Combinatorics, Probability and Computing (2014) 23, 269–289.
doi:10.1017/S0963548313000552

Greedy Random Walk

TAL ORENSHTEIN and IGOR SHINKAR

Department of Mathematics, Weizmann Institute of Science, Rehovot, Israel
(e-mail: tal.orenshtein@weizmann.ac.il and igor.shinkar@weizmann.ac.il)

Received 9 August 2012; revised 25 September 2013; first published online 20 November 2013

We study a discrete time self-interacting random process on graphs, which we call greedy random walk. The walker is located initially at some vertex. As time evolves, each vertex maintains the set of adjacent edges touching it that have not yet been crossed by the walker. At each step, the walker, being at some vertex, picks an adjacent edge among the edges that have not traversed thus far according to some (deterministic or randomized) rule. If all the adjacent edges have already been traversed, then an adjacent edge is chosen uniformly at random. After picking an edge the walker jumps along it to the neighbouring vertex. We show that the expected edge cover time of the greedy random walk is linear in the number of edges for certain natural families of graphs. Examples of such graphs include the complete graph, even degree expanders of logarithmic girth, and the hypercube graph. We also show that GRW is transient in $\mathbb{Z}^d$ for all $d \geqslant 3$.

2010 *Mathematics subject classification*: Primary 05C85
Secondary 60K35

1. Introduction

A *greedy random walk* (GRW) on a graph is a discrete time random process, with transition law defined as follows. The walker is located initially at some vertex of the graph. As time evolves each vertex in the graph maintains the set of all adjacent edges that the walker has not yet crossed. At each step the walker picks an unvisited edge among the edges adjacent to its current location arbitrarily according to some rule. If all the adjacent edges have already been visited, an adjacent edge is picked uniformly at random. The walker then jumps to a neighbouring vertex along the chosen edge. We think of the process as trying to cover the graph as rapidly as possible by using a greedy rule that prefers to walk along an unvisited edge whenever possible. This suggests the name *greedy random walk*.

Formally, for an undirected graph $G = (V, E)$, a GRW with a (possibly randomized) rule $\mathcal{R}$ on G is a sequence $X_0, X_1, X_2, \dots$ of random variables defined on V with the following transition probabilities. For each $t \geqslant 0$, define

$$H_t = \{(X_{s-1}, X_s) \in E : 0 < s \leqslant t\} \tag{1.1}$$

to be the set of all the edges traversed by the walk up to time t. For every vertex $v \in V$ and time $t \geqslant 0$, define

$$J_t(v) = \{e \in E : v \in e \text{ and } e \notin H_t\} \tag{1.2}$$

to be the set of all the edges touching v that have not been traversed by the walk up to time t. Letting N_v denote the set of neighbours of v in G, the transition probabilities are given by

$$\mathbb{P}[X_{t+1} = w|(X_i)_{i\leqslant t}] = \begin{cases} \mathcal{R}(w|(X_i)_{i\leqslant t}) & J_t(X_t) \neq \emptyset \text{ and } \{X_t, w\} \in J_t(X_t), \\ \frac{1}{|N_{X_t}|} & J_t(X_t) = \emptyset \text{ and } w \in N_{X_t}, \\ 0 & \text{otherwise,} \end{cases}$$

where $\mathcal{R}(w|(X_i)_{i\leqslant t})$ denotes the probability of choosing $w \in N_{X_t}$ conditioned on the information regarding the process so far. A natural rule $\mathcal{R}$ is to choose uniformly at random an edge among the adjacent unvisited edges $J_t(v)$ of the current vertex $v = X_t$. We shall denote this rule by $\mathcal{R}_{\mathsf{RAND}}$.

One can think of GRW as a random walk where the walker wishes to cover the graph as fast as possible and is allowed to make some local computation at each vertex she visits (*e.g.*, mark the last edge that the walker used to reach the current vertex, and also mark the edge that the walker is going to use in the next step), but is not allowed to transfer information between vertices. A motivation for the study of GRW arises from distributed computation in which an agent sits on every vertex of a graph. Each agent has a list of neighbours and is allowed to communicate only with them. The goal is to let all the agents use some resource as fast as possible, while using only the local information for each vertex, and no extra information regarding the graph and the vertices that have already been visited. An agent has a list of neighbours who have communicated with her thus far during the process, and each time the agent receives the resource, she is allowed to perform only local computations before moving it to one of her neighbours. We will see that the GRW protocol performs better than a simple random walk (SRW) on some families of graphs.

The main difficulty in analysing this random process comes from the fact that GRW is self-interacting, *i.e.*, it is not a Markov chain (meaning that the probability distribution of the next step depends not only on the current position of the walker, but also on the entire walk thus far). Although in many cases a certain property of self-interacting random walks can be observed in simulations or seems to be suggested by 'heuristical proof', typically it is much harder to give robust proofs for random walks that do not have the Markov property. Related models include RW with choice [5], non-backtracking RW [4], RW with neighbourhood exploration [8], excited RW [7], reinforced RW [17], rotor router RW [14], and more. Recently this model has been considered independently by Berenbrink, Cooper and Friedetzky [9]. They showed that if G is an even degree expander graph such that every vertex is contained in a vertex-induced cycle of logarithmic length, then the expected vertex cover time by GRW is linear for any rule $\mathcal{R}$.

Our results. In Section 2 we study the edge cover time of GRW on finite graphs. Obviously, the edge cover time of any graph $G = (V, E)$ is at least $|E|$, as the walker must cross every edge at least once. We prove bounds on the edge cover time of GRW by analysing the 'overhead' of the walk, *i.e.*, the difference between the expected edge cover time of the walk, and the number of edges in a graph. For example, we establish that the expected time it takes for GRW to go via all edges of K_n, the complete graph on n vertices, is $\binom{n}{2} + (1+o(1))n\log(n)$. Therefore, the aforementioned 'overhead' in the case of K_n is $(1+o(1))n\log(n)$. In particular, all edges of K_n are covered by GRW in time $(1+o(1))\cdot\binom{n}{2}$, which is asymptotically faster than $\Theta(n^2\log n)$, the expected edge cover time of SRW.

We show that for certain families of graphs the expected edge cover time of GRW is asymptotically faster than that of SRW. In particular, we establish that expected edge cover time of GRW is *linear* in the number of edges for the complete graph, for the hypercube graph, and for constant even degree expanders with logarithmic girth. The latter result is claimed in the paper of Berenbrink, Cooper and Friedetzky [9].

Another interesting result is given in Lemma 2.9 that bounds the *edge* cover time of an even degree graph by GRW in terms of its *vertex* cover time by SRW. Specifically, we show that for any graph $G=(V,E)$ whose vertices have even degrees and whose expected *vertex* cover time by SRW is C, the expected *edge* cover time of G using GRW is at most $|E|+C$. Therefore, for even degree graphs of logarithmic degree whose vertex cover time is $O(n\log(n))$, we obtain a bound on the edge cover time which is linear in the number of edges.

These results should be compared with the general lower bound on the expected cover time of graphs by SRW. Recall that Feige [13] has shown that for any graph with n vertices the expected vertex cover time by a simple random walk is at least $(1-o(1))n\log n$. Analogously, for all graphs the expected edge cover is at least $\Omega(|E|\log(|E|))$ (see [20], [1]). In this direction, a result of Benjamini, Gurel-Gurevich and Morris [6] says that for bounded degree graphs linear cover time is exponentially unlikely.

We are also interested in the behaviour of GRW on infinite graphs. It is well known that SRW on $\mathbb{Z}^d$ is transient if $d\geqslant 3$, and recurrent otherwise. We prove that GRW is transient on $\mathbb{Z}^d$ for $d\geqslant 3$. The case of $d=2$ remains open, and it is shown to be equivalent to the notorious two-dimensional mirror model problem [18, 12]. Our proof holds for all graphs with even degrees on which SRW is transient. This leaves unsolved the question of transience of GRW in lattices with odd degrees. These and other related results are discussed in Section 3, which can be read independently of the rest of the paper.

We need to make two general remarks.

Choice of the rule $\mathcal{R}$. In the first version of this paper we considered GRW that uses only the rule $\mathcal{R}_{\mathsf{RAND}}$. After our work was uploaded to `arxiv.org`, Berenbrink, Cooper and Friedetzky [9] independently published their work in which they consider GRW with *any* (deterministic or randomized) rule, even adversarial ones that try to slow the process down. After reading their results, we noticed that in fact our proofs for bounding from above the edge cover time are independent of $\mathcal{R}$ and hold for any rule as well.

Choice of the starting vertex. In all of our results on cover time the bounds are independent of the starting vertex. Also, in most cases the considered graphs are vertex-transitive, and therefore specification of the starting vertex is unnecessary.

Notation. We use the standard notations of asymptotic growth rates. For two functions $f, g : \mathbb{N} \to \mathbb{R}_+$, we write $f = O(g)$ when there is a positive constant $C \in \mathbb{R}$ such that $f(n) < Cg(n)$ for all sufficiently large values of n. The notation $f = \Omega(g)$ means that there is a positive constant $c > 0$ such that $f(n) > cg(n)$ for all sufficiently large values of n, and $f = \Theta(g)$ means both $f = O(g)$ and $f = \Omega(g)$. We write $f = o(g)$ if

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0.$$

2. Edge cover time of finite graphs

In GRW the choice of the next move depends on the history of the walk with respect to the adjacent edges of the current vertex. Hence, it seems more natural to ask about the edge cover time rather than the vertex cover time. We show that for some common families of graphs the greedy walk covers the edges asymptotically faster than the simple random walk.

Let $G = (V, E)$ be a connected undirected graph on n vertices. Let $C_E(G)$ denote the edge cover time of GRW, *i.e.*, the number of steps it takes for GRW to traverse all edges of G. Note that since the graph G is finite, the edge cover time $C_E(G)$ is almost surely finite.

The basic idea behind the analysis is as follows. Divide the random discrete time interval $[0, C_E(G)]$ into two (random) parts.

(1) The greedy part: all times for which the walker is at a vertex that has an adjacent edge yet to be covered, *i.e.*, all times $t \in [0, C_E(G)]$ such that $\{X_t, X_{t+1}\} \notin H_t$.

(2) The simple part: all times for which the walker is positioned at a vertex all of whose adjacent edges have already been covered previously, *i.e.*, all times $t \in [0, C_E(G)]$ such that $\{X_t, X_{t+1}\} \in H_t$. For these times the choice of the next move has the same distribution as that of a simple random walk.

Roughly speaking, the GRW typically looks as follows. It starts at $t_0 = 0$ in a greedy time part. This time part lasts until the walk reaches, at time s_1, a vertex v_1 all of whose adjacent edges have already been covered. In this situation we say that the walk has 'got stuck'. This means that the last step before it got stuck covered the last edge touching v_1. Since, at time s_1, all edges touching v_1 have already been covered, the walker picks an edge at random among these edges. In other words the walk is now in a simple time part, which started at time s_1. This time part lasts until the walk reaches, at time t_1, a vertex u_1 that has an adjacent edge which has not yet been covered. By definition, the next step will belong to a greedy part, and will continue until the walk reaches, at time s_2, some vertex v_2 all of whose adjacent edges have already been covered, thus starting the second simple part. The walk continues in this way until all edges are covered, and then it becomes a simple random walk.

Formally, define the times $t_0, s_1, t_1, s_2, t_2, \ldots, s_n$ recursively, where the intervals $[t_{i-1}, s_i)$ denote the ith greedy part, and the intervals $[s_i, t_i)$ denote the ith simple part of the walk:

$$t_0 = 0,$$

$$s_{i+1} = \begin{cases} \inf\{t_i < t \leqslant C_E(G) : J_t(X_t) = \emptyset\} & \text{if there is such a } t, \\ C_E(G) & \text{otherwise,} \end{cases}$$

$$t_{i+1} = \begin{cases} \inf\{s_{i+1} < t \leqslant C_E(G) : J_t(X_t) \neq \emptyset\} & \text{if there is such a } t, \\ C_E(G) & \text{otherwise.} \end{cases}$$

We say the walk has *got stuck* at time t if $t = s_i$ for some $i \in \mathbb{N}$. It should be clear from the description that the vertices X_{s_i} must all be distinct, since X_{s_i} is the ith time that the walk got stuck, and it is impossible to get stuck at the same vertex twice. Therefore, it is enough to define the times t_i and s_i only for $i \leqslant n$ (where n denotes the number of vertices in G). This gives a random partition $(0 = t_0 < s_1 < t_1 < s_2 < t_2 < \cdots < t_{k-1} < s_k = t_k = \cdots = s_n = C_E(G))$ of the time segment $[0, C_E(G)]$, where the random variable $k \leqslant n$ is the first i for which $s_i = C_E(G)$, *i.e.*, all edges of G are covered.

Note that the total time the walker spends in the greedy parts is equal to the number of edges $|E|$, implying the following expression for the edge cover time:

$$C_E(G) = |E| + \sum_{i=1}^{n} (t_i - s_i).$$

By linearity of expectation we have the following simple expression for the expected edge cover time, which will be the key formula in our proofs.

Proposition 2.1 (key formula). *Let $G = (V, E)$ be a graph with n vertices, and let $t_0, s_1, t_1, s_2, t_2, \ldots$ be random times as above. Then, the expected edge cover time of GRW on G is*

$$\mathbb{E}[C_E(G)] = |E| + \sum_{i=1}^{n} \mathbb{E}[t_i - s_i]. \tag{2.1}$$

Thus, in order to bound $\mathbb{E}[C_E(G)]$, it is enough to bound the expected total size of all simple parts, *i.e.*, $\mathbb{E}[\sum_{i=1}^{k} (t_i - s_i)]$. In order to apply Proposition 2.1 the following notation will be convenient. For $i = 1, \ldots, n$ let

$$B_i = \{v \in V : J_{s_i}(v) = \emptyset\}$$

be the set of vertices all of whose adjacent edges are covered by time s_i. (Here B stands for 'bad'. If the walker is in some vertex in B, then the next step will be along an edge that has already been crossed, thus increasing the edge cover time.) By the definition of s_i and t_i, we note that $B_i = \{v \in V : J_t(v) = \emptyset\}$ for every $t \in [s_i, t_i]$. Note also that $B_i \subseteq B_j$ for all $i < j$, and the vertex $v_j = X_{s_j}$ at which the walker got stuck at time s_j does not belong to B_i for $i < j$, since at any time $t < s_j$ the vertex v_j still had an adjacent edge which had not yet been covered. Thus the containment $B_i \subsetneqq B_j$ is strict for all $i < j \leqslant k$,

https://doi.org/10.1017/S0963548313000552 Published online by Cambridge University Press

i.e., the sets B_i form a strictly increasing chain until it stabilizes at $B_k = V$:

$$B_1 \subsetneq B_2 \subsetneq \cdots \subsetneq B_k = B_{k+1} = \cdots = B_n = V. \tag{2.2}$$

In particular,

$$|B_i| < n \text{ if and only if } i < k. \tag{2.3}$$

Conditioned on B_i and X_{s_i}, the length of the time segment $[s_i, t_i]$ is distributed as the escape time of a simple random walk from B_i, when started at X_{s_i}. That is, conditioned on B_i and X_{s_i}, the random variable $(t_i - s_i)$ has the same distribution as $T(X_{s_i}, B_i)$, where

$$T(v, B) = \min\{t : Y_t \notin B | Y_0 = v\},$$

and $Y_0, Y_1, \ldots$ is a simple random walk on G that started at $Y_0 = v$. By applying known bounds of the expected escape time of SRW, we shall use Proposition 2.1 to upper-bound the expected edge cover time of GRW.

2.1. The complete graph

We prove in this section that for the complete graph with n vertices the expected edge cover time is $(1 + o(1))\binom{n}{2}$. Specifically, we prove the following result.

Theorem 2.2. *For any rule $\mathcal{R}$ the expected edge cover time of GRW on K_n is bounded by*

$$\mathbb{E}[C_E(K_n)] \leqslant |E| + (1 + o(1))n \log n.$$

This is an improvement over the $\Theta(n^2 \log n)$ time of the SRW, which follows from the coupon collector argument.

Proof. Consider the complete n-vertex graph $G = K_n$. The proof relies on the following simple observation. For any set of vertices $B \subseteq V$, the escape time of SRW from B depends only on the size of B, and has geometric distribution. Specifically, for each $i = 1, \ldots, n$, the quantity $t_i - s_i$ *conditioned on* B_i is distributed geometrically:

$$t_i - s_i \sim \begin{cases} G\left(\frac{n-|B_i|}{n-1}\right) & \text{if } |B_i| < n, \\ 0 & \text{otherwise.} \end{cases} \tag{2.4}$$

Let T_i denote the expected escape time from the subset B_i. Then,

$$T_i = \mathbb{E}(t_i - s_i | B_i) = \begin{cases} \frac{n-1}{n-|B_i|} & \text{if } |B_i| < n, \\ 0 & \text{otherwise.} \end{cases} \tag{2.5}$$

By averaging over the B_i, the quantity $\sum_{i=1}^n \mathbb{E}[t_i - s_i]$ is equal to

$$\sum_{i=1}^n \mathbb{E}[t_i - s_i] = \sum_{i=1}^n \mathbb{E}[\mathbb{E}(t_i - s_i | B_i)] = \sum_{i=1}^n \mathbb{E}[T_i] = \mathbb{E}\left[\sum_{i=1}^{k-1} \frac{n-1}{n-|B_i|}\right],$$

where the last equality follows from linearity of expectation, together with (2.5). In order to bound the sum in the expectation, let $b_i = |B_i|$, and note that we have an increasing

https://doi.org/10.1017/S0963548313000552 Published online by Cambridge University Press

sequence of natural numbers $b_1 < b_2 < \cdots < b_k$ so that $b_1 \geqslant 1$ and $b_k = n$ for some $k \leqslant n$. For any such sequence it holds that

$$\sum_{i=i}^{k-1} \frac{n-1}{n-b_i} \leqslant \sum_{i=i}^{n-1} \frac{n-1}{n-i}. \tag{2.6}$$

To see this, note that all summands are positive, and each one on the left-hand side of the inequality also appears on the right-hand side. Therefore, we can upper-bound the quantity $\sum_{i=1}^{n} \mathbb{E}[t_i - s_i]$ by

$$\sum_{i=1}^{n} \mathbb{E}[t_i - s_i] \leqslant \sum_{i=1}^{n-1} \frac{n-1}{n-i} = (1+o(1))n\log n.$$

Applying Proposition 2.1 gives the desired result. □

Remark. We conjecture that if the rule in the greedy part is $\mathcal{R}_{\mathsf{RAND}}$ (in which an edge is chosen uniformly at random among the adjacent unvisited edges of the current vertex), then for odd values of n, *i.e.*, when the degree is even, the overhead for cliques is $O(n)$, *i.e.*, $\mathbb{E}[C_E(K_n)] \leqslant |E| + O(n)$. For a related discussion see Section 4.

2.2. Expander graphs

We apply the same method as in the previous section on expander graphs. Let $G = (V,E)$ be a d-regular graph on n vertices and let $A = A(G) \in \{0,1\}^{V\times V}$ be its normalized adjacency matrix, namely

$$A(u,v) = \begin{cases} 1/d & (u,v) \in E, \\ 0 & (u,v) \notin E. \end{cases}$$

It is a standard fact that A has real eigenvalues, all lying in the interval $[-1,1]$. Denote the eigenvalues by $1 = \lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n \geqslant -1$, and let $\lambda(G)$ be the spectral radius of G, defined by

$$\lambda(G) = \max_{i=2,\ldots,n} |\lambda_i|.$$

We say that a d-regular graph G is an (n,d,λ)-expander if $\lambda(G) < \lambda < 1$ (for more details see the excellent survey by Hoory, Linial and Wigderson [15]).

We are able to show that for $d = \Omega(\log n)$, the expected edge cover time of the GRW is linear in the number of edges. This is faster than a simple random walk, which covers the edges in $\Omega(|E|\log|E|)$ steps, as mentioned in the Introduction. Specifically, we prove the following theorem.

Theorem 2.3. *Let G be a (n,d,λ)-expander graph. Then, for any rule $\mathcal{R}$ the expected edge cover time is*

$$\mathbb{E}[C_E(G)] \leqslant |E| + O\left(\frac{n\log n}{1-\lambda}\right).$$

In particular, for an expander with $d = \Omega(\log n)$, the expected edge cover time of the GRW is linear in the number of edges.

Proof. The key observation here is that, as in the case of the complete graph, $\mathbb{E}(t_i - s_i | B_i)$ can be bounded in terms of the *size* of B_i, independently of its structure. We use the following lemma of Broder and Karlin [11].

Lemma 2.4 ([11, Lemma 3]). *Let G be an (n, d, λ)-expander and let $S \subsetneq V$ be a non-empty set of vertices. Consider a simple random walk $Y_0, Y_1 \ldots$ on G, starting at some $v \in S$ (i.e., $Y_0 = v$). Let $T(v, S)$ be the escape time of the walk from S when started from v. Then*

$$\mathbb{E}[T(v,S)] \leqslant \frac{C}{1-\lambda}\left(\log n + \frac{n}{n-|S|}\right)$$

for some absolute constant C.

Denoting by T_i the expected escape time from the subset B_i, by Lemma 2.4, for all $i = 1, \ldots, n$ we have

$$T_i := \mathbb{E}(t_i - s_i | B_i) \leqslant \begin{cases} \frac{C}{1-\lambda}\left(\log n + \frac{n}{n-|B_i|}\right) & \text{if } |B_i| < n, \\ 0 & \text{otherwise,} \end{cases} \tag{2.7}$$

for some absolute constant $C \in \mathbb{R}$. In order to upper-bound $\sum_{i=1}^{n} \mathbb{E}[t_i - s_i]$, we apply an analysis similar to that in the proof of Theorem 2.2. Specifically, by averaging over the B_i, the quantity $\sum_{i=1}^{n} \mathbb{E}[t_i - s_i]$ is equal to

$$\sum_{i=1}^{n} \mathbb{E}[t_i - s_i] = \sum_{i=1}^{n} \mathbb{E}[\mathbb{E}(t_i - s_i|B_i)] = \sum_{i=1}^{n} \mathbb{E}[T_i] = \mathbb{E}\left[\sum_{i=1}^{n} T_i\right],$$

where the last equality follows from linearity of expectation. Using (2.7) we obtain

$$\begin{aligned} \sum_{i=1}^{n} \mathbb{E}[t_i - s_i] &\leqslant \mathbb{E}\left[\sum_{i=1}^{k-1} \frac{C}{1-\lambda}\left(\log n + \frac{n}{n-|B_i|}\right)\right] \\ &\leqslant \frac{C}{1-\lambda} \cdot n\log(n) + \frac{C}{1-\lambda}\mathbb{E}\left[\sum_{i=1}^{k-1} \frac{n}{n-|B_i|}\right] \\ &\leqslant O\left(\frac{n\log n}{1-\lambda}\right), \end{aligned}$$

where the bound

$$\sum_{i=1}^{k-1} \frac{n}{n-|B_i|} \leqslant O(n\log(n))$$

in the last inequality follows using the same proof as (2.6). Using Proposition 2.1, we have

$$\mathbb{E}[C_E(G)] \leqslant |E| + \sum_{i=1}^{n} \mathbb{E}[t_i - s_i] = |E| + O\left(\frac{n\log n}{1-\lambda}\right),$$

which completes the proof of the theorem. □

Next, we strengthen Theorem 2.3 by showing that for constant degree expanders with logarithmic girth whose vertices have even degrees, the expected edge cover time is linear in the number of vertices. Recall that girth of a graph G, denoted by $\mathrm{girth}(G)$, is the minimal length of a cycle in G. This result is claimed in [9] without proof.

Theorem 2.5. *Let G be a (n,d,λ)-expander graph such that $d \in \mathbb{N}$ is even, and* $\mathrm{girth}(G) = g$. *Then, for any rule $\mathcal{R}$ the expected edge cover time is*

$$\mathbb{E}[C_E(G)] \leqslant |E| + O\left(|E| \cdot \frac{\log(n)}{(1-\lambda)g}\right).$$

In particular, if $G = (V,E)$ is an expander of constant even degree with $\mathrm{girth}(G) = \Omega(\log(n))$, *then the expected edge cover time of the GRW is linear in the number of vertices.*

The proof relies on the following simple observation. Suppose that the ith greedy part starts at some vertex $v = X_{s_i}$. Then, using the fact that all degrees of G are even, we conclude that this greedy part will end at the same vertex v. Indeed, by an Euler-path-type argument, if a vertex has even degree and the walker entered this vertex along a new edge that has not yet been visited, then by parity there must be another unvisited edge for the walker to leave the vertex. In particular, the range covered by each greedy part forms a (not necessarily simple) cycle. We summarize this observation as follows.

Observation 2.6. *If the all degrees of a graph $G = (V,E)$ are even, then in each greedy time part $[t_i, s_{i+1}]$ it holds that $X_{t_i} = X_{s_{i+1}}$, i.e., every greedy part ends at the same vertex it started from.*

Therefore, since in the greedy time parts the walker crosses no edge twice, in each greedy part $[t_i, s_{i+1}]$ the walker traverses along some (not necessarily simple) cycle, and thus the number of steps in each greedy time part is at least $\mathrm{girth}(G)$.

We now turn to the proof of Theorem 2.5.

Proof. As in the proof of Theorem 2.3, the expected edge cover time can be bounded from above by

$$\mathbb{E}[C_E(G)] = |E| + O\left(\frac{1}{1-\lambda}\right) \cdot \mathbb{E}\left[\sum_{i=1}^{k-1}\left(\log n + \frac{n}{n-|B_i|}\right)\right]. \tag{2.8}$$

By Observation 2.6 it follows that the random number k of greedy parts is upper-bounded by $|E|/g$.

In order to bound the terms $n/(n-|B_i|)$, note that for all $i \leqslant k$ we have

$$k \leqslant i + \frac{d \cdot (n-|B_i|)}{g}.$$

Indeed, if in time s_i the number of vertices all of whose adjacent edges have already been covered is $|B_i|$, then the number of edges that have not yet been traversed is at most $d \cdot (n-|B_i|)$, and hence, by the assumption on the girth of G, the number of remaining

https://doi.org/10.1017/S0963548313000552 Published online by Cambridge University Press

greedy parts is at most

$$\frac{d\cdot(n-|B_i|)}{g}.$$

Therefore, for all $i\leqslant k$ we have

$$\frac{n}{n-|B_i|}\leqslant\frac{dn}{(k-i)g}=\frac{2|E|}{(k-i)g}.$$

By (2.8) we have

$$\begin{aligned}\mathbb{E}[C_E(G)] &= |E| + O\left(\frac{1}{1-\lambda}\right)\cdot\mathbb{E}[k\log(n)] + O\left(\frac{1}{1-\lambda}\right)\cdot\left(\sum_{i=1}^{k-1}\frac{n}{n-|B_i|}\right)\\ &= |E| + O\left(|E|\cdot\frac{\log(n)}{(1-\lambda)g}\right) + O\left(\frac{1}{1-\lambda}\right)\cdot\left(\sum_{i=1}^{k-1}\frac{|E|}{(k-i)\cdot g}\right)\\ &\leqslant |E| + O\left(|E|\cdot\frac{\log(n)}{(1-\lambda)g}\right),\end{aligned}$$

where the last inequality uses the assumption that $k\leqslant n$ and the facts that

$$\sum_{i=1}^{k-1}\frac{1}{k-i}\leqslant\log(k)+O(1).$$

Theorem 2.5 follows. □

We show below that the assumption that a graph has logarithmic girth in Theorem 2.5 is necessary. Specifically, we present a 6-regular expander graph G, and a rule $\mathcal{R}$, such that GRW with the rule $\mathcal{R}$ coves all the edges of G in expected time $\Omega(n\log(n))$. In fact, the graph G satisfies an additional property that every vertex of G is contained in some induced cycle of logarithmic length. This should be compared with the result of Berenbrink, Cooper and Friedetzky [9], who have shown that if G is an even degree expander such that every vertex of G is contained in some induced cycle of logarithmic length, then the expected *vertex* cover time by GRW is linear for any rule $\mathcal{R}$. This shows a gap between the edge cover time and the vertex cover time of GRW.

Theorem 2.7. *For every $n=0 \pmod 3$ there exists a 6-regular expander graph $G=(V,E)$ with $|V|=n$ vertices such that every vertex of G is contained in an induced cycle of logarithmic length, and there exists a rule $\mathcal{R}$ such that the expected edge cover time of G by GRW with the rule $\mathcal{R}$ is $\Omega(n\log(n))$.*

Proof. Let $H=(U,F)$ be a 4-regular expander graph on $n/3$ vertices such that every vertex of G is contained in an induced cycle of length $\epsilon\log(n)$ for some constant $\epsilon>0$.[1] Define a graph $G=(V,E)$ to be the Cartesian product of H with the graph K_3. Namely, the vertices of G are $V=U\times\{1,2,3\}$ and $((u,i),(u',j))\in E$ if and only if either (i) $(u,u')\in F$

[1] Such a graph can be obtained by choosing a random 4-regular graph. For reference see [10, Chapter II.4].

https://doi.org/10.1017/S0963548313000552 Published online by Cambridge University Press

and $i = j$ or (ii) $u = u'$ and $i \neq j$. By the properties of H, the graph G is a 6-regular expander and it satisfies the property that every vertex of G is contained in some induced cycle of length at least $\epsilon \log(n)$.

The vertices of G are naturally partitioned into three subsets $V = V_1 \cup V_2 \cup V_3$ where $V_i = \{(U,i) : u \in U\}$ for $i = 1,2,3$. The rule $\mathcal{R}$ is defined so that the first greedy part will cover all edges of the form $((u,i)(v,i))$ for all $(u,v) \in F$ and $i \in \{1,2,3\}$. Assume now that GRW starts from some arbitrary vertex $(u_0,1) \in V_1$. The walker walks along some Eulerian cycle of V_1, covering all edges induced by V_1. Indeed, this can be done, as the graph induced by V_1 is isomorphic to H, and hence its vertices have even degrees. After completing the cycle in V_1 and returning to the initial vertex $(u_0,1)$, the walker moves to $(u_0,2)$, performs a walk along some Eulerian cycle on V_2, and returns to $(u_0,2)$. Similarly, the walker then moves to $(u_0,3)$, covers all edges induced by V_3, and returns to $(u_0,3)$. Finally, the walker moves back to $(u_0,1)$, and gets stuck for the first time. Note that at this point all edges induced by each of the V_i have already been covered by GRW, and the remaining edges form disjoint triangles of the form $\{(u,1),(u,2),(u,3)\}$ induced by each of the vertices $u \in U \setminus \{u_0\}$. Hence, each subsequent greedy part will consist of three steps, covering one triangle at each part, and the order is defined by the first time that the SRW reaches some vertex of a triangle $\{(u,i) : i = 1,2,3\}$. Noting that by ignoring all steps from (u,i) to (u,j) in G, the random walk on G induces a simple random walk on H, it follows that in order to cover all triangles, the SRW needs to cover all the vertices of a copy of H. Since by the theorem of Fiege [13] the expected vertex cover time of every graph by SRW is at least $\Omega(n\log(n)$, this bound also holds for the edge cover time of G. This completes the proof of the theorem. □

2.3. Hypercube $\{0,1\}^d$

The hypercube graph $G = (V,E)$ is a graph whose vertices are $V = \{0,1\}^d$ and $(u,v) \in E$ if and only if $d(u,v) = 1$, where $d(\cdot,\cdot)$ is the Hamming distance between two strings. We show that for even dimension d the edge cover time of the hypercube is linear in the number of edges.

Proposition 2.8. *Let $d \in \mathbb{N}$ be even, and let $Q_d = (V,E)$ be the d-dimensional hypercube graph. Then, for any rule $\mathcal{R}$ the expected edge cover time of Q_d is bounded by*

$$\mathbb{E}[C_E(Q_d)] = O(|E|).$$

□

Proof. The proposition follows from the next lemma.

Lemma 2.9. *Let $G = (V,E)$ be a graph whose vertices have even degrees. Suppose that for the graph G the expected* vertex *cover time of SRW is C. Then, the expected* edge *cover time of GRW of G is at most*

$$\mathbb{E}[C_E(G)] \leqslant |E| + C.$$

https://doi.org/10.1017/S0963548313000552 Published online by Cambridge University Press

Since the number of edges in Q_d is $|E| = \frac{1}{2}d \cdot 2^d$, and using the fact that the expected vertex cover time of the hypercube by SRW is $C = O(d \cdot 2^d)$, Proposition 2.8 follows by Lemma 2.9. □

We now turn to the proof of Lemma 2.9.

Proof of Lemma 2.9. The proof proceeds by coupling between a simple random walk and a greedy random walk so that the number of steps made by the GRW is larger than the number of steps made by the SRW by at most $|E|$.

As observed above in Observation 2.6, for graphs of even degrees we have $X_{t_i} = X_{s_{i+1}}$ for all $i \leqslant k$, *i.e.*, every greedy part finishes at the same vertex that it started from. This implies that the simple parts can be concatenated, as the end of the ith simple part is X_{t_i}, and the beginning of the $(i+1)$st part is $X_{s_{i+1}}$. The coupling between the SRW and the GRW is the natural one, where the SRW performs all the steps that the GRW makes in its simple parts. Clearly, the number of steps made by the GRW is larger than the number of steps made by the SRW by at most the total number of steps made in the greedy parts, which is bounded by $|E|$.

Observe that whenever the SRW reaches some vertex v, it is either the case that (i) all edges adjacent to v have already been covered by the GRW, or (ii) the vertex v is the last vertex in the current simple part, and thus, using the property $X_{t_i} = X_{s_i}$ for all i, the next greedy part will cover all edges adjacent to v. This implies that by the time the SRW covers all vertices of G, the GRW has either already covered all edges of G, or will do so in the number greedy part. Therefore, the *edge* cover time of the GRW is larger than the *vertex* cover time of SRW by at most $|E|$. This completes the proof of the lemma. □

We also remark (without a proof) on the edge cover time of a generalization of the hypercube graph.

Remark. Define a generalization of the hypercube by connecting two vertices in $\{0,1\}^d$ if the distance between them is at most some parameter $\ell \geqslant 2$. Specifically, for $\ell \geqslant 2$, let $Q_d^{(\leqslant \ell)} = (V, E_\ell)$, where $V = \{0,1\}^d$ and $(x,y) \in E$ if and only if $d(x,y) \leqslant \ell$. Denoting the number of vertices in the graph by $n = 2^d$, the spectral radius of $Q_d^{(\leqslant \ell)}$ is bounded from above by $\lambda \leqslant 1 - \ell/\log n$. Therefore, by Theorem 2.3 for $\ell \geqslant 2$ the expected edge cover time of GRW on $Q_d^{(\leqslant \ell)}$ is $|E_\ell| + O(n \log^2 n)$, where the constant in the $O()$ notation depends on ℓ.

Noting that the number of edges in $Q_d^{\leqslant \ell}$ is $|E_\ell| = O(n \cdot \log^\ell n)$, this implies that for $\ell = 2$ the edge cover time is linear in the number of edges $|E_2| = O(n \cdot \log^2 n)$, and for $\ell \geqslant 3$ the edge cover time is $(1 + o(1)|E_\ell|$.

2.4. *d*-regular trees

In this section we provide an upper bound for the edge cover time of GRW on trees. We are able to describe the behaviour of GRW quite accurately, and subsequently provide a tight bound on the cover time.

Theorem 2.10. *Let $G = (V, E)$ be a tree rooted at a vertex denoted by r. For any $v \in V$ let T_v denote the subtree rooted at v and let $|T_v|$ denote the number of edges in T_v. Then, for any rule $\mathcal{R}$ the GRW edge cover time of G is*

$$\mathbb{E}[C_E(G)] = |E| + O\left(\sum_{u \in G \setminus \{r\}} |T_u|\right).$$

If the rule for GRW is $\mathcal{R}_{\mathsf{RAND}}$ and $\deg(r) \geqslant 2$, then there is a matching lower bound, namely

$$\mathbb{E}[C_E(G)] = |E| + \Theta\left(\sum_{u \in G \setminus \{r\}} |T_u|\right).$$

The following corollary is immediate from Theorem 2.10.

Corollary 2.11. *If G is a d-regular tree with n vertices, then the expected edge cover time is $O(n \log_d n)$.*

Comparing Corollary 2.11 to the cover time of SRW on d-regular trees, we again see an asymptotic speed-up over the $\Theta(n \log_d^2 n)$ time of the SRW [2].

Proof. In order to use the tree structure of the graph, let us first give an overview of the behaviour of GRW on trees. The walker starts at the root r and goes down greedily (*i.e.*, an unvisited edge is traversed at every new step), until she reaches a leaf. Since she got stuck at a leaf, she performs a simple random walk until she reaches the lowest ancestor with an adjacent edge that has not yet been covered. The non-covered edge is necessarily from the ancestor to one of the children (as its parent has already been visited on the way down). The walker continues by moving down greedily until she reaches another leaf not covered thus far, and then performs a simple random walk until she again reaches the lowest ancestor with a child that has not been visited thus far by the walk. The walk continues in the same manner until all edges have been covered, getting stuck only at the leaves. In fact the walk gets stuck exactly once at each leaf, and the time $C_E(G)$ is the time when the walker visits the last leaf of the tree. Note that when visiting some vertex v, the walk will cover the entire subtree of v before returning to v's parent. This property is what makes the cover time of GRW asymptotically faster than the cover time of SRW.

The order in which the vertices are visited for the first time defines some preorder traversal on the tree (first the root, then the subtrees), where for each vertex the order of the subtrees is chosen according to the rule $\mathcal{R}$. We observe that the vertices $(X_{s_1}, X_{s_2}, \dots, X_{s_k})$ define some order on the leaves of the tree, induced by the preorder traversal as described above (and in particular, k is equal to the number of leaves). In addition, for every $i < k$, the vertex X_{t_i} is the lowest ancestor of X_{s_i} such that at time s_i not all of its descendants have been visited by the walk. Hence, $\mathbb{E}[t_i - s_i]$ equals the expected time it takes for the simple random walk starting at X_{s_i} to visit this ancestor. This implies that for every edge (u, v), where u is the parent of v, there is at most one $i \in \{1, \dots, k\}$ such that the edge (u, v) lies on the shortest path from X_{s_i} to X_{t_i}. Therefore, if w is the leaf where the walk got stuck for the ith time, that is, $X_{s_i} = w$, and v is its lowest ancestor whose subtree is not yet covered, then the expected time to reach v starting from w is

$$\mathbb{E}[t_i - s_i] = H(w,v) = \sum_{(u_1,u_2)\in P_{(w,v)}} H(u_1,u_2),$$

where $H(x,y)$ denotes the expected number of steps required for SRW starting at x to visit y, and the sum is over all edges on the shortest path from w to v (using the convention that the edge (u_1,u_2) means that u_2 is a parent of u_1).

Going over all leaves in the graph, using the observation that the walk gets stuck at each leaf exactly once (stopping at the last visited leaf at time s_k), and finishing the corresponding simple part at the lowest ancestor whose tree has not yet been covered, we observe that for each $i < k$ the shortest paths from X_{s_i} to X_{t_i} are disjoint. Furthermore, the union of all these paths covers all edges of the graph except for the path from the last covered leaf, denoted by $l = X_{s_k}$, to the root of the tree. Let $P_{(r,l)}$ denote the shortest path from l to r. Then

$$\mathbb{E}\left[\sum_{i=1}^{k}(t_i - s_i)\right] = \mathbb{E}\left[\sum_{(u,v)\in E\setminus P_{(r,l)}} H(u,v)\right] \leqslant \sum_{(u,v)\in E} H(u,v), \tag{2.9}$$

where $H(v,u)$ denotes the expected number of steps required for SRW starting at v to visit u for the first time, and the summation is over all edges (u,v), where v is the parent of u. It is well known (see, *e.g.*, [3, Lemma 1]) that if (u,v) is an edge in a tree, then $H(u,v) = 2|T_u| + 1$. Proposition 2.1, together with (2.9), proves the upper bound of the theorem.

Note that if we allow the walker return to the origin after covering the tree, then the expected return time is equal to

$$|E| + \sum_{(u,v)\in E} (2|T(u)| + 1) = 2\sum_{u\in V} |T(u)| = 2\sum_{u\in V} \mathrm{depth}(u),$$

where depth(u) is the distance of the vertex u from the root.

If GRW uses the rule $\mathcal{R}_{\mathsf{RAND}}$, then the subtrees rooted at the children of r are explored completely one after another (the order of the children is random), and the walk will return to r from all but the last subtree. Therefore, for each u child of r the subtree rooted at u is completely explored by GRW with probability $(\deg(r) - 1)/\deg(r)$, and hence, every edge of the tree belongs to $P_{(r,l)}$ with probability at most $1/\deg(r)$. Therefore, by applying the formula in (2.9) we get

$$\begin{aligned}\mathbb{E}\left[\sum_{i=1}^{k}(t_i - s_i)\right] &= \mathbb{E}\left[\sum_{(u,v)\in E\setminus P_{(r,l)}} H(u,v)\right]\\ &\geqslant \left(1 - \frac{1}{\deg(r)}\right) \sum_{(u,v)\in E} H(u,v)\\ &\geqslant \left(1 - \frac{1}{\deg(r)}\right) \sum_{u\in V\setminus\{r\}} 2|T_u|.\end{aligned}$$

This completes the proof of the theorem. □

https://doi.org/10.1017/S0963548313000552 Published online by Cambridge University Press

Remark. Note that in the second part we can remove the condition $\deg(r) \geqslant 2$, and obtain a bound of $\mathbb{E}[C_E(G)] = |E| + \Theta\big(\sum_u |T_u|\big)$, where the sum is over all vertices $u \in V$ that either have at least two children or have an ancestor with at least two children.

3. Greedy random walk on $\mathbb{Z}^d$

In this section we study the behaviour of GRW on infinite graphs. Specifically we ask whether the walk is recurrent or transient in different graphs.

3.1. GRW on $\mathbb{Z}^d$ for $d \neq 2$

Obviously, GRW on $\mathbb{Z}$ visits every vertex at most once. We show that for $d \geqslant 3$ the greedy random walk on $\mathbb{Z}^d$ is transient.

Theorem 3.1. *Let $G = (V, E)$ be an infinite graph all of whose vertices are of even degree. If the simple random walk on G is transient, then for any rule $\mathcal{R}$ the greedy random walk is also transient.*

In particular, for $d \geqslant 3$, the greedy random walk on $\mathbb{Z}^d$ returns to the origin only finitely many times almost surely.

Proof. Partition the time $[0, +\infty]$ into two types of parts, greedy parts and simple parts, by defining times $t_0 = 0, s_1, t_1, s_2, t_2, \ldots \in \mathbb{N} \cup \{+\infty\}$ as follows:

$$t_0 = 0,$$

$$s_{i+1} = \begin{cases} \inf\{t_i \leqslant t < +\infty : J_t(X_t) = \emptyset\} & \text{if there is such a } t, \\ +\infty & \text{otherwise,} \end{cases}$$

$$t_{i+1} = \begin{cases} \inf\{s_{i+1} \leqslant t < +\infty : J_t(X_t) \neq \emptyset\} & \text{if there is such a } t, \\ +\infty & \text{otherwise.} \end{cases}$$

(An analogous partition underlies the results in Section 2. The difference here is that the times can have the value $+\infty$.)

For the reader's convenience we restate Observation 2.6 adapted for the case of infinite graphs.

Observation 3.2. *If all degrees of a graph $G = (V, E)$ are even and $s_{i+1} < \infty$, then $X_{t_i} = X_{s_{i+1}}$.*

Assume that the event that s_i or t_i equals $+\infty$ for some $i \geqslant 1$, and t_k is the first such time, has a positive probability. Conditioning on this event, the walk remains in a simple part starting from time s_k, and hence performs a simple random walk from this time onwards. Since SRW is transient on G, the walk will return to X_0 only finitely many times almost surely. In fact, since the random range $R = \{X_t : 0 \leqslant s_k\}$ is finite, and the SRW is transient, conditioning on R, the SRW will leave R in finite time almost surely, and so t_k is almost surely finite, contradicting the assumption.

Similarly, if the event that s_i or t_i equals $+\infty$ for some $i \geqslant 1$, and s_k is the first such time, has a positive probability, then, conditioning on this event, the walk is in a greedy part from t_{k-1} onwards. In other words, from time t_{k-1} onwards the walker crosses each edge at most once. Hence, as the degree of X_0 is finite, the maximal number of returns to X_0 is at most

$$\frac{\deg(X_0)}{2} + \frac{t_{k-1}}{2},$$

and in particular almost surely finite.

Assume now that the event that $s_i, t_i < +\infty$ for all $i \geqslant 1$ has a positive probability, and condition on this event. Using the assumption that all vertices of the graph have even degrees, it follows from Observation 3.2 that $X_{t_i} = X_{s_{i+1}}$ for all $i \geqslant 0$. Therefore, for all $i \geqslant 0$ the walk in time segments $[s_i, t_i]$ and $[s_{i+1}, t_{i+1}]$ can be concatenated. Hence, the walk restricted to time $\bigcup_{i \geqslant 0}[s_i, t_i]$ is distributed as a simple random walk on G, and so, by transience, returns to X_0 finitely many times almost surely. Since in the overall greedy parts the walker can visit X_0 at most $\deg(X_0)/2$ times, the entire walk returns to X_0 finitely many times almost surely. □

Note that we have made great use of the fact that every vertex in our graph has even degree. The following proposition shows a similar result by slightly relaxing this assumption.

Proposition 3.3. *Let $G = (V, E)$ be a graph obtained from $\mathbb{Z}^d$ by removing at most $r^{d-2-\epsilon}$ edges from any box of radius r centred at the origin for some $\epsilon > 0$. Then the greedy random walk on G is transient.* □

The proof generalizes the concatenation argument of Theorem 3.1. Unlike the previous proof, which relied on the fact that all vertices had even degrees, in our case some vertices have odd degrees. Hence it is possible that the simple parts cannot be concatenated into one walk. However, we can divide the simple parts into classes, such that in each class the parts can be concatenated into one simple random walk. The proof uses the fact that a simple random walk starting from a point at distance r from the origin then visits the origin with probability $O(1/r^{d-2})$. Therefore, if there are $r^{d-2-\epsilon}$ independent simple random walkers starting on a sphere of radius r centred at the origin, then the total number of visits to the origin by all the walkers is almost surely finite.

Proof. We start with a time partition $(t_0 = 0, s_1, t_1, s_2, t_2, \ldots)$ as in the proof of Theorem 3.1. Call a vertex v a *new start* if $v = X_{s_i}$ for some i and $X_{t_j} \neq v$ for all $j < i$. As in the proof of Theorem 3.1, the concatenation argument implies that every new start vertex must be either the origin or have an odd degree.

Consider the walk restricted to the segments $[s_i, t_i]$. The indices $i \geqslant 1$ can be partitioned into classes $C_1, C_2, \ldots$ such that in each class C_j the segments $[s_i, t_i]$, $i \in C_j$, can be concatenated into one walk that starts with a new start vertex. Namely, if $C_j = \{i_1 < i_2 < i_3 < \cdots\}$, then $X_{s_{i_1}}$ is a new start and $X_{s_{i_2}} = X_{t_{i_1}}, X_{s_{i_3}} = X_{t_{i_2}}, \ldots$ Letting m_j denote $\min C_j$,

https://doi.org/10.1017/S0963548313000552 Published online by Cambridge University Press

we have that $X_{s_{m_j}}$ is necessarily a new start, and is therefore either the origin or a vertex of odd degree. Moreover, the times $\{s_{m_j}\}_j$ are all distinct.

For each C_j, restricting the walk to times $\bigcup\{[s_i, t_i] : i \in C_j\}$ gives us a simple random walk (possibly finite) starting from $X_{s_{m_j}}$. Therefore there are at most $O(r^{d-2-\epsilon})$ simple random walks, starting from a box of radius r centred at the origin. Using the fact that a random walk in $\mathbb{Z}^d$ starting from a vertex at distance r from the origin hits the origin with probability $O(1/r^{d-2})$, we conclude that the sum of probabilities of hitting zero converges, when summing over all random walks. More precisely, let P_v be the probability that a simple random walk starting at v reaches the origin, and let ODD be the set of all vertices of odd degree. Then

$$\sum_{v\in \mathrm{ODD}} P_v = \sum_{n=1}^{\infty} \sum_{\substack{v\in \mathrm{ODD} \\ 2^{n-1}\leqslant \|v\| < 2^n}} P_v \leqslant \sum_n (2^n)^{(d-2-\epsilon)} \cdot O\left(\frac{1}{(2^n)^{d-2}}\right) = O\left(\sum_{n=1}^{\infty} 2^{-\epsilon n}\right) < \infty.$$

By the first Borel–Cantelli lemma, the event that only finitely many of the walks will reach the origin occurs with probability 1. Therefore, GRW on this graph is transient, as required. □

3.2. GRW on $\mathbb{Z}^2$ and the mirror model

The following observation relating the behaviour of GRW on $\mathbb{Z}^2$ to the mirror model is due to Omer Angel.

In the mirror model, introduced by Ruijgrok and Cohen [18], a mirror is placed randomly on $\mathbb{Z}^2$ by aligning a mirror along either one of the diagonal directions with probability 1/3 each, or placing no mirror with probability 1/3. A particle moves along the edges of the lattice and is reflected by the mirrors according to the law of reflection: see, *e.g.*, [12] for details. A major open problem in this area is to determine whether every orbit is periodic almost surely. We claim below that this question is equivalent to determining whether GRW with rule $\mathcal{R}_{\mathsf{RAND}}$ is recurrent in $\mathbb{Z}^2$. (Recall that in the rule $\mathcal{R}_{\mathsf{RAND}}$ an edge is chosen uniformly at random among the adjacent unvisited edges of the current vertex.)

Let $(X_t)_{t\geqslant 0}$ be a GRW on $\mathbb{Z}^2$ with the rule $\mathcal{R}_{\mathsf{RAND}}$. Then there exists a coupling between $(X_t)_{t\geqslant 0}$ and the particle motion in the planar mirror model until the first time they return to the origin. Indeed, if at time $t \geqslant 0$ the GRW reaches a vertex X_t that we have not visited so far, then in both the GRW and in the mirror model the next step will be chosen in a non-backtracking manner, giving equal probabilities of 1/3 to each of the adjacent vertices (except for X_{t-1}). In the mirror model this uniquely defines the alignment of the mirror at vertex X_t and hence the next move of the particle in the next visit to this place, given that the orbit is not periodic: it will go to the unvisited neighbouring vertex. On the other hand, if at time $t \geqslant 0$ we reach a vertex X_t that has already been visited previously, then the next step is uniquely determined: it is to make a move along the edge that has not been traversed so far. This defines a coupling of the two models up to the first returning time to zero.

Claim 3.4. *The probability that GRW with the rule $\mathcal{R}_{\mathsf{RAND}}$ on $\mathbb{Z}^2$ returns to the origin at least once is equal to the probability that a particle returns to the origin in the planar mirror model.*

From Claim 3.4 we infer the following theorem.

Theorem 3.5. *A GRW with rule $\mathcal{R}_{\mathsf{RAND}}$ on $\mathbb{Z}^2$ returns to the origin infinitely often almost surely if and only if the orbit in the mirror model on $\mathbb{Z}^2$ is periodic almost surely.*

Proof. Note first that a GRW on $\mathbb{Z}^2$ returns to the origin infinitely often if and only if every greedy part is finite. Indeed, if there is an infinite greedy part, then there are finitely many returns to the origin as every vertex is visited at most twice in total in all greedy time parts. In the other direction, assume that all greedy time parts are finite. Then, by the concatenation argument, which follows by Observation 3.2, the simple parts form an infinite subsequence distributed as SRW on $\mathbb{Z}^2$ starting at the origin. The latter returns to the origin infinitely often almost surely, and hence, so does GRW. Therefore, it is enough to show that the orbit in the mirror model on $\mathbb{Z}^2$ is periodic almost surely if and only if every greedy part is finite.

Suppose first that every orbit in the mirror model on $\mathbb{Z}^2$ is periodic almost surely, and suppose that GRW starts the ith greedy part at some time t_i. Then, conditioning on the t_i steps of GRW so far, we define the orientation of the mirrors in the vertices visited up to now. Since the number of visited vertices is finite, it follows that the conditioning is on a non-zero event, and so the trajectory of the particle starting from X_{t_i} is almost surely periodic. Therefore, by considering the coupling between GRW and the mirror model conditioned on that event, analogously to Claim 3.4, it follows that, with probability 1, the ith greedy part is finite.

Assume now that every greedy part of GRW is finite. Note that by translation invariance it is enough to show that the trajectory of a single particle starting at the origin is periodic almost surely.[2] Indeed, since GRW returns to the origin twice almost surely, it follows from the coupling in Claim 3.4 that the trajectory of a particle starting at the origin is periodic almost surely, as required. □

4. Remarks and open problems

4.1. A conjecture regarding Theorem 2.2

Recall Observation 2.6 used in the proof of Theorem 2.5. It seems to be potentially useful for proving stronger bounds on the edge cover time of GRW. To illustrate how this observation can be useful, let us consider the GRW on the complete graph K_n for odd

[2] Indeed, if the trajectory of a particle starting at the origin is periodic almost surely, then, by translation invariance the trajectory of a particle starting at any vertex and moving in any direction is periodic almost surely. Thus, by placing four particles at each vertex of the graph and letting them move in the four possible directions, it follows that with probability 1, the trajectory of each of them is periodic, as this event is an intersection of countably many probability 1 events. Therefore, the trajectory of a particle is periodic almost surely if and only if all orbits are periodic almost surely.

https://doi.org/10.1017/S0963548313000552 Published online by Cambridge University Press

values of n. In the proof of Theorem 2.2 we only used the assumption that the 'bad' sets B_i grow at least by one each time, thus allowing us to bound the 'overhead' by

$$\mathbb{E}\left[\sum_{i=1}^{k-1} \frac{n-1}{n-|B_i|}\right] \leqslant \mathbb{E}\left[\sum_{i=1}^{k-1} \frac{n-1}{n-i}\right] \leqslant n\log(n).$$

We suspect, however, that the sets B_i grow linearly in n, since by the time the walker gets stuck for the first time, *i.e.*, visits the starting vertex $n/2$ times, the number of vertices that have already been visited $n/2$ times will be linear in n. The situation, however, becomes more complicated when trying to analyse the set B_2, as it seems to require some understanding regarding the subgraph of K_n that has not been covered by the time s_1 when the walker got stuck for the first time. If this is indeed true, and the sets B_i grow linearly at each step, we would obtain a stronger bound $\mathbb{E}[\sum(t_i - s_i)] = O(n)$. We make the following, rather bold, conjecture.

Conjecture 4.1. *The expected edge cover time of GRW on K_n is*

$$\mathbb{E}[C_E(K_n)] = |E| + \Theta(n).$$

An interesting result in this direction is a recent result of Omer Angel and Yariv Yaari. They showed that for the complete graph K_n for odd values of n, *i.e.*, when the graph K_n is of even degree, the expected number of unvisited edges in K_n until the first time the walk got stuck (*i.e.*, up to time s_1) is linear in n [19].

4.2. Rules on vertices instead of edges

In this paper we have considered the *edge* cover time of graphs, rather than the *vertex* cover time. This seems to be a natural quantity to analyse due to the transition rule of GRW. A naïve modification of GRW to speed up the *vertex* cover time is as follows. At each step, the walker at vertex v picks an unvisited neighbour of v according to some rule and jumps there. If all neighbours have already been visited, the next move is chosen uniformly at random among the neighbours of v. For example, it is obvious that in the complete graph K_n, this walk covers all vertices in n steps.

Note that, when the walker is allowed to make some local computations at a vertex, and each vertex has information regarding its neighbours, then one can define a rule that will force the walk to perform depth-first search on the graph, by letting each vertex use only the information regarding its neighbours. Such a walk crosses each edge of some spanning tree at most twice, thus visiting all vertices of the graph in less than $2n$ steps.

4.3. Open problems

In order to avoid trivialities, in the questions below consider GRW with the rule $\mathcal{R}_{\mathsf{RAND}}$.

(1) Give a tight bound for the 'overhead' of GRW on the complete graph. Specifically, is it true that $\mathbb{E}[C_E(K_n)] = \binom{n}{2} + \Theta(n)$?

(2) Show upper bounds on $C_E(G)$ for other families of graphs. One interesting example to look at would be the d-dimensional torus.

(3) It would also be interesting to analyse the GRW on graphs with power-law degree distribution. On such graphs there are hubs of very large degrees, and when visiting them, the GRW is expected to be efficient.
(4) Show that for any transitive graph the expected edge cover time of the GRW cannot be asymptotically larger than that of the SRW for any finite graph. We know that this is true for vertex-transitive graphs of even degree.
(5) Give bounds on the expected *vertex* cover time of the GRW for finite graphs.
(6) Give bounds on the expected hitting time of GRW for different graphs.
(7) Define the GRW mixing time and show that the GRW mixing time is as fast as that of SRW. Here [4] is relevant, and [16] may also be found to be useful.

The remaining problems are regarding recurrence/transience of GRW on infinite graphs.

(8) Is GRW on $\mathbb{Z}^2$ recurrent? Is GRW diffusive on $\mathbb{Z}^d$, for all $d \geqslant 2$? (See the discussion in Section 3.2.)
(9) Is GRW on the ladder $\mathbb{Z} \times \mathbb{Z}_2$ recurrent?
(10) Prove that GRW is transient on any graph that is roughly isometric to $\mathbb{Z}^3$. In particular, show it for odd degree lattices.
(11) Show that GRW is transient on non-amenable infinite graphs.
(12) Consider GRW on a vertex-transitive graph. Is there a zero-one law for the event that the walker returns to the initial location infinitely often? Note that, as in the argument in the proof of Theorem 3.5, the event above occurs almost surely if and only if the walker returns to the initial location almost surely.

Acknowledgements

We are grateful to Itai Benjamini for proposing the model and for many valuable discussions, and to Gady Kozma for his useful ideas. We thank Shlomo Jozeph for suggesting that we use the concatenation argument in Section 3. We also thank Eviatar Procaccia for pointing out Omer Angel's argument relating GRW to the mirror model in $\mathbb{Z}^2$. T.O.'s research is supported by the Israel Science Foundation. I.S.'s research is supported by ERC grant number 239985.

References

[1] Aldous, D. J. (1989) Lower bounds for covering times for reversible Markov chains and random walks on graphs. *J. Theoret. Probab.* **2** 91–100.
[2] Aldous, D. J. (1991) Random walk covering of some special trees. *J. Math. Anal. Appl.* **157** 271–283.
[3] Aleliunas, R., Karp, R. M., Lipton, R. J., Lovasz, L. and Charles, R. (1979) Random walks, universal traversal sequences, and the complexity of maze problems. In *Proc. 20th Annual Symposium on Foundations of Computer Science*, pp. 218–223.
[4] Alon, N., Benjamini, I., Lubetzky, E. and Sodin, S. (2007) Non-backtracking random walks mix faster. *Commun. Contemp. Math.* **9** 585–603.
[5] Avin, C. and Krishnamachari, B. (2008) The power of choice in random walks: An empirical study. *Comput. Networks* **52** 44–60.
[6] Benjamini, I., Gurel-Gurevich, O. and Morris, B. (2013) Linear cover time is exponentially unlikely. *Probab. Theory Rel. Fields* **155** 451–461.

[7] Benjamini, I. and Wilson, D. B. (2003) Excited random walk. *Elect. Comm. Probab.* **8** 86–92.

[8] Berenbrink, P., Cooper, C., Elsässer, R., Radzik, T. and Sauerwald, T. (2010) Speeding up random walks with neighborhood exploration. In *Proc. 21st Annual ACM–SIAM Symposium on Discrete Algorithms*, pp. 1422–1435.

[9] Berenbrink, P., Cooper, C. and Friedetzky, T. (2012) Random walks which prefer unvisited edges, and exploring high girth even degree expanders in linear time. In *Proc. 31st Annual ACM SIGACT–SIGOPS Symposium on Principles of Distributed Computing*.

[10] Bollobás, B. (2001) *Random Graphs*, Cambridge University Press.

[11] Broder, A. Z. and Karlin, A. R. (1989) Bounds on the cover time. *J. Theoret. Probab.* **2** 101–120.

[12] Bunimovich, L. A. and Troubetzkoy, S. E. (1992) Recurrence properties of Lorentz lattice gas cellular automata. *J. Statist. Phys.* **67** 289–302.

[13] Feige, U. (1995) A tight lower bound on the cover time for random walks on graphs. *Random Struct. Alg.* **6** 433–438.

[14] Friedrich, T. and Sauerwald, T. (2010) The cover time of deterministic random walks. In *Proc. 16th Annual International Conference on Computing and Combinatorics: COCOON'10*, Springer, pp. 130–139.

[15] Hoory, S., Linial, N. and Wigderson, A. (2006) Expander graphs and their applications. *Bull. Amer. Math. Soc.* **43** 439–561.

[16] Madras, N. and Wu, C. C. (2005) Self-avoiding walks on hyperbolic graphs. *Combin. Probab. Comput.* **14** 523–548.

[17] Pemantle, R. (2007) A survey of random processes with reinforcement. *Probability Surveys* **4** 1–79.

[18] Ruijgrok, T. and Cohen, E. (1988) Deterministic lattice gas models. *Phys. Lett.* A **133** 415–418.

[19] Yaari, Y. (2011) MSc thesis, Weizmann Institute of Science, Israel.

[20] Zuckerman, D. (1990) A technique for lower bounding the cover time. In *Proc. 22nd annual ACM Symposium on Theory of Computing: STOC '90*, ACM, pp. 254–259.