Random walks which prefer unvisited edges. Exploring high girth even degree expanders in linear time

Petra Berenbrink[*] Colin Cooper[†] Tom Friedetzky[‡]

March 13, 2016

Abstract

Let $G = (V, E)$ be a connected graph with $|V| = n$ vertices. A simple random walk on the vertex set of G is a process, which at each step moves from its current vertex position to a neighbouring vertex chosen uniformly at random.

We consider a modified walk which, whenever possible, chooses an unvisited edge for the next transition; and makes a simple random walk otherwise. We call such a walk an *edge-process* (or E-process). The rule used to choose among unvisited edges at any step has no effect on our analysis. One possible method is to choose an unvisited edge uniformly at random, but we impose no such restriction.

For the class of connected even degree graphs of constant maximum degree, we bound the vertex cover time of the E-process in terms of the edge expansion rate of the graph G, as measured by eigenvalue gap $1 - \lambda_{\max}$ of the transition matrix of a simple random walk on G.

A vertex v is ℓ-good, if any even degree subgraph containing all edges incident with v contains at least ℓ vertices. A graph G is ℓ-good, if every vertex has the ℓ-good property. Let G be an even degree ℓ-good expander of bounded maximum degree. Any E-process on G has vertex cover time

$$C_V(E\text{-process}) = O\left(n + \frac{n \log n}{\ell}\right).$$

This is to be compared with the $\Omega(n \log n)$ lower bound on the cover time of any connected graph by a weighted random walk. Our result is independent of the rule used to select the order of the unvisited edges, which could, for example, be chosen on-line by an adversary.

As no walk based process can cover an n vertex graph in less than $n - 1$ steps, the cover time of the E-process is of optimal order when $\ell = \Theta(\log n)$. With high probability random r-regular graphs, $r \geq 4$ even, have $\ell = \Omega(\log n)$. Thus the vertex cover time of the E-process on such graphs is $\Theta(n)$.

[*]School of Computing Science, Simon Fraser University, Burnaby, Canada
[†]Department of Informatics, King's College, London, U.K.
[‡]School of Engineering and Computing Sciences, Durham University, Durham, U.K.

1 Introduction

Let $G = (V, E)$ be a connected graph with $|V| = n$ vertices and $|E| = m$ edges. A simple random walk on the vertex set of G is a process, which at each step moves from its current vertex position to a neighbouring vertex chosen uniformly at random. At step $t = 0$, the walk is located at a start vertex $v \in V$, and we write $W_v(0) = v$. Let $N(v) = \{u : \{v, u\} \in E\}$ be the disjoint neighbour set of v. At step $t \geq 0$, if the walk is at vertex $x = W_v(t)$, it chooses a vertex y uniformly at random (u.a.r.) from $N(x)$, and makes the edge transition (x, y) to $y = W_v(t+1)$. The trajectory of the walk is $W_v = (W_v(t),\ t \geq 0)$.

We consider a related walk process $X = (X(t),\ t \geq 0)$ on the vertex set of G which we call an *edge-process* (or E-process). Initially all edges E of G are marked as unvisited. At each step the edge-process makes a transition to a neighbour of the currently occupied vertex as follows: If there are *unvisited edges* incident with the current vertex pick one, make a transition along this edge and mark the edge as visited. If there are no unvisited edges incident with the current vertex, move to a u.a.r. neighbour using a simple random walk. We assume there is a rule $\mathcal{A}$, which tells the walk how to choose among unvisited edges. In the simplest case, $\mathcal{A}$ chooses u.a.r. among unvisited edges incident with the current vertex occupied by the walk. However we do not exclude arbitrary choices of rule $\mathcal{A}$. For example, the rule could be deterministic, or decided on-line by an adversary, or could vary from vertex to vertex.

For any process $Y = (Y(t),\ t \geq 0)$ which explores a graph G by moving from vertex to vertex, the *vertex cover time*, $C_V(Y, G)$, is defined as follows. For $v \in V$, let C_v be the expected time taken for a walk Y_v starting at v, to visit every vertex of G. The vertex cover time is defined as $C_V(Y, G) = \max_{v \in V} C_v$. A similar definition holds for *edge cover time*, $C_E(Y, G)$. When it is clear from the context, we write $G_V(G)$ for the vertex cover time of a simple random walk. If G is fixed we write $C_V(Y)$ for the vertex cover time of G by process Y, and $C_V(SRW)$ for the vertex cover time of G by a simple random walk.

It was shown by Feige [8], that for any connected n-vertex graph G, the vertex cover time of a simple random walk W satisfies $C_V(G) \geq (1 - o(1))n \log n$. The comparable result that edge cover time of any connected m-edge graph is $C_E(G) = \Omega(m \log m)$ is due to [1], [17].

The idea that the vertex cover time of a random walk could be reduced by choosing unvisited neighbour vertices whenever possible seems attractive and often arises in discussion. For sparse graphs it seems plausible that a random walk which prefers unvisited edges might also perform well. In the same vein, we might seek to reduce edge cover time by using a random walk which prefers unvisited edges. The edge cover time of this version of the E-process, (in which the next unvisited edge is chosen u.a.r.), was studied by Orenshtein and Shinkar [13], under the name of *Greedy Random Walk*. A random process which generalized the idea of choosing an unvisited vertex was studied experimentally by Avin and Krishnamachari [3]. This process, the *Random Walk with Choice*, (RWC(d)), selects d neighbours uniformly at random at each step, and moves to the least visited vertex among them. Avin and Krishnamachari investigate the process RWC(d) on geometric random graphs and the toroidal grid.

They report reductions in cover time, and improved concentration of experimental results.

The E-process has similarities with deterministic walks such as the rotor-router, or Propp machine model. (See [6] for an introduction to this topic.) In the case of the rotor-router process, the graph is turned into an Eulerian digraph by replacing each edge with a pair of oppositely directed edges. The vertex cover time of a connected graph in the rotor-router model is $O(mD)$, where m is the number of edges of G, and D is the diameter, (see [16]). The analysis of the rotor-router process depends on the underlying Eulerian properties of the graph. The E-process can be seen as a hybrid between a rotor-router and a random walk. Other deterministic processes to explore graphs based on edge utilization were studied in [5] under the theme of *locally fair exploration*. At a given vertex the process either picks the least frequently traversed edge (Least-Used-First), or the edge which has waited the longest time to be visited (Oldest-First). It is shown that Oldest-First can lead to exponential cover times on some connected graphs, where Least-Used-First covers all vertices in $O(m|D|)$ steps and traverses all edges with the same frequency in the long run.

The graphs we consider in this paper are connected n vertex, m edge graphs of even degree. We will henceforth always assume this is the case unless explicitly stated otherwise. We define a local expansion property of vertices on even degree graphs. We say a vertex v is ℓ-good, if any even degree subgraph containing all edges incident with v contains at least ℓ vertices. A graph G is ℓ-good, if every vertex has the ℓ-good property. Let the eigenvalues of the transition matrix of a simple random walk on G be ordered $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, and let $\lambda_{\max} = \max(\lambda_2, |\lambda_n|)$. We bound the cover time of the E-process in terms of the edge expansion rate of G, as measured by eigenvalue gap $1 - \lambda_{\max}$. We will assume henceforth that $\lambda_{\max} = \lambda_2$. If this is not the case (e.g. the graph is bipartite), we can make the random walk lazy, at most doubling the cover time.

We use the notation *with high probability* (**whp**) to mean with probability $1 - \epsilon(n)$, where $\epsilon(n) \to 0$ as the size of the vertex set $n \to \infty$. We often round real values to integers in our proofs. We do this without mention, as long as no significant error is introduced.

Vertex cover time of E-process

Any walk-based process must take at least $n - 1$ steps to visit every vertex, so if we can find a process and a structural property of a graph which gives an $O(n)$ upper bound on the cover time of the graph, this result is of optimal order.

The following theorem is a general statement on the vertex cover time for any E-process.

Theorem 1 *Let G be a connected n vertex even degree graph, with finite maximum degree, and with the additional property that that G is ℓ-good. Then,* **any** *E-process on G has cover time*

$$C_V(E\text{-}process) = O\left(n + \frac{n \log n}{\ell(1 - \lambda_{\max})}\right).$$

We briefly list a series of remarks and corollaries which arise from Theorem 1

The upper bound on the cover time given in Theorem 1 is independent of the rule $\mathcal{A}$ used to select unvisited edges, even if this choice is decided on-line by an adversary.

For expander graphs, which have positive constant eigenvalue gap, Theorem 1 becomes

$$C_V(E\text{-process}) = O\left(n + \frac{n \log n}{\ell}\right). \tag{1}$$

In particular, for ℓ-good even degree expanders where $\ell = \Omega(\log n)$, the E-process covers the graph in $\Theta(n)$ steps, which is best possible.

The lower bound on the vertex cover time of G by *any* reversible random walk W is $C_V(W, G) = \Omega(n \log n)$. A proof of the $\Omega(n \log n)$ lower bound on the cover time of reversible random walks, due to T. Radzik [14], is given in Section 2.2. For expander graphs, the comparable vertex cover time of the E-process is given by (1). This gives a speed up of $\Omega(\min(\log n, \ell))$ compared to *any* reversible random walk.

Examples of ℓ-good graphs where $\ell = \Omega(\log n)$ **whp** include random r-regular graphs of even degree, and fixed degree sequence random graphs, with all vertex degrees $d(v) \geq 4$, even and finite. The following corollary, proved in Section 4 is typical.

Corollary 2 *Let $r \geq 4$ even. Let $\mathcal{G}_r$ denote the class of random r-regular graphs. Let G be sampled uniformly at random from $\mathcal{G}_r$, then with high probability $C_V(E\text{-process}) = \Theta(n)$.*

In the case of random r-regular graphs, $r \geq 3$ odd, there is good reason to believe that the cover time of the u.a.r. E-process is $\Theta(n \log n)$. A further discussion of the reasons for this, and of experimental results are given in Section 5.

Edge cover time of E-process

A version of the E-process, in which the next unvisited edge is chosen u.a.r., was studied by Orenshtein and Shinkar [13], under the name of *Greedy Random Walk*. They give the following upper bound for the edge cover time $C_E(GRW)$ of m edge, n vertex r-regular graphs G using a Greedy Random Walk:

$$C_E(GRW) = m + O\left(\frac{n \log n}{1 - \lambda_{\max}}\right). \tag{2}$$

For an expander with $1 - \lambda_{\max} > 0$ constant, and $r = \Omega(\log n)$ the expected edge cover time of the Greedy Random Walk is $\Theta(m)$ and hence linear in the number of edges. The result of [13] holds for any r, and not just r even.

For any even degree graph, a general bound on the edge cover time of any E-process can be found a follows. Let $C_V(SRW)$ denote the cover time of graph G by a simple random walk, then

$$m \leq C_E(E\text{-process}) \leq m + C_V(SRW). \tag{3}$$

These bounds are explained as follows. Clearly m steps are need to cover the edges. The embedded random walk of the E-process visits every vertex in $C_V(SRW)$ expected time. Whenever the random walk first visits a vertex u with unexplored edges, the E-process begins an edge exploration which completes when all edges incident with u are explored. The formal details of this argument are given in Section 3.1.

An example where the upper bound in (3) is tight but (2) is not, is an E-process on the n-vertex hypercube H_r. The E-process on H_r has edge cover time $\Theta(n \log n)$ which improves on the $\Theta(n \log^2 n)$ edge cover time for H_r obtained by a simple random walk. This can be argued as follows. H_r has degree $r = \log_2 n$ so $2m = n \log_2 n$. The vertex cover time of the H_r is $C_V(SRW) = \Theta(n \log n)$ ([12]). The second eigenvalue of the hypercube is $1 - 2/\log_2 n$, so the upper bound in (2) is $O(n \log^2 n)$. We note that the edge cover time of H_r by a simple random walk is $\Theta(m \log m) = \Theta(n \log^2 n)$.

An upper bound for the cover time of the E-process on any connected graph can be obtained using a result of Ding, Lee and Peres [7] on the the blanket time of a simple random walk. Let $\pi_v = d(v)/2m$ be the stationary probability of v. For any $\delta \in (0,1)$ define $\tau_{bl}(\delta)$ to be the first step t when every vertex v of G has been visited $N_v \geq \delta \pi_v t$ times. Let $t_{bl}(\delta) = \mathbf{E}\tau_{bl}(\delta)$. It follows from Theorem 1.1 of [7] that $t_{bl}(\delta) = O(C_V(SRW))$.

Provided a vertex v of graph G is visited at least $d(v)$ times by the embedded random walk of the E-process, then all edges incident with v must be explored. For any connected r-regular graph, $r \geq 3, r = O(\log n)$, the blanket time $t_{bl}(\delta) = O(C_V(SRW))$. This, and the facts that $\pi_v = 1/n$, $C_V(SRW) = \Omega(n \log n)$, imply that the time $T(r)$ for a random walk to visit all vertices at least r times has expected value $\mathbf{E}T(r) = O(C_V(SRW))$. Thus

$$C_E(E\text{-process}) = O(m + C_V(SRW)). \tag{4}$$

The upper bounds given in (2), (3), (4) are never better than $O(n \log n)$. For $m = \Theta(n)$ an $o(n \log n)$ upper bound can be given for some classes of even degree graphs, such as high girth expanders and random regular graphs. The following general result, based on girth, holds for any even degree graph.

Theorem 3 *Let G be a connected n vertex even degree graph, with maximum degree Δ, girth g and m edges Then,* **any** *E-process on G has edge cover time*

$$C_E(E\text{-}process) = O\left(m + \frac{m}{(1-\lambda_{\max})^2}\left(\frac{\log n}{g} + \log \Delta\right)\right).$$

Thus even degree expanders of constant maximum degree have edge cover time $C_E = O(n + (n \log n)/g)$. See [11] for the construction of high girth expanders.

Better results for sparse random graphs can be obtained by an inspection of the graph structure. For example, random regular graphs have constant girth **whp**, but the total number of small cycles can be bounded above to give the following result.

Corollary 4 *Let $r \geq 4$ even. Let $\mathcal{G}_r$ denote the class of random r-regular graphs. Let G be sampled uniformly at random from $\mathcal{G}_r$, then for any $\omega \to \infty$, with high probability $C_E(E\text{-process}) = O(\omega n)$.*

Outline of the paper

It is helpful to think of the E-process re-colouring the edges of the graph as it proceeds. We consider unvisited edges as coloured blue, and explored edges as coloured red. Suppose at some point all edges incident with the vertex currently occupied by the walk are red (previously explored). It can be shown that any remaining blue edges in the graph (unexplored edges) form even degree edge induced components (blue components). Any unvisited vertex must be part of such a blue component. However not every blue component need contain unvisited vertices.

For connected graphs of even degree, our method to upper bound the vertex cover time of the E-process can be summarized as follows.

The E-process walk starts at some vertex and follows unexplored edges. This continues until the walk returns to the start vertex. This follows from the even degree assumption, and a simple parity argument. See Section 3.1 for details.

If all edges at the currently occupied vertex are red, the process makes a random walk. As soon as this random walk visits some blue component C, it will begin to follow unexplored edges with the possibility of reaching an unvisited vertex. If a blue component C contains at least ℓ vertices, and the graph is an expander, it can be proved that **whp** a random walk visits this component in $O(n \log n/\ell)$ random walk steps. This is established in Section 3.2. Unfortunately, the E-process walk on unexplored edges of C may only re-colour part of this component before exit, and never reach an unvisited vertex. To get around this, we prove that there is no edge induced subgraph of size $\Theta(\ell)$ rooted at any unvisited vertex which is left untouched by the random walk in $O(n \log n/\ell)$ steps.

In Section 2 we establish some general bounds on the time taken by a random walk to visit sets of vertices S of a given size. In particular the exponentially strong bound given in Lemma 8 can be used to ensure that no edge induced subgraph of size $\Theta(\ell)$ is left untouched by the walk in $O(n \log n/\ell)$ steps.

2 Properties of random walks

In this section we give various results on random walks needed for our proofs.

2.1 Convergence to stationarity

Let $G = (V, E)$ denote a connected graph, $|V| = n$, $|E| = m$, and let $d(v)$ be the degree of a vertex v. A *simple random walk* W_u, $u \in V$, on graph G is a Markov chain modeled by a particle moving from vertex to vertex according to the following rule. The probability of transition from vertex v to vertex w is equal to $1/d(v)$, if w is a neighbour of v, and 0 otherwise. The walk W_u starts from vertex u at $t = 0$. Denote by $W_u(t)$ the vertex reached at step t; $W_u(0) = u$.

Let P be the transition matrix of a simple random walk on a graph G. Thus $P_{i,j} = 1/d(i)$ if and only if there is an edge between i and j in G. Let $P_u^{(t)}(v) = \mathbf{Pr}(W_u(t) = v)$ be the t-step transition probability. We assume the random walk W_u on G is ergodic with stationary distribution π, where $\pi_v = d(v)/(2m)$. If this is not the case, e.g. G is bipartite, then the walk can be made ergodic, by making it lazy. A random walk is *lazy*, if it moves from v to one of its neighbours w with probability $1/(2d(v))$, and stays where it is (at vertex v) with probability $1/2$.

Let $1 \geq \lambda_2 \geq, \ldots \geq \lambda_n$, be the eigenvalues of P, and let $\lambda_{\max} = \min(|\lambda_2|, |\lambda_n|)$. We henceforth assume that $\lambda_2 = \lambda_{\max}$. This can be achieved by making the chain lazy, which has no significant effect on our analysis.

The convergence to stationarity of a simple random walk is bounded by

$$|P_u^{(t)}(x) - \pi_x| \leq (\pi_x/\pi_u)^{1/2} \lambda_{\max}^t. \tag{5}$$

2.2 Lower bound on cover time for reversible random walks

For a random variable Z depending on W, define $\mathbf{E}_u Z$ as the expectation of Z for the walk W_u started from u at time $t = 0$. For a walk W_u, started from vertex u, the first return time to u is defined as $T_u^+ = \min(t \geq 1 : W_u(t) = u)$. The expected first return time $\mathbf{E}_u T_u^+$ is given by $\mathbf{E}_u T_u^+ = 1/\pi_u$. (See e.g. [2] Chapter 2, Lemma 5 for a proof of this.) The first visit time from u to v is similarly defined as $T_{uv} = \min(t \geq 1 : W_u(0) = u, W_u(t) = v)$. The (expected) hitting time of vertex v is $\mathbf{E}_u T_{uv}$ and the commute time $K(u, v)$ between vertices u and v, is the expected time taken to go from vertex u to vertex v and then back to vertex u. Thus $K(u, v) = \mathbf{E}_u T_{uv} + \mathbf{E}_v T_{vu}$.

A reversible weighted random walk W is defined by a set of weights $w(e) > 0$ on the edges $e \in E(G)$, and transition probabilities $p(x, y)$ from x to $y \in N(x)$ given by

$$p(x, y) = \frac{w(x, y)}{\sum_{z \in N(x)} w(x, z)}.$$

For an introduction to weighted random walks see [2] Chapter 3.

The following proof that the cover time of any weighted random walk W is $\Omega(n \log n)$, is due to T. Radzik [14].

Theorem 5 *For any weighted random walk W and n vertex graph G,*

$$C_v(W,G) \geq \frac{n}{4}\log\frac{n}{2}.$$

Proof Any walk starting from u either visits v before returning to u or it does not. Thus $\mathbf{E}T_u^+$ is at most the commute time $K(u,v)$ between u and v.

Let S be the subset of vertices with $\pi_u \leq 2/n$. Thus $|S| \geq n/2$. This follows because $\sum_{u\in V}\pi_u = 1$. As $\mathbf{E}T_u^+ = 1/\pi_u$, it follows that for $u \in S$, $\mathbf{E}T_u^+ \geq n/2$.

Let $K_S = \min_{i,j\in S} K(i,j)$ then, $K_S \geq \mathbf{E}T_u^+ \geq n/2$. From [10], we have the following lower bound on the vertex cover time

$$C_V(W,G) \geq (\max_{A\subseteq V} K_A \log|A|)/2.$$

However for the set S given above we have $K_S \log|S| \geq (n/2)\log(n/2)$. □

Visits to a Single Vertex Let H_v be the first visit time of vertex v by a random walk W. For a walk W_u starting from vertex u, let $\mathbf{E}_u(H_v)$ denote the expected value of H_v. If $v = u$ then $H_v = 0$. If the distribution of the random walk at some step $t \geq 0$ is $\rho = (\rho(u), u \in V)$, we define the hitting time from starting distribution ρ as $\mathbf{E}_\rho(H_v) = \sum_{u\in V}\rho(u)E_u(H_v)$. In particular, for a random walk starting at a vertex chosen according to the stationary distribution π, let $\mathbf{E}_\pi(H_v)$ denote the expected hitting time of vertex v from stationarity.

The quantity $\mathbf{E}_\pi(H_v)$ can be expressed in the following way, (see e.g. [2], Chapter 2.2)

$$\mathbf{E}_\pi(H_v) = Z_{vv}/\pi_v, \tag{6}$$

where

$$Z_{vv} = \sum_{t=0}^{\infty}(P_v^{(t)}(v) - \pi_v). \tag{7}$$

Using (5), we can upper bound $\mathbf{E}_\pi(H_v)$ as follows.

Lemma 6

$$\mathbf{E}_\pi(H_v) \leq \frac{1}{(1-\lambda_{\max})\pi_v}. \tag{8}$$

Proof Using (5) with $x = u = v$, then

$$|P_v^t(v) - \pi_v| \leq (\lambda_{\max})^t,$$

and

$$Z_{vv} = \sum_{t\geq 0}(P_v^t(v) - \pi_v) \leq \sum_{t\geq 0}(\lambda_{\max})^t = \frac{1}{1-\lambda_{\max}}.$$

□

Lemma 7 *Let G be an n vertex (multi)-graph with eigenvalue gap $(1-\lambda_{\max})$, and maximum degree $\Delta \leq n^2$. Let*

$$T = K \log n/(1-\lambda_{\max}). \tag{9}$$

For $K \geq 6$, T is a mixing time for a random walk on G, such that, for $t \geq T$,

$$\max_{u,x \in V} |P_u^{(t)}(x) - \pi_x| \leq \frac{1}{n^3}. \tag{10}$$

Proof

For $\lambda \leq 1$, $\lambda \leq e^{-(1-\lambda)}$. It follows from (5), for given u, x that

$$|P_u^t(x) - \pi_x| \leq \Delta^{1/2} e^{-(1-\lambda_{\max})t}, \tag{11}$$

Let

$$T = K \log n/(1-\lambda_{\max}),$$

where $K \geq 6$. As there are at most n^2 pairs u, x, and $\Delta \leq n^2$, then using (11)

$$\sum_{u,x} |P_u^t(x) - \pi_x| \leq n^2 \Delta^{1/2} e^{-T(1-\lambda_{\max})} \leq \frac{1}{n^3}. \tag{12}$$

□

Let $\boldsymbol{A}_{t,u}(v)$ denote the event that W_u does not visit vertex v in steps $0, ..., t$. Lemma 8 gives an upper bound on $\mathbf{Pr}(\boldsymbol{A}_{t,u}(v))$ for any start vertex u in terms of $\mathbf{E}_\pi(H_v)$ and the mixing time T.

Lemma 8 *Let T_G be a mixing time of a random walk W_u on G satisfying* (10). *Then for any start vertex u*

$$\mathbf{Pr}(\boldsymbol{A}_{t,u}(v)) \leq e^{-\left\lfloor t/(T_G + 3\mathbf{E}_\pi(H_v)) \right\rfloor}.$$

Proof Let $\rho = (\rho_w)$ be the distribution of W_u on G after $T = T_G$ steps, where $\rho_w = P_u^{(T)}(w)$. Let $\mathbf{E}_\rho(H_v)$ be the expected time to hit v starting from ρ. As T satisfies (10), and $\pi_x \geq 1/n^2$ for any connected graph, then $\rho_w = (1+o(1))\pi_w$. It follows that

$$\mathbf{E}_\rho(H_v) = (1+o(1))\mathbf{E}_\pi(H_v). \tag{13}$$

Let $H_v(\rho)$ be the time to hit v starting from ρ, then

$$\mathbf{Pr}[H_v(\rho) \geq 3\mathbf{E}_\pi(H_v)] \leq \frac{(1+o(1))}{3} \leq \frac{1}{e}. \tag{14}$$

Let $\tau = T + 3\mathbf{E}_\pi(H_v)$. By considering the walk W_u for intervals of τ steps starting at $s = 0, \tau, 2\tau, \cdots, \lfloor t/\tau \rfloor \tau$, and applying (14) to each interval gives

$$\mathbf{Pr}(\boldsymbol{A}_{t,u}(v)) \leq e^{-\lfloor t/\tau \rfloor}.$$

□

Visits to Vertex Sets We extend the results presented above to any nonempty subset S of vertices in the following way. Let $d(S) = \sum_{v \in S} d(v)$ be the degree of set $S \subseteq V$. From G we obtain a (multi)-graph $\Gamma = \Gamma_S$ by contracting S to a single vertex γ. We retain multiple edges and loops in Γ, so that $d(S) = d(\gamma)$, and $|E(\Gamma)| = |E(G)| = m$. Let $\widehat{\pi}$ be the stationary distribution of a random walk on Γ. If $v \notin S$ then $\widehat{\pi}_v = \pi_v$, and $\widehat{\pi}_\gamma = \pi_S \equiv \sum_{x \in S} \pi_x$.

For $u \notin S$ let W_u be a walk starting from u in G, and let $\widehat{W}_u$ be the equivalent walk starting in Γ. Let $\boldsymbol{A}_{t,u}(S,G)$ be the event that W_u does not visit S in t steps, and let $\boldsymbol{A}_{t,u}(\gamma,\Gamma)$ be the event that $\widehat{W}_u$ does not visit γ in t steps. Up to the first visit to S the walks W_u and $\widehat{W}_u$ have the same transition probabilities and can be identically coupled. Thus,

$$\mathbf{Pr}(\boldsymbol{A}_{t,u}(S,G)) = \mathbf{Pr}(\boldsymbol{A}_{t,u}(\gamma,\Gamma)).$$

Let H_S be the first visit time to set S, by a random walk. For a walk starting from $u \in S$, define $H_S = 0$. For any $u \notin S$ the above coupling gives

$$\mathbf{E}_\pi(H_S) \;=\; \mathbf{E}_{\widehat{\pi}}(H_\gamma). \tag{15}$$

It is a known result that contracting vertex sets increases the eigenvalue gap. (For a proof see e.g. [2] Chapter 3, Corollary 27.) Thus

$$1 - \lambda_{\max}(G) \le 1 - \lambda_{\max}(\Gamma). \tag{16}$$

Thus Lemma 6 applies equally to Γ with $\lambda_{\max} = \lambda_{\max}(G)$. The choice of $K \ge 6$ for the mixing time T in Lemma 7 ensures that (10) holds in both G and Γ. The upper bound in Lemma 8 also holds in Γ with this mixing time T. We note the following consequence of (15) and Lemma 6.

Corollary 9 *Let $G = (V,E)$, let $|E| = m$. Let $S \subseteq V$, and let $d(S)$ be the degree of S. Then $\mathbf{E}_\pi H_S$, the expected hitting time of S from stationarity satisfies*

$$\mathbf{E}_\pi H_S \le \frac{2m}{d(S)(1-\lambda_{\max}(G))}.$$

3 Proof of main results

3.1 Properties of the edge-process on even degree graphs

Let $X(t)$ be the position at step t of a particle moving according to an E-process. It is helpful to think of the progress of the E-process as a re-colouring of the edges of the graph G. We consider unvisited edges as coloured blue, and visited edges as coloured red.

Initially, the particle is at $X(0) = u$, the start vertex, and all edges of the graph G are coloured blue (unvisited). Given $X(t) = v$, $X(t+1)$ is chosen as follows. If all edges incident with v are red (previously visited) the walk chooses $X(t+1)$ u.a.r. from $N(v)$. If however, there are any blue (unvisited) edges incident with v, then the walk picks a blue edge (v, w) according to the rule $\mathcal{A}$. The walk then moves to $X(t+1) = w$, and re-colours the edge (v, w) red (visited). We assume that the edge (v, w) is re-coloured red at the start of step $t+1$, the instant at which the walk arrives at w. Thus we regard the transition (v, w) as being along a blue edge.

At each step t the next transition is either along a blue or a red edge. We speak of the sequence of these edge transitions as the blue (sub)-walk and the red (sub)-walk. The walk thus defines red and blue phases which are maximal sequences of edge transitions when the walk is the given colour. For any vertex v, and step t, the blue (resp. red) degree of v is the number of blue (resp. red) edges incident with v at the start of step t.

Observation 10 *Assume all vertices of G are of even degree. Then a blue phase of the E-process which starts at a vertex u (at some step t), must end at u (at some step $t + \tau$).*

Proof This follows from a simple parity argument. The first blue phase starts at $t = 0$, at the start vertex $X(0) = u$. At $t = 0$ every vertex has even blue degree. Suppose that at step t we have $X(t) = w$, where $w \neq u$. Inductively every vertex, apart from the start vertex u and the current position w have even blue degree, whereas the blue degree of u and w is odd, and hence greater than zero. The particle can thus exit w along a blue edge. When the particle leaves $w = X(t)$ making the transition $(X(t), X(t+1))$, then the blue degree of $w = X(t)$ becomes even. If $X(t+1) = u$, then the degree of u is even and the particle has returned to the start. If $X(t+1) \neq u$, then the blue degree of $X(t+1)$ and u is odd.

If the particle returns to u at step t, and the blue degree of u is zero, then the blue phase at u is completed at (the start of) step t. The particle now leaves u along a red edge $(u, v) = (X(t), X(t+1))$, and this is the beginning of a red phase. Inductively, the blue degree of v is even when the particle arrives at v. If v has blue edges incident with it, then a blue phase begins. Otherwise the red phase continues. □

Note that it is possible that all edges incident with a vertex v are coloured red by transitions made during the blue sub-walk, and but that v has not yet been visited by a red sub-walk.

Let $G[S]$ denote the subgraph of G induced by the set of vertices $S \subseteq V$. The following summarizes the consequences of Observation 10.

Observation 11 *Assume vertex v is unvisited at step t, and that the E-process is in a red phase.*

1. *All edges incident with v are blue at step t.*

2. *The blue degree of all vertices at step t is even.*

3. *Let S_v^* be the maximal blue (unvisited), edge induced subgraph obtained by fanning out in a breadth first manner from v using only blue edges. Let U^* be the vertex set of S_v^*. Then*

   (a) *The degree of v in S_v^* is $d(v)$, the degree of vertex v in G. All vertices of S_v^* have positive even degree.*

   (b) *All edges between U^* and $G \setminus U^*$ are red.*

   (c) *$G[U^*]$ may induce red edges, but these are not part of S_v^*.*

In the simplest case S_v^* consists of $d(v)/2$ blue cycles with common root vertex v, but otherwise vertex disjoint.

It follows from Observation 10, that if we ignore the blue phases of the E-process X, then the resulting red phases describe a simple random walk W on the graph G. If the E-process X starts at $X(0) = u$, then W must also start at vertex u after the completion of the first blue phase. Each step t_R of the walk $W_u(t_R)$ corresponds to some step $X_u(s)$, $s > t_R$ of the E-process X_u.

At step t of the E-process, we have $t = t_R + t_B$, where t_R, t_B are the (unknown) number of red and blue edge transitions. One thing is certain however; the length of the blue walk can be at most the number of edges m of G. Moreover, no vertex which is visited at step t_R of the red sub-walk can have unvisited edges at the start of step $t_R + 1$ of the red sub-walk. These observations are formalized as:

Observation 12 *Let $X_u(t)$ be the walk defined by an E-process starting from vertex u. Let $W_u(t_R)$ be the corresponding simple random walk on the graph G defined by the red phase of the E-process. Then $t_R < t < t_R + m$.*

In particular the edge cover time of the E-process is bounded by

$$m \le C_E(E\text{-process}) \le m + C_V(SRW),$$

where $C_V(SRW)$ is the vertex cover time of a simple random walk on G.

3.2 Vertex cover time of the E-process

Lemma 13 *Let W_u be a random walk starting from u in G. Let $\lambda_{\max} = \lambda_{\max}(G)$. Let S be a set of vertices of G. Let $d(S)$ be the sum of the degrees of the vertices in S. Suppose that*

$$d(S) \le m/6 \log n,$$

and that

$$t \ge 7m/d(S)(1 - \lambda_{\max}). \tag{17}$$

Then

$$\mathbf{Pr}(S \text{ is unvisited by } W_u \text{ at step } t) \le e^{-td(S)(1-\lambda_{\max})/14m}.$$

Proof Contract S to a single vertex $\gamma = \gamma(S)$, retaining all loops and parallel edges. Denote the resulting graph by Γ. Let Δ be the maximum degree in G or Γ as appropriate. In either case, $\Delta \leq 2m \leq n^2$. Let

$$T = 6 \log n/(1 - \lambda_{\max}(G)).$$

From (16) and Lemma 7, we have that T is a mixing time satisfying (10) in both G and Γ.

For $u \notin S$ let W_u be a walk starting from u in G, and let $\widehat{W}_u$ be the equivalent walk starting at u in Γ. Provided W_u does not visit S in t steps, (the event $\boldsymbol{A}_t(S,G)$), then $\widehat{W}_u$ does not visit γ (the event $\boldsymbol{A}_t(\gamma,\Gamma)$), and the walks have the same probabilities. Thus

$$\mathbf{Pr}(\boldsymbol{A}_t(S,G)) = \mathbf{Pr}(\boldsymbol{A}_t(\gamma,\Gamma)).$$

From Lemma 8 we have

$$\mathbf{Pr}(\boldsymbol{A}_t(\gamma,\Gamma)) \leq \exp\left(-\lfloor t/(T + 3\mathbf{E}_{\widehat{\pi}}(H_\gamma))\rfloor\right).$$

Assuming $d(S) \leq m/6\log n$, we have that $T \leq m/(d(S)(1-\lambda_{\max})$. From Lemma 6 and (16)

$$\mathbf{E}_{\widehat{\pi}}(H_\gamma)) \leq \frac{2m}{d(S)(1-\lambda_{\max})},$$

which gives

$$T + 3\mathbf{E}_{\widehat{\pi}}(H_\gamma) \leq \frac{7m}{d(S)(1-\lambda_{\max})}.$$

Provided $t \geq \tau$ then $\lfloor t/\tau \rfloor \geq t/2\tau$. As t satisfies (17) we have

$$\mathbf{Pr}(\boldsymbol{A}_t(S,G)) \leq \exp\left(-t\frac{d(S)(1-\lambda_{\max})}{14m}\right).$$

$\square$

Lemma 14 *Let G be a graph of maximum degree Δ. Let $\beta(s,v)$ be the number of connected edge induced subgraphs containing s vertices, and rooted at vertex v in G. Then*

$$\beta(s,v) \leq 2^{s\Delta}.$$

Proof We make a crude estimate for $\beta(s,v)$ by building a digraph H_v in a breadth first manner as follows. Initially $H_v = \emptyset$ and all adjacent edges of v in G are labeled unvisited. Mark v as processed and add it to H_v. For each edge incident with v, we label it as retained or excluded. Starting from v there are $d(v)$ unvisited edges, and so at most $2^{d(v)}$ choices for the subset of edges incident with v to retain. We process each retained edge (v,u) in increasing endpoint label order. Mark u as processed and add the retained edge (v,u) to H_v. There are at most $2^{d(u)-1}$ choices for labels (retained, excluded) of any unvisited edges incident with u.

Thus we fan out from v in a breadth-first manner using only retained edges, (u, w). We add w to H_v, and also any retained edges (x, w), where x was processed earlier than w. In general there are some number of retained and excluded edges incident with w in G, resulting from processing earlier vertices; and the remaining at most $(d(w) - 1)$ edges are unvisited. We continue until H_v has s processed vertices, and the choices at these vertices have been evaluated. The vertices of H_v and any retained edges between them define a connected subgraph of size s rooted at v, and every subgraph of size s rooted at v is found by this construction. $\square$

Lemma 15 *Let G be an ℓ-good even degree graph of minimum degree δ and maximum degree Δ. With probability $1 - O(n^{-3})$, after*

$$\tau^* = m\left(1 + \frac{14(\Delta + 4)\log n}{\delta \min(\ell, \log n)(1 - \lambda_{\max})}\right)$$

steps of the E-process, no vertex of G remains unvisited. The value of τ^* is independent of the choice of rule $\mathcal{A}$ used by the E-process.

In particular, if G has constant maximum degree, there exists a constant $B > 0$ such that

$$\tau^* = Bn[1 + (\log n)/\min(\ell, \log n)(1 - \lambda_{\max})].$$

Proof Let S_v^* be the maximal connected even degree blue subgraph rooted at v, as described in Observation 11. Let S_v be any connected subgraph of S_v^* of vertex size

$$s = \min(\ell, \log n),$$

rooted at v. By Lemma 14, there are at most $2^{\Delta s}$ such possible subgraphs.

Suppose some vertex v is unvisited at step t of the embedded random walk on red (visited) edges. The event that v is unvisited occurs if and only if there exists a blue (unvisited) edge induced subgraph S_v of vertex size s rooted at v, none of whose vertices have been visited. For a random walk W_u starting from vertex u, let $P_v(s, t)$ be the probability of this event. Let $d(S)$ be the minimum degree of any subgraph S_v. From Lemmas 13 and 14,

$$P_v(s, t) \leq 2^{\Delta s} e^{-t \frac{d(S)(1-\lambda_{\max})}{14m}}.$$

As $s = \min(\ell, \log n)$, on choosing

$$t^* = (\Delta + 4)\log n \, \frac{14m}{\delta s(1 - \lambda_{\max})},$$

we find that

$$\sum_v P_v(s, t^*) = O(1/n^3). \tag{18}$$

From Observation 12, the length of the E-process walk on unvisited edges is at most m, the number of edges of G, and the step $\tau^* = \tau(t^*)$ in the E-process corresponding to the step t^*

in the red phase random walk W_u is bounded by $\tau^* \leq m + t^*$. In particular, if Δ is constant then $m = cn$, and

$$\tau^* \leq m + t^* = B(n + (n \log n)/(\min(\ell, \log n)(1 - \lambda_{\max}))).$$

Suppose some vertex v is unvisited at τ^*. Then a blue (unvisited) edge induced subgraph S_v^* rooted at v exists at τ^*. However, from (18), **whp** any $S_v \subseteq S_v^*$ of size s, contains a vertex z already visited by the red sub-walk W at t^*. Suppose this visit occurs at $t \leq t^*$, but that, at step t^*, some edges incident with z are unvisited, a necessary condition for $z \in S_v^*$. On arriving at z, the E-process completes the exploration of all edges incident with z, after which the random walk $W(t)$ continues up to step t^*. Thus at τ^* all edges adjacent to z are red, which is a contradiction. □

The probability of a failure to reach an unvisited (blue) component after t^* steps is $O(n^{-3})$ (see (18)). To complete the proof of Theorem 1, we note that

$$C_V(G) \leq m + t^* + \sum_{k \geq 1} k t^* O(n^{-3k}) = O(\tau^*).$$

3.3 Edge cover time of the E-process

The conductance Φ of a graph G is defined as

$$\Phi(G) = \min_{\substack{X \subseteq V \\ d(X) \leq m(G)}} \frac{e(X : \overline{X})}{d(X)},$$

where $m(G) = |E(G)|$ is the number of edges of G, $d(X)$ is the degree of set X, and $e(X : \overline{X})$ is the number of edges between X and $\overline{X} = V \setminus X$. For a simple random walk on a graph G, the second eigenvalue λ_2 of the Markov chain satisfies

$$1 - 2\Phi \leq \lambda_2 \leq 1 - \frac{\Phi^2}{2}. \tag{19}$$

Using a lazy walk, we can assume that $\lambda_{\max} = \lambda_2$.

Recall that $\Gamma = \Gamma(S)$ is the graph obtained from G by contracting the set of vertices S to a vertex γ. From the construction of Γ it follows that $\Phi(\Gamma) \geq \Phi(G)$; every set of vertices in V_Γ corresponds to a set in V_G, and edges and degrees of vertices are preserved on contraction.

The proof of Theorem 3 is based on the following construction. Let G be a graph of girth g. Suppose there is an unvisited (blue) edge incident with a vertex v of G. There must be an unvisited cycle of length at least g passing through v. We examine the breadth first search tree of depth at most half the girth rooted at v for the existence of blue paths from leaf to leaf of this tree which pass through v.

For a given vertex v let $B_\ell(v)$ be the subgraph induced by the vertices at distance at most $\ell = \min(\lfloor g/2 \rfloor - 1, \log n)$ from v, and let $L(v)$ be the vertices at distance ℓ. The set $B_\ell(v)$

induces a tree, of which $L(v)$ are the leaves. For $x, y \in L(v)$ let xPy be the unique path from x to y in $B_\ell(v)$, and let $Q_v = \{xPy : xPy \text{ passes through } v\}$.

Let W_u be a random walk starting from u in G. We say xPy is unvisited at t, if the edge transitions $(W(s-1), W(s))$ of the walk up to step t are disjoint from the edges of xPy.

Lemma 16 *Let W_u be a random walk starting from u in G. For vertex v and $xPy \in Q$, if $t = \Omega(m/\ell(1-\lambda_{\max})^2)$, then*

$$\mathbf{Pr}(xPy \textit{ is unvisited by } W_u \textit{ at step } t) = O\left(e^{-t\ell(1-\lambda_{\max})^2/128m}\right).$$

Proof Subdivide the edges of xPy by inserting a vertex z_i of degree 2 in each edge $e_i, i = 1, ..., 2\ell$ of xPy. This gives a path $xP'y$, with an extra set of vertices $S = \{z_1, ..., z_{2\ell}\}$. The rest of G unaltered. Let G' be the resulting graph. Note that $|S| = 2\ell$, and $d(S) = 4\ell$. The edge set of G' has size $m + 2\ell$. Let Γ be obtained from G' by contracting S to a single vertex γ.

Let $\boldsymbol{A}_t(xPy, G)$ be the event that a walk $W_u[G]$ has not visited any edge of xPy at or before step t. Let $\boldsymbol{A}_t(S, G')$ be the event that a walk $W_u[G']$ has not visited any vertex of S at or before step t. Let $\boldsymbol{A}_t(\gamma, \Gamma)$ be the event that a walk $W_u[\Gamma]$ has not visited vertex γ at or before step t. The walks which start from $u \in V(G)$ and satisfy these conditions are identically coupled, so

$$\mathbf{Pr}(\boldsymbol{A}_t(xPy, G)) = \mathbf{Pr}(\boldsymbol{A}_t(S, G')) = \mathbf{Pr}(\boldsymbol{A}_t(\gamma, \Gamma)).$$

From here, the proof follows that of Lemma 13. As $\widehat{\pi}_\gamma = d(\gamma)/(2(m + 2\ell))$, and $d(\gamma) = 4\ell$ we upper bound $\mathbf{E}_{\widehat{\pi}}(H_\gamma)$ by

$$\mathbf{E}_{\widehat{\pi}}(H_\gamma) \leq \frac{m + 2\ell}{2\ell(1 - \lambda_{\max}(\Gamma))} = C(\gamma),$$

say. As $\ell \leq \log n = o(m)$ this gives $\tau = T_\Gamma + 3\mathbf{E}_{\widehat{\pi}}(H_\gamma) \leq 4C(\gamma)$. Provided $t \geq \tau$

$$\mathbf{Pr}(\boldsymbol{A}_t(\gamma, \Gamma)) \leq \exp\left(-t\frac{\ell(1 - \lambda_{\max}(\Gamma))}{4m}\right).$$

Using $d(xPy) \geq 4\ell$ and $d(\gamma) = 4\ell$ we find that $\Phi_{G'} \geq \Phi_G/2$. Finally (16) and (19) give

$$1 - \lambda_{\max}(\Gamma) \geq 1 - \lambda_{\max}(G') \geq \frac{\Phi^2(G)}{8} \geq \frac{1}{32}(1 - \lambda_{\max}(G))^2.$$

$\square$

Lemma 17 *Let G be an graph of maximum degree Δ and girth g With probability $1-O(n^{-1})$, after*

$$\tau^* = O\left(m + \frac{m}{(1 - \lambda_{\max})^2}\left(\frac{\log n}{\min(g, \log n)} + \log \Delta\right)\right)$$

steps of the E-process, no edge of G remains unvisited. The value of τ^* is independent of the choice of rule $\mathcal{A}$ used by the process.

In particular, if G has constant maximum degree, there exists a constant $B > 0$ such that

$$\tau^* = Bn[1 + (\log n)/\min(g, \log n)(1 - \lambda_{\max})^2]. \tag{20}$$

Proof

For $C \geq 260$ let t^* be given by

$$t^* = C\frac{m}{(1-\lambda_{\max}(G))^2}\left(\frac{\log n}{\ell} + 2\log\Delta\right).$$

At any vertex v, the number of choices $|Q_v|$ for paths xPy from leaf to leaf of $B_\ell(v)$ passing through v is at most $|Q_v| = \Delta^{2\ell}$. Thus

$$\begin{aligned}&\mathbf{Pr}(\text{There exists a vertex } v \text{ and path } xPy \in Q_v \text{ unvisited by } W_u \text{ at } t^*)\\ &\leq n\Delta^{2\ell}e^{-t^*\ell(1-\lambda_{\max})^2/128m} \qquad = O(n^{-1}).\end{aligned}$$

Let $\tau^* = m + t^*$. If the edge cover time of G is not $O(\tau^*)$, there exists an induced cycle $U, |U| \geq g$ all of whose edges are unvisited. Let v be some vertex of this cycle, and x, y vertices at distance ℓ from v in U. Then the corresponding path xPy is unvisited, an event of probability $O(n^{-1})$. □

4 Results for random regular graphs

Random r-regular graphs, $\mathcal{G}_r$, with $r \geq 4$ even, are an example of a class of graphs for which, with high probability $C_V(E\text{-process}) = \Theta(n)$. Let $\mathcal{G}'_r$ be the subset of $\mathcal{G}_r$ with the following properties.

(P1) G is connected, and the second eigenvalue of the adjacency matrix of G is at most $2\sqrt{r-1} + \varepsilon$, where $\varepsilon > 0$ is an arbitrarily small positive constant.

(P2) Let $s = O(\log n)$, and let $a = \lfloor 2s(\log re)/\log n \rfloor$. No set of vertices S of size s induces more than $s + a$ edges. In particular, for $s \leq (\log n)/(4\log re)$ no set of vertices S of size s induces more than s edges.

Lemma 18 *Let $\mathcal{G}'_r \subseteq \mathcal{G}_r$ be the r-regular graphs satisfying (P1), (P2). Then $|\mathcal{G}'_r| \sim |\mathcal{G}_r|$.*

Proof Friedman [9], shows the deep result that (P1) holds **whp** for random regular graphs. It is straightforward to establish that (P2) holds **whp**. □

4.1 Proof of vertex cover time (Corollary 2)

Let $\ell = \epsilon \log n$ for some $\epsilon > 0$. Property (P2) implies the graph is ℓ-good as follows. For any vertex v of the graph G, let U^* be the smallest non-trivial connected, even degree, vertex induced subgraph rooted at v. As $r \geq 4$, this subgraph contains at least two cycles. Let $|U^*| = k$, then U^* induces at least $k+1$ edges. By property (P2), no subgraph on less than $\ell = \log n/(4 \log re)$ vertices induces more than ℓ edges, and we conclude that $|U^*| \geq \ell$.

4.2 Proof of edge cover time (Corollary 4)

If a graph contains small cycles, these cycles can occur as unvisited even degree edge induced components, and the methods of Section 3.3 cannot be used unmodified.

Define a cycle as small if it is of length less than $\ell = \epsilon \log n$ for some $\epsilon > 0$ constant. Suppose vertex v is contained in exactly one small cycle C. For the edges $e_x = \{v, x\}$ of v not contained in C, there is no even degree edge induced component of girth less than ℓ containing e_x, and (20) of Lemma 17 applies. Suppose that after some step t the only unvisited cycles remaining are those of length less than ℓ. We assume this is the case. (By Lemma 17 this will occur after at most $\tau^* = \Theta(n)$ expected steps.) If the E-process is in a red-phase, and the random walk arrives at a vertex v on a small unvisited cycle C, the E-process will traverse C before resuming the random walk at v.

Referring to property (P2) above, let $\epsilon \leq 1/4 \log re$. This implies that **whp** all cycles of length k in G, for $3 \leq k \leq \epsilon \log n$, are vertex disjoint. We estimate how long it takes a random walk to visit all isolated cycles of size at most $\epsilon \log n$.

Let N_k denote the number of cycles of length k in G. Then $\mathbf{E} N_k = \theta_k r^k / k$ for some $\theta_k > 0$ constant, and

$$\mathbf{Pr}(N_k > 2^k \mathbf{E} N_k) \leq 2^{-k}.$$

Let $\omega \to \infty$ arbitrarily slowly. For all $\omega \leq k \leq \epsilon \log n$, with probability $1 - O(2^{-\omega})$ we have $N_k \leq 2^k \mathbf{E} N_k$. Using Lemma 13, and (P1), we have that the probability some cycle length k is not visited by a random walk before step $t = cn$ is at most

$$\sum_{k \geq \omega} \theta_k (2r)^k e^{-akt/n} \leq e^{-a'\omega} = o(1),$$

for some $a', a, c > 0$ constant.

Let $3 \leq k < \omega$. Then $\mathbf{Pr}(N_k \geq \omega^2 \mathbf{E} N_k) \leq 1/\omega^2$. The probability some cycle is not visited by a random walk before step $t = n\omega'$ is at most

$$\sum_{k=3}^{\omega} \omega^2 \theta_k r^k e^{-akt/n} \to 0,$$

provided $a\omega' \geq 2\omega \log r + 3 \log \omega$.

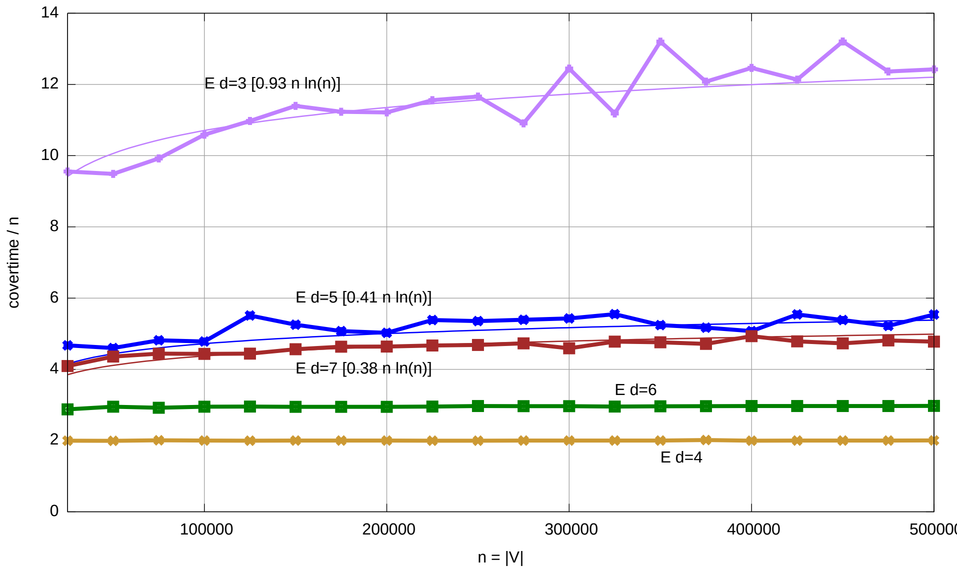

Figure 1: Normalised cover time of E-process on d-regular graphs as function of $n = |V|$

5 Removing the even degree constraint?

The only place where we used the even degree assumption we in Observation 10 which proves that a walk on unvisited edges must terminates at the start vertex. How important is the even degree constraint?

We consider the experimental evidence for the performance of the E-process on both even degree, and odd degree regular graphs. In our experiments unvisited edges are chosen uniformly at random. For a wider range of experimental results on the cover time of regular graphs by random walks which prefer unvisited edges or vertices see [4].

To conduct our experiments, we generated graphs of size up to half a million vertices, using the random regular graph generator from the *NetworkX* package (`http://networkx.lanl.gov/`) for the programming language Python. This package implements the Steger/Wormald approach, see [15]. We used Python's built-in random number generator which is based upon the Mersenne Twister. Each data point is the average of five actual experiments.

In Figure 1 we plot the normalised cover time of the E-process, in the case where the choice of unvisited edges is random. The *normalised* cover time is the actual cover time divided by n, as a function of n. Thus, linear functions of n appear flat. The labeling on the graphs is as follows: The first letter indicates an E-process, and this is followed by the degree $d = r$ of the graph. In the case where the plot appears to be non-linear, a curve of the form $c \log n$, is drawn behind the normalised experimental data, and labeled $[cn \ln(n)]$. The constant c used to draw the curve was determined by inspection.

It would appear the plots for even degrees 4 and 6 are constant, i.e. the cover time is $O(n)$. On the basis of experimental evidence, the normalised cover time of 3-regular graphs is $\omega(n)$;

see Figure 1. This $\omega(n)$ growth appears to be $0.93n \log n$. For degrees 5 and 7 the plot also appears to show logarithmic growth. We note, however, that it is notoriously difficult to quantify such growth on the basis of finite n, and we make no claims other than to present our experiments.

We give an intuitive argument to suggest why the cover time is $\Omega(n \log n)$ when r is odd. We use the notation blue walk, to mean the walk on unvisited edges, and red walk to mean the random walk on visited edges. When $r = 3$ there is a set of isolated vertices I of expected size $|I| \sim n/8$, left behind by the blue walk. This can be seen as follows. Fix a vertex v, and assume that v is tree-like to some fixed depth. All but $o(n)$ vertices satisfy this condition **whp**. The probability that a random blue walk turns away from v each time it visits $N(v)$ is $(1/2)^3$. If this occurs then v is at the center of an isolated blue star $\{v, w, x, y\}$. Let I be the set of such stars. By a coupon collecting argument, it should take $\Omega(n \log n)$ steps for the red walk to visit all of I.

References

[1] D. Aldous. *Lower Bounds for Covering Times for Reversible Markov Chains and Random Walks on Graphs.* Journal of Theoretical Probability, 2, 91-100, (1989).

[2] D. Aldous, J. Fill. *Reversible Markov Chains and Random Walks on Graphs.* (2001). `http://stat-www.berkeley.edu/users/aldous/RWG/book.html`

[3] C. Avin and B. Krishnamachari. *The power of choice in random walks: An empirical study.* Proceedings of 9th ACM/IEEE International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems, (MSWiM-06), 219-228, (2006).

[4] P. Berenbrink, C. Cooper and T. Friedetzky. *Speeding up random walks by choosing unvisited edges or vertices.* ArXiv. (2012)

[5] C. Cooper, D. Ilcinkas, R. Klasing and A. Kosowski. *Derandomizing Random Walks in Undirected Graphs Using Locally Fair Exploration Strategies.* Distributed Computing 24: 91-99 (2011)

[6] J. Cooper, B. Doerr, T. Friedrich, and J. Spencer. *Deterministic random walks on regular trees.* Proceedings of 19th ACM-SIAM Symposium on Discrete Algorithms (SODA'08), 766–772, (2008).

[7] J. Ding, J. Lee and Y. Peres. *Cover times, blanket times, and majorizing measures.* Proceedings of STOC 2011, 61–70, (2011).

[8] U. Feige, *A tight lower bound for the cover time of random walks on graphs.* Random Structures and Algorithms 6 433–438, (1995).

[9] J. Friedman. *A proof of Alon's second eigenvalue conjecture.* Proceedings of 35th ACM Symposium on Theory of Computing, 720–724, (2003).

[10] J. Kahn, J. H. Kim, L. Lovasz, and V. H. Vu. *The cover time, the blanket time, and the Matthews bound.* Proceedings of FOCS 2000, 467–475, (2000).

[11] A. Lubotzky, R. Phillips, and P. Sarnak, *Ramanujan graphs.* Combinatorica, 8, 261–277, (1988).

[12] P. Matthews. *Some sample path properties of a random walk on the cube.* J. Theoret. Probab. 2.1, 129-146, (1989).

[13] T. Orenshtein and I. Shinkar. *Greedy random walk.* `arXiv:1101.5711v3 [math.PR]`

[14] T. Radzik. Private communication. (2010).

[15] A. Steger and N.C. Wormald. *Generating random regular graphs quickly.* Combinatorics, Probability & Computing 8(4): 377-396 (1999).

[16] V. Yanovski, I. A. Wagner, and A. M. Bruckstein. *A distributed ant algorithm for efficiently patrolling a network.* Algorithmica, 37:165–186, (2003).

[17] D. Zuckerman. *A technique for lower bounding the cover time.* SIAM J. Disc. Math, 5, 81-87, (1992).