepsilon > 0$ and $(1 - \log^{-2} n)\frac{3n}{2} \le t \le \frac{3n}{2}$,*

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

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

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

Expectations in Lemma 4 are taken over the full probability space. In particular, if $\mathcal{G}$ denotes the set of graphs,

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

We can strengthen Lemma 4 to stating that almost every G satisfies $\mathbb{E}_G(C(t)) \approx \mathbb{E}(C(t))$, and similarly for $C_V(s)$ (proof omitted in this extended abstract). Theorem 2 will then follow.

An essential part of the proof of Lemma 4 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, $t = 1, 2, ..., 3n/2$ (note that the graph contains exactly $3n/2$ edges). Ignoring in this extended abstract the set $X_2(t)$, which can only contain the tail vertex, the recurrences are

$$\mathbb{E}(X_3(t+1) \mid W(t)) = X_3(t) - \frac{3X_3(t)}{3n-2t+1}, \tag{9}$$

$$\mathbb{E}(X_1(t+1) \mid W(t)) = X_1(t) - \frac{2X_1(t)}{3n-2t+1} + \frac{3X_3(t)}{3n-2t+1}, \tag{10}$$

and we have $X_0(t) = n - X_1(t) - X_2(t) - X_3(t)$. These recurrences suggest that at time $t = (1-\delta)\frac{3n}{2}$ with $\delta = o(1)$ we have $X_1(t) \approx 3n\delta$ and $X_3(t) \approx n\delta^{3/2}$, and this is proven in the full paper version.

We will argue that for most of the process, it takes approximately $3n/(3n - 2t + 1)$ 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 3.

Given that $X_3(t) \approx n\delta^{3/2}$, 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 3. In this extended abstract we omit further details in calculating the vertex cover time.

We separate the proof of Lemma 4 into phases. Define

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

and set

$$t_i = (1-\delta_i)\frac{3n}{2} \text{ for } i = 0, 1, 2, 3, 4. \tag{11}$$

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

▶ **Lemma 5.** *Let* $\delta_1 = \log^{-1/2} n$ *and* $t_1 = (1-\delta_1)\frac{3n}{2}$. *Then*

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

Between times t_1 and t_4 we bound the time taken between discovering new edges. The proof, in Section 6, will be split into the ranges $t_1 \le t \le t_3$ and $t_3 \le t \le t_4$.

▶ **Lemma 6.** *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 7.** *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 4 follows from Lemmas 5, 6 and 7.

3 Structural properties of random cubic graphs

The random cubic graph is chosen according to the *configuration model*, introduced by Bollobás [5]. Each vertex $v \in [n]$ is associated with a set $\mathcal{P}(v)$ of 3 *configuration points*, and we let $\mathcal{P} = \cup_v \mathcal{P}(v)$. We choose u.a.r. a perfect matching μ of the points in $\mathcal{P}$. Each μ induces a multigraph G on $[n]$ in which u is adjacent to v if and only if $\mu(x) \in \mathcal{P}(v)$ for some $x \in \mathcal{P}(u)$, allowing parallel edges and self-loops. Here we collect some properties of random cubic graphs, chosen according to the configuration model. Any simple cubic graph is equally likely to be chosen under this model.

▶ **Lemma 8.** *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 w.h.p.,*

(i) *In absolute value, the second largest eigenvalue of the transition matrix for a simple random walk on G is at most* 0.99*.*
(ii) *G contains at most $\omega 3^\omega$ cycles of length at most ω,*
(iii) *The probability that G is simple is $\Omega(1)$.*

Friedman [11] showed that for any $\varepsilon > 0$, the second largest eigenvalue of the transition matrix is at most $2\sqrt{2}/3 + \varepsilon$ w.h.p., which gives (i). Property (ii) follows from the Markov inequality, given that the expected number of cycles of length $k \leq \omega$ can be bounded by $O(3^k)$. For the proof of (iii) see Frieze and Karoński [12], Theorem 10.3. Note that (iii) implies that any property which holds w.h.p. for a configuration multigraph chosen u.a.r., also holds w.h.p. for a simple cubic graph chosen u.a.r.

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 if Γ is a graph obtained from G by contracting a set of vertices, retaining all edges, then $\lambda(\Gamma) \leq \lambda(G)$, see [2, Corollary 3.27]. 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, as the only possible probability measure on a singleton is the trivial measure.

4 Hitting times for simple random walks

We are interested in calculating $\mathbb{E}\,(C(t+1)-C(t))$, i.e. the expected 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 $\overline{X}$. We will be able to calculate the hitting time as a consequence of $\overline{X}$ having a special structure as in the following definition.

▶ **Definition 9.** 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 contain an almost-perfect matching, and most of whose vertices are otherwise separated by a large distance. We can calculate the expected hitting time for such sets.

▶ **Lemma 10.** *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 $\omega 3^\omega$ cycles of length at most ω. If S is a root set of order ω, then the expected hitting time of S for a simple random walk starting at a uniformly chosen vertex is*

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

5 The structure of $\overline{X}$

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 $\overline{X}=X_1\cup X_2\cup X_3$, 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 between the two endpoints of the bridge, 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 1. Thus the only way that W,W' differ is as to where the vertices in Y are placed on the green bridges.

▶ **Lemma 11.** *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 $(a)_b=a(a-1)\cdots(a-b+1)$.

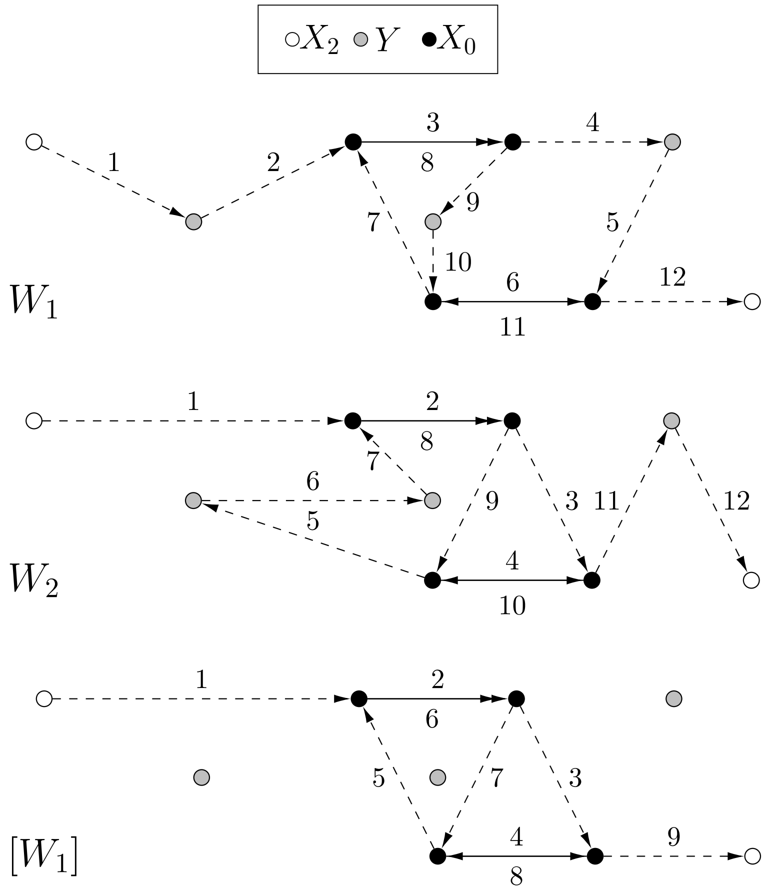

Figure 1 Two equivalent walks, and a visual representation of their equivalence class. Numbers represent order of traversal. Unvisited edges and vertices are not displayed, and edges visited exactly once are dashed. Lemma 11 shows that the walks are equiprobable.

We can now view the biased random walk as a walk on the equivalence class $[W(t)]$. Any time a green edge in $[W(t)]$ is visited, the probability that the edge corresponds to a nontrivial path in a randomly chosen $W(t) \in [W(t)]$ is about $X_1(t)/\Phi(t)$, where $\Phi(t)$ denotes the number of green edges in $W(t)$. This provides a precise recursion for $\mathbb{E}(\Phi(t))$ similar to those for $X_1(t), X_3(t)$, which we use to prove the following. Recall $\delta_0 = 1/\log\log n$. W.h.p.,

$$|X_1(t)| \sim 3n\delta \quad \text{when } \delta \le \delta_1, \tag{12}$$

$$|X_3(t)| \sim n\delta^{3/2} \quad \text{when } \delta \le \delta_1, \tag{13}$$

$$\Phi(t) \ge n(\delta_0\delta)^{1/2} \quad \text{when } \delta_3 \le \delta \le \delta_1. \tag{14}$$

Suppose $\delta_3 \le \delta \le \delta_1$. As $X_1(t) = o(\Phi(t))$, when $W(t) \in [W(t)]$ is chosen uniformly at random, the vertices of $X_1(t)$ are sprinkled into the much larger set of green edges, and are expected to be spread far apart. This will imply that $X_1(t)$ is a root set of order ω, and as $X_1(t)$ makes up almost all of $\overline{X}(t)$ by (12), the latter is also a root set of order ω. When $\delta \le n^{-2/3}$, the same technique can be applied with a little more work.

6 Calculating the cover time

6.1 Early stages

With $t_1 = (1 - \log^{-1/2} n)\frac{3n}{2}$, 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

$$\mathbb{E}(C(t+1) - C(t)) \leq 1 + \mathbb{E}(C(t+1) - C(t) \mid x_{2k} \in \mathcal{P}(X_1)) \Pr\{x_{2k} \in \mathcal{P}(X_1)\}, \tag{15}$$

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 12.** *Let $G = (V, E)$ be a d-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 $(1-c)n((1-c)d + c\lambda_G)^\ell$.*

The set Z of Theorem 12 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 12, we split $X_1(t)$ into two disjoint sets A, A' of (almost) equal size. For $u \in A$, 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

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

where $\lambda \leq .99$ (see Lemma 8) is the absolute value of the second largest eigenvalue of the transition matrix of Z. Thus

$$\mathbb{E}_A(H(B')) \leq \sum_{\ell \geq 1} p_\ell \leq \frac{2(1-c)n}{|X_1(t)|}\frac{1}{c(1-\lambda)}. \tag{16}$$

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

$$\mathbb{E}(C(t+1) - C(t) \mid x_{2k} \in \mathcal{P}(X_1(t))) = O\left(\frac{(n - |X_3|)n}{|X_1|\,(|X_1| + |X_3|)}\right). \tag{17}$$

Using (12), (13), and other bounds for $|X_1(t)|, |X_3(t)|$,

$$\mathbb{E}(C(t_1)) = \sum_{t=1}^{t_1} \mathbb{E}(C(t) - C(t-1)) = o(n \log n).$$

Details are omitted in this extended abstract.

6.2 Later Stages

We will now use Lemmas 10 and 11, together with Definition 9 and equations (12) – (14). For $t = (1-\delta)\frac{3n}{2}$ with $\delta \le \delta_1 = \log^{-1/2} n$ we set $\omega = \omega(t) = \log(-\log\delta)$ and define the events (with $\overline{X}(t) = X_1(t) \cup X_2(t) \cup X_3(t)$)

$$\mathcal{A}(t) = \{|X_1(t) - 3n\delta| = O(\omega^{-1}\delta n)\}, \tag{18}$$

$$\mathcal{B}(t) = \{\overline{X}(t) \text{ is a root set of order } \omega\}. \tag{19}$$

and set $\mathcal{E}(t) = \mathcal{A}(t) \cap \mathcal{B}(t)$. As a consequence of Lemma 10, equation (16) and the fact that $\mathbb{E}\left(\overline{X}(t)\right) = 3n - 2t + o(3n-2t)$, we have

$$\mathbb{E}\left(C(t+1) - C(t)\right) = (3 \pm \varepsilon)\frac{n}{3n-2t}\Pr\{\mathcal{E}(t)\} + O\left(\frac{n}{3n-2t}\right)\Pr\left\{\overline{\mathcal{E}(t)}\right\} + O(\log n). \tag{20}$$

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 π, at which time we apply Lemma 10. Here we rely on $\lambda^*(t) \le 0.99$. In the event of $\overline{\mathcal{E}(t)}$ we use the fact that $\overline{X}(t) = \Omega(3n-2t)$, which follows from (13) and the well-known hitting time bound $\frac{1}{1-\lambda}\frac{n}{\overline{X}(t)}$ (see e.g. Jerrum and Sinclair [14]) to conclude that the hitting time is $O(n/(3n-2t))$.

The bound (12) for $|X_1(t)|$ implies that $\mathcal{A}(t)$ occurs w.h.p. for any fixed $t \ge t_1$ and we will prove that $\mathcal{B}(t)$ also occurs w.h.p. Lemma 6 will follow. The relatively simple proof of Lemma 7 is sketched at the end.

▶ **Lemma 13.** *Fix t and let $\delta = (3n-2t)/3n$. If $\delta_1 = \log^{-1/2} n \ge \delta \ge \delta_4 = n^{-1}\log^{11} n$ then,*

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

Proof. Fix some t, δ with $t_1 \le t \le t_3$. Expose $[W(t)]$. As in (12) and (14), w.h.p.,

$$\Phi(t) \ge (\delta_0\delta)^{1/2} n, \tag{21}$$

$$|X_1(t)| = 3\delta n + O(\omega^{-1}\delta n). \tag{22}$$

As already remarked, this shows that $\Pr\{\mathcal{A}(t)\} = 1 - o(1)$. By (13), w.h.p. $X_3(t) \approx n\delta^{3/2} = o(X_1(t))$. 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 w.h.p. Here ω is chosen to satisfy (25) 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

$$e(\overline{X}(t)) = |E_t^c| + |E(X_1 \cup X_2) \cap E_t| \tag{23}$$

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 $\omega^3 = o(\delta_0^{-1/2})$.

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

$$\sum_{u,v\in X_1} \Pr\{d(u,v) \le \omega\} = O\left(\frac{|X_1|^2 3^\omega}{n(\delta_0\delta)^{1/2}}\right) = O(n\delta_0^{-1/2}\delta^{3/2}3^\omega) = o(|X_1|/\omega^3), \tag{24}$$

if we choose

$$\omega^3 3^\omega = o\left((\delta_0/\delta)^{1/2}\right). \tag{25}$$

w.h.p. 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 w.h.p.

We show in the full paper version that w.h.p., $\mathcal{E}(t_3)$ holds with enough room to spare so that $\mathcal{E}(t)$ must hold for $t_3 \le t \le t_4$. ◀

For $t \ge t_4$, we use the bound

$$\mathbb{E}\left(C(t+1) - C(t)\right) \le \frac{1}{1-\lambda}\frac{n}{|\overline{X}(t)|},$$

see e.g. Jerrum and Sinclair [14], to conclude that $\mathbb{E}\left(C(3n/2) - C(t_4)\right) = o(n \log n)$.

References

1. Miklós Ajtai, János Komlós, and Endre Szemerédi. Deterministic simulation in LOGSPACE. In Alfred V. Aho, editor, *Proceedings of the 19th Annual ACM Symposium on Theory of Computing, 1987, New York, New York, USA*, pages 132–140. ACM, 1987. doi:10.1145/28395.28410.
2. David Aldous and James Allen Fill. Reversible markov chains and random walks on graphs, 2002. Unfinished monograph, recompiled 2014, available at http://www.stat.berkeley.edu/~aldous/RWG/book.html.
3. Noga Alon, Itai Benjamini, Eyal Lubetzky, and Sasha Sodin. Non-backtracking random walks mix faster. *Communications in Contemporary Mathematics*, 09(04):585–603, 2007. doi:10.1142/S0219199707002551.
4. Petra Berenbrink, Colin Cooper, and Tom Friedetzky. Random walks which prefer unvisited edges: Exploring high girth even degree expanders in linear time. *Random Struct. Algorithms*, 46(1):36–54, 2015. doi:10.1002/rsa.20504.
5. Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Eur. J. Comb.*, 1(4):311–316, 1980. doi:10.1016/S0195-6698(80)80030-8.
6. Colin Cooper and Alan M. Frieze. The cover time of random regular graphs. *SIAM J. Discrete Math.*, 18(4):728–740, 2005. doi:10.1137/S0895480103428478.
7. Colin Cooper and Alan M. Frieze. Vacant sets and vacant nets: Component structures induced by a random walk. *SIAM J. Discrete Math.*, 30(1):166–205, 2016. doi:10.1137/14097937X.
8. Colin Cooper, Alan M. Frieze, and Samantha Petti. The cover time of a biased random walk on $G_{n,\,p}$. In Markus E. Nebel and Stephan G. Wagner, editors, *Proceedings of the Fifteenth Workshop on Analytic Algorithmics and Combinatorics, ANALCO 2018, New Orleans, LA, USA, January 8-9, 2018.*, pages 158–167. SIAM, 2018. doi:10.1137/1.9781611975062.14.

9 Uriel Feige. A tight lower bound on the cover time for random walks on graphs. *Random Struct. Algorithms*, 6(4):433–438, 1995. doi:10.1002/rsa.3240060406.

10 Uriel Feige. A tight upper bound on the cover time for random walks on graphs. *Random Struct. Algorithms*, 6(1):51–54, 1995. doi:10.1002/rsa.3240060106.

11 Joel Friedman. A proof of Alon's second eigenvalue conjecture. In *Proceedings of the Thirty-fifth Annual ACM Symposium on Theory of Computing*, STOC '03, pages 720–724, New York, NY, USA, 2003. ACM. doi:10.1145/780542.780646.

12 Alan M. Frieze and Michal Karonski. *Introduction to Random Graphs.* Cambridge University Press, Cambridge, UK, 2015. doi:10.1017/CBO9781316339831.011.

13 Tal Orenshtein and Igor Shinkar. Greedy random walk. *Combinatorics, Probability & Computing*, 23(2):269–289, 2014. doi:10.1017/S0963548313000552.

14 Alistair Sinclair and Mark Jerrum. Approximate counting, uniform generation and rapidly mixing markov chains. *Inf. Comput.*, 82(1):93–133, 1989. doi:10.1016/0890-5401(89)90067-9.

15 Hui Tian, Hong Shen, and Teruo Matsuzawa. RandomWalk routing for wireless sensor networks. In *Sixth International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT 2005), 5-8 December 2005, Dalian, China*, pages 196–200. IEEE Computer Society, 2005. URL: http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=10544, doi:10.1109/PDCAT.2005.193.