A Random Environment for Linearly Edge-Reinforced Random Walks on Infinite Graphs

Franz Merkl[1] Silke W.W. Rolles [2]

August 19, 2005

Abstract

We consider linearly edge-reinforced random walk on an arbitrary locally finite connected graph. It is shown that the process has the same distribution as a random walk in a time-independent random environment given by strictly positive weights on the edges. Furthermore, we prove bounds for the random environment, uniform, among others, in the size of the graph. [34]

1 Introduction

Linearly edge-reinforced random walk (errw) is the following model: Consider a locally finite connected graph $G = (V, E)$ with vertex set V and edge set E. The edges are *undirected*. A random walker moves randomly on the vertices of the graph, traversing an edge between each discrete time $t = 0, 1, 2, \ldots$. Let X_t denote the random location of the random walker at time t. At time 0, the random walker starts in a distinguished vertex $X_0 = \mathbf{0} \in V$. We realize the X_t as canonical projections defined on the set $\Omega_{\mathbf{0}} \subseteq V^{\mathbb{N}_0}$ of all admissible paths.

The random walk $(X_t)_{t\in\mathbb{N}_0}$ is non-Markovian. The relevant memory of the random walker is encoded in random weights $w_e(t)$, $e \in E$, $t \in \mathbb{N}_0$, which change with time. Initially, the weights take prescribed values $w_e(0) := a_e > 0$, $e \in E$, possibly depending on the edge e. Each time the random walker traverses the edge $e = \{u, v\} \in E$, the weight of e is increased by 1, and the weight of all the other edges $e' \in E \setminus \{e\}$ remain unchanged. In other words:

$$w_e(t) := a_e + \sum_{s=1}^{t} 1_e(\{X_{s-1}, X_s\}), \qquad (e \in E,\ t \in \mathbb{N}_0). \tag{1.1}$$

The random weights $w_e(t)$, representing the memory of the random walker at time t, determine the transition probabilities of the random walker as follows: For any $u \in V$,

[1]Mathematical Institute, University of Munich, Theresienstr. 39, D-80333 Munich, Germany. e-mail: merkl@mathematik.uni-muenchen.de

[2]Eindhoven University of Technology, Department of Mathematics, P.O. Box 513, 5600 MB Eindhoven, The Netherlands. e-mail: srolles@win.tue.nl

[3]MSC 2000 subject classifications: Primary 60K35; secondary 60K37.

[4]Key words: Reinforced random walk, random environment.

one has

$$P^G_{\mathbf{0},a}[X_{t+1}=u \mid X_0,\ldots,X_t] = \frac{w_{\{X_t,u\}}(t)}{w_{X_t}(t)} \quad \text{whenever } \{X_t,u\}\in E, \tag{1.2}$$

where for $v \in V$, we set

$$w_v(t) := \sum_{e \ni v} w_e(t) \quad \text{and} \quad a = (a_e)_{e\in E}. \tag{1.3}$$

Whenever $\{X_t,u\} \notin E$, then

$$P^G_{\mathbf{0},a}[X_{t+1}=u \mid X_0,\ldots,X_t] = 0. \tag{1.4}$$

In other words, the transition probabilities are proportional to the weight of the traversed edge at that time. For an overview of the history of the model see e.g. [MR05c].

Linearly edge-reinforced random walk is partially exchangeable in the following sense: Two finite paths with the same starting point are traversed with the same probability by the reinforced random walker provided every (undirected) edge is traversed the same number of times in both paths. If the reinforced random walk is recurrent, that is, if it returns to its starting point infinitely often with probability one, then a de Finetti theorem for Markov chains due to Diaconis and Freedman [DF80] implies that the reinforced random walk is a mixture of Markov chains. In particular, this applies to finite graphs. Even more, on any finite graph, the reinforced random walk is a mixture of *reversible* Markov chains; this follows e.g. by a de Finetti theorem for reversible Markov chains (see [Rol03]). However, on many infinite, locally finite graphs with some initial weights, edge-reinforced random walk is *not* recurrent, and on some infinite, locally finite graphs, including $\mathbb{Z}^d$ for $d \geq 2$, it is not known whether edge-reinforced random walk is recurrent. In this paper, we prove that for *any locally finite* graph and any initial weights, edge-reinforced random walk is a mixture of Markov chains, irrespectively whether it is recurrent or not. The Markov chains are defined in terms of random, time-independent weights on the (undirected) edges. For finite graphs, there is an explicit but complicated formula for the joint distribution of these random weights. They show a complicated dependence structure. This description was first given by Coppersmith and Diaconis in [CD86] and refined by Keane and Rolles in [KR00]. It has already been useful to analyze edge-reinforced random walk on certain infinite graphs, including ladders of arbitrary width, by taking the infinite volume limit of finite subgraphs; see [MR05b], [Rol05], and [MR05a]. However, in this paper, we do *not* make use of the explicit form of the mixing measure for finite graphs.

2 Results

For ordinary (Markovian) random walks, the following is known: If the random walker returns to its starting point a.s. at least once, then it also returns a.s. infinitely often. For general non-Markovian random walks, this breaks down. However, for linearly edge-reinforced random walks, the implication is still true, as the following theorem shows:

Theorem 2.1 *For the edge-reinforced random walk on any locally finite connected graph, the following statements are equivalent:*

(a) The edge-reinforced random walker returns to its starting point with probability one.

(b) The edge-reinforced random walker returns to its starting point infinitely often with probability one.

(c) With probability one, every vertex is visited at least twice by the edge-reinforced random walker.

(d) The edge-reinforced random walker visits all vertices infinitely often with probability one.

For $x = (x_e)_{e\in E} \in \mathbb{R}_+^E$, let $Q_{\mathbf{0},x}^G$ denote the distribution of the Markov chain on G over $\Omega_{\mathbf{0}}$ with starting vertex $\mathbf{0}$ and *time-independent* transition probabilities

$$Q_{\mathbf{0},x}^G[X_{t+1} = u'|X_t = u] = \frac{x_{\{u,u'\}}}{\sum_{e\in E: u\in e} x_e} 1_{\{\{u,u'\}\in E\}}. \tag{2.1}$$

The following theorem generalizes the known representation of linearly edge-reinforced random walk on *finite* graphs to general locally finite, possibly infinite graphs.

Theorem 2.2 *For edge-reinforced random walk on any locally finite graph G with any starting vertex $\mathbf{0}$ and any initial edge weights $a = (a_e)_{e\in E} \in \mathbb{R}_+^E$, there exists a probability measure $\mathbb{Q}_{\mathbf{0},a}^G$ on the set $(0,\infty)^E$ of strictly positive edge weights such that for all events $A \subseteq \Omega_{\mathbf{0}}$, one has*

$$P_{\mathbf{0},a}^G[A] = \int_{\mathbb{R}_+^E} Q_{\mathbf{0},x}^G[A]\, \mathbb{Q}_{\mathbf{0},a}^G(dx), \tag{2.2}$$

i.e. the edge-reinforced random walk has the same distribution as a random walk in a random environment given by (time-independent) random weights on the edges.

We prove bounds for the distribution of the transition probabilities of edge-reinforced random walk. These estimates hold for any locally finite, possibly infinite graph and any initial weights. They depend only on the local structure of the graph and on the initial weights nearby.

Theorem 2.3 *For edge-reinforced random walk on any locally finite graph G with any starting vertex $\mathbf{0}$ and any initial edge weights a, there is a measure $\mathbb{Q}_{\mathbf{0},a}^G$ with the properties specified in Theorem 2.2, such that the following holds: For any vertex v and any edge e incident to v, there are constants $c_1 = c_1(a_v, a_e) > 0$ and $c_2 = c_2(a_v, a_e) > 0$, depending continuously only on a_v and a_e, $a_v \geq a_e > 0$, but not depending on any other details of G, $\mathbf{0}$, and a, such that for all $\varepsilon > 0$, the estimates*

$$\mathbb{Q}_{\mathbf{0},a}^G\left[\frac{x_e}{x_v} \leq \varepsilon\right] \leq c_1 \varepsilon^{a_e/2} \quad \text{and} \quad \mathbb{Q}_{\mathbf{0},a}^G\left[\frac{x_e}{x_v} \geq 1-\varepsilon\right] \leq c_2 \varepsilon^{(a_v-a_e)/2} \tag{2.3}$$

hold.

The random variables x_e/x_f; $e, f \in E$ are tight with algebraically bounded tails, uniformly in the choice of the graph G, provided that the graph distance from e to f is bounded, and provided that the initial weights are bounded and bounded away from 0.

More precisely:

Theorem 2.4 *For all compact sets $K \subset (0,\infty)$, there exists a constant $c_3(K) > 0$, such that for all locally finite graphs $G = (V, E)$, any starting vertex $\mathbf{0}$ and all initial edge weights $a = (a_e)_{e\in E}$, there is a measure $\mathbb{Q}^G_{\mathbf{0},a}$ with the properties specified in Theorems 2.2 and 2.3, such that the following holds: For all edges $e, f \in E$, for all paths*

$$\stackrel{e=e_0}{\longrightarrow} v_1 \stackrel{e_1}{\longrightarrow} v_2 \stackrel{e_2}{\longrightarrow} \dots \stackrel{e_{l-1}}{\longrightarrow} v_l \stackrel{e_l=f}{\longrightarrow} \tag{2.4}$$

of length $l+1$ from e to f with the property $a_{e_i} \in K$ and $a_{v_j} = \sum_{e'\ni v_j} a_{e'} \in K$ ($i = 0, \dots, l$; $j = 1, \dots, l$) and for all $M > 0$, one has

$$\mathbb{Q}^G_{\mathbf{0},a}[x_e \geq M x_f] \leq c_3(K) l M^{-\gamma}, \tag{2.5}$$

where

$$\gamma = \frac{1}{2l} \min\{a_{e_i} \ : \ 1 \leq i \leq l\}. \tag{2.6}$$

In general, it is not known whether the measure $\mathbb{Q}^G_{\mathbf{0},a}$ is unique. If the underlying graph is of the form $\mathbb{Z} \times T$ with a finite tree T and the initial weights are large, then, up to multiplication of all x_e by the same constant, there is a unique measure $\mathbb{Q}^G_{\mathbf{0},a}$ satisfying (2.2) for all events A; see Theorem 2.4 in [MR05a]. Furthermore, in that case, $\sum_{e\in E} x_e < \infty$ holds $\mathbb{Q}^G_{\mathbf{0},a}$-almost surely.

The bounds in Theorems 2.3 and 2.4 are first proven on finite graphs. In the sense of convex order, we bound the jump probability distribution for edge-reinforced random walk by transition probabilities of suitable Polya urn models. The notion of convex order plays an important role in our argument. Therefore, we review the basic properties of convex order in Section 3. The comparison with Polya urn models is a local construction. It therefore yields uniform bounds in the size of the graph. We approximate an infinite locally finite graph by growing finite pieces, using compactness and tightness arguments. The uniformity of the estimates is essential in this approximation to get tightness.

3 Preliminaries

In this section, we collect just the basic properties of convex order that we need.

Definition 3.1 *Let X and Y be random variables with finite expectation, not necessarily defined on the same probability space. We say that X* and *Y* are in convex order*, in symbols $X \lhd Y$, if there are random variables $X_1 \stackrel{d}{=} X$ and $Y_1 \stackrel{d}{=} Y$ defined on a common probability space $(\Omega, \mathcal{A}, P)$ such that (X_1, Y_1) is a 1-step martingale, i.e., if there is a σ-field $\mathcal{F} \subseteq \mathcal{A}$ such that X_1 is $\mathcal{F}$-measurable and*

$$X_1 = E_P[Y_1|\mathcal{F}] \tag{3.1}$$

holds a.s. The relation $X \lhd Y$ depends only on the laws $\mathcal{L}_X$ and $\mathcal{L}_Y$ of X and Y; we therefore write also $\mathcal{L}_X \lhd \mathcal{L}_Y$.

Note that we may always replace $\mathcal{F}$ by $\sigma(X_1)$ in (3.1).

Lemma 3.2 *(a) The relation $\lhd$ is transitive.*

(b) Assume that $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ are uniformly integrable martingales with respect to filtrations $(\mathcal{F}_n)_{n\in\mathbb{N}}$ and $(\mathcal{G}_n)_{n\in\mathbb{N}}$, respectively, not necessarily defined on the same probability space. Let X and Y denote the a.s. limit of $(X_n)_n$ and $(Y_n)_n$, respectively. Assume that for all $n \in \mathbb{N}$, the relation $X_n \lhd Y_n$ holds. Then $X \lhd Y$ also holds.

(c) Let $\mu_1, \dots, \mu_n$ and $\nu_1, \dots, \nu_n$ be distributions on $\mathbb{R}$ such that $\mu_i \lhd \nu_i$ holds for all $i = 1, \dots, n$. Let p_i, $i = 1, \dots, n$, be nonnegative numbers with sum 1. Then

$$\sum_{i=1}^{n} p_i \mu_i \lhd \sum_{i=1}^{n} p_i \nu_i \tag{3.2}$$

also holds.

(d) Let $X \lhd Y$. For any convex function $f : I \to \mathbb{R}$, defined at least on an interval $I \subseteq \mathbb{R}$ containing the range of X and Y, such that $E[f(Y)] < \infty$ is valid, the inequality

$$E[f(X)] \leq E[f(Y)] \tag{3.3}$$

holds.

Remark. As a consequence of Strassen's theorem (see e.g. Theorem 2 in Strassen's classical paper [Str65]), the converse of part (d) is also true. This theorem could be used to prove the lemma. However, in this paper, we rather base the proof of the lemma on more direct, elementary arguments, using only Jensen's inequality, i.e. the trivial direction of Strassen's theorem.

Proof.

(a) Assume that $X \lhd Y$ and $Y \lhd Z$. Let $X_1 \stackrel{d}{=} X$ and $Y_1 \stackrel{d}{=} Y$ be random variables on a common probability space $(\Omega_1, \mathcal{A}_1, P_1)$, such that $X_1 = E_{P_1}[Y_1|X_1]$ holds. Let $P_1[Y_1 \in \cdot|X_1 = \cdot]$ denote a regular conditional distribution of Y_1 conditional on X_1, i.e. it is a stochastic kernel defined on $\mathbb{R} \times \mathcal{B}(\mathbb{R})$ such that for all $A, B \in \mathcal{B}(\mathbb{R})$, one has

$$P_1[X_1 \in A, Y_1 \in B] = \int_A P_1[Y_1 \in B \mid X_1 = x]\mathcal{L}_X(dx). \tag{3.4}$$

Similarly, let $Y_2 \stackrel{d}{=} Y$ and $Z_2 \stackrel{d}{=} Z$ be random variables on a common probability space $(\Omega_2, \mathcal{A}_2, P_2)$, such that $Y_2 = E_{P_2}[Z_2|Y_2]$ holds. Let $P_2[Z_2 \in \cdot|Y_2 = \cdot]$ denote a regular conditional distribution of Z_2 conditional on Y_2. Let P_3 be the law of a time-inhomogeneous Markov chain with 3 time points, starting distribution $\mathcal{L}_X$, first transition kernel $P_1[Y_1 \in \cdot | X_1 = \cdot]$, and second transition kernel $P_2[Z_2 \in \cdot | Y_2 = \cdot]$, i.e. P_3 is the probability measure on $\mathcal{B}(\mathbb{R}^3)$ which fulfills

$$P_3[A \times B \times C] = \int_A \int_B P_2[Z_2 \in C | Y_2 = y] \, P[Y_1 \in dy | X_1 = x] \, \mathcal{L}_X(dx) \qquad (3.5)$$

for all cylinder events $A \times B \times C \subseteq \mathbb{R}^3$. Let X_3, Y_3, and Z_3 denote the projections to the first, second, and third coordinate of $\mathbb{R}^3$, respectively. By construction, with respect to P_3, $X_3 \stackrel{d}{=} X$, $Y_3 \stackrel{d}{=} Y$, $Z_3 \stackrel{d}{=} Z$, and we have P_3-a.s.:

$$E_{P_3}[Z_3 | X_3, Y_3] = E_{P_3}[Z_3 | Y_3] = Y_3 \qquad (3.6)$$

and therefore

$$E_{P_3}[Z_3 | X_3] = E_{P_3}[E_{P_3}[Z_3 | X_3, Y_3] | X_3] = E_{P_3}[Y_3 | X_3] = X_3. \qquad (3.7)$$

This shows $X \triangleleft Z$.

(b) Fix $n \in \mathbb{N}$. Since the martingale $(Y_m)_{m \in \mathbb{N}}$ is uniformly integrable, it also converges in L^1 to Y, and we have $E[Y | \mathcal{G}_n] = Y_n$ a.s.; in particular $Y_n \triangleleft Y$ holds. Since $X_n \triangleleft Y_n$ holds by assumption, we conclude $X_n \triangleleft Y$. Therefore, we can take $X'_n \stackrel{d}{=} X_n$ and $Y'_n \stackrel{d}{=} Y$ on some common probability space $(\Omega_n, \mathcal{A}_n, P_n)$ such that $X'_n = E_{P_n}[Y'_n | X'_n]$. Hence, for any bounded continuous function $f : \mathbb{R} \to \mathbb{R}$, the identity

$$E_{P_n}[f(X'_n) X'_n] = E_{P_n}[f(X'_n) Y'_n] \qquad (3.8)$$

holds. Since $(X'_n)_{n \in \mathbb{N}}$ is uniformly integrable, $(X'_n, Y'_n)_{n \in \mathbb{N}}$ is tight. Thus, there exists a subsequence $(X'_{n_k}, Y'_{n_k})_{k \in \mathbb{N}}$ that converges weakly to a limit (X', Y'). Taking the limit as $k \to \infty$ and using uniform integrability yields

$$E[f(X') X'] = E[f(X') Y'], \qquad (3.9)$$

which is equivalent to $X' = E[Y' | X']$. Since by construction $\mathcal{L}_{X'} = \mathcal{L}_X$ (recall that $X_n \to X$ almost surely and hence in distribution) and $\mathcal{L}_{Y'} = \mathcal{L}_Y$, we have shown that $X \triangleleft Y$.

(c) For $1 \leq i \leq n$, let X_i and Y_i be random variables on some probability space $(\Omega_i, \mathcal{A}_i, P_i)$ with $\mathcal{L}_{X_i} = \mu_i$ and $\mathcal{L}_{Y_i} = \nu_i$ such that $X_i = E[Y_i | \mathcal{F}_i]$ for some σ-algebra $\mathcal{F}_i \subseteq \mathcal{A}_i$. Without loss of generality, we assume that the sets Ω_i are pairwise disjoint. We set $\Omega = \cup_{i=1}^n \Omega_i$, $\mathcal{A} = \sigma(\cup_{i=1}^n \mathcal{A}_i)$, $P[A] = \sum_{i=1}^n p_i P_i[A \cap \Omega_i]$ for $A \in \mathcal{A}$ and $\mathcal{F} = \sigma(\cup_{i=1}^n \mathcal{F}_i)$. Furthermore, we define random variables $X = \sum_{i=1}^n X_i 1_{\Omega_i}$ and $Y = \sum_{i=1}^n Y_i 1_{\Omega_i}$ on $(\Omega, \mathcal{A}, P)$. Then, $\mathcal{L}_X = \sum_{i=1}^n p_i \mu_i$, $\mathcal{L}_Y = \sum_{i=1}^n p_i \nu_i$, and $E[Y | \mathcal{F}] = X$. Hence, (3.2) holds.

(d) The claim follows immediately from the conditional version of Jensen's inequality.

■

4 Comparison of ERRW on finite graphs with Polya urns

Throughout this section, we consider linearly edge-reinforced random walk on a *finite* graph $G = (V, E)$ with starting point $\mathbf{0} \in V$ and initial weights $a = (a_e)_{e \in E}$. Let X_t, $t \in \mathbb{N}_0$, denote the location of the random walker at time t. Recall that we realize X_t as the t-th projection defined on the subset $\Omega_{\mathbf{0}} \subseteq V^{\mathbb{N}_0}$ of all admissible paths within the set $V^{\mathbb{N}_0}$ of all paths. Recall the definitions (1.1) and (1.3) of $w_e(t)$ and $w_v(t)$. For $v \in V$ and $t \in \mathbb{N}_0$, we abbreviate

$$a_v = \sum_{\substack{e \ni v \\ e \in E}} a_e = w_v(0). \tag{4.1}$$

Fix a vertex $v^* \in V$ and an edge $e^* \in E$ incident to v^*. For $n \in \mathbb{N}_0$, we define T_n, $n \in \mathbb{N}_0$, to denote the time of the $n+1$-st visit of the vertex v^*. Note that all these times are almost surely well-defined. Define the filtration $\mathcal{F} = (\mathcal{F}_n)_{n \in \mathbb{N}_0}$ with

$$\mathcal{F}_n := \sigma(X_t : \ 0 \le t \le T_n), \tag{4.2}$$

encoding the observable information up to time T_n. We set

$$M_n^{\mathrm{errw}} := \frac{w_{e^*}(T_n)}{w_{v^*}(T_n)} = \frac{w_{e^*}(T_n)}{w_{v^*}(T_0) + 2n}. \tag{4.3}$$

Here, we use that $w_{v^*}(T_n) = w_{v^*}(T_0) + 2n$, since between two subsequent visits of v^*, two (maybe coinciding) edges adjacent to v^* have been crossed, one to leave v^*, and one to enter v^* again. The random variable M_n^{errw} equals the conditional probability given $\mathcal{F}_n$ that the edge-reinforced random walker leaves the vertex v^* at time T_n via the edge e^*.

Lemma 4.1 *The sequence $(M_n^{\mathrm{errw}})_{n \in \mathbb{N}_0}$ is a martingale with respect to $\mathcal{F}$.*

Proof. By its definition (4.3), M_n^{errw} is $\mathcal{F}_n$-measurable. Define a partition of $\Omega_{\mathbf{0}}$ by

$$A = \{\{X_{T_n}, X_{T_n+1}\} = e^*, \ \{X_{T_{n+1}-1}, X_{T_{n+1}}\} \neq e^*\}, \tag{4.4}$$
$$B = \{\{X_{T_n}, X_{T_n+1}\} \neq e^*, \ \{X_{T_{n+1}-1}, X_{T_{n+1}}\} = e^*\}, \tag{4.5}$$
$$C = \{\{X_{T_n}, X_{T_n+1}\} = e^*, \ \{X_{T_{n+1}-1}, X_{T_{n+1}}\} = e^*\}, \tag{4.6}$$
$$D = \{\{X_{T_n}, X_{T_n+1}\} \neq e^*, \ \{X_{T_{n+1}-1}, X_{T_{n+1}}\} \neq e^*\}. \tag{4.7}$$

The events A, B, C, D distinguish whether the reinforced random walker leaves the vertex v^* at time T_n via e^* or not and whether it enters v^* at time T_{n+1} via e^* or not. We observe that

$$w_{e^*}(T_{n+1}) = \begin{cases} w_{e^*}(T_n) + 1 & \text{on } A \cup B, \\ w_{e^*}(T_n) + 2 & \text{on } C, \\ w_{e^*}(T_n) & \text{on } D. \end{cases} \tag{4.8}$$

Let us consider the transformation $\Pi_{n+1} : \Omega_{\mathbf{0}} \to \Omega_{\mathbf{0}}$ which reverts the orientation of the $n+1$-st excursion from v^*, i.e. the excursion between time T_n and time T_{n+1}, while keeping the rest of the path unchanged. By partial exchangeability (see e.g. Lemma 2 of [KR00]), this transformation is measure preserving. Furthermore, it maps A bijectively to B. Hence, abbreviating $P := P^G_{\mathbf{0},a}$, the formula $P[A|\mathcal{F}_n] = P[B|\mathcal{F}_n]$ is valid. Note that

$$P[A|\mathcal{F}_n] + P[C|\mathcal{F}_n] = M_n^{\text{errw}} \quad \text{and} \tag{4.9}$$

$$P[B|\mathcal{F}_n] + P[D|\mathcal{F}_n] = 1 - M_n^{\text{errw}} \tag{4.10}$$

hold. Consequently,

$$\begin{aligned} E[M_{n+1}^{\text{errw}}|\mathcal{F}_n] =& \frac{w_{e^*}(T_n)+1}{w_{v^*}(T_0)+2n+2}\left(P[A|\mathcal{F}_n] + P[B|\mathcal{F}_n]\right) + \frac{w_{e^*}(T_n)+2}{w_{v^*}(T_0)+2n+2} P[C|\mathcal{F}_n] \\ &+ \frac{w_{e^*}(T_n)}{w_{v^*}(T_0)+2n+2} P[D|\mathcal{F}_n] \\ =& \frac{w_{e^*}(T_n)+1}{w_{v^*}(T_0)+2n+2} \cdot 2P[A|\mathcal{F}_n] + \frac{w_{e^*}(T_n)+2}{w_{v^*}(T_0)+2n+2}\left(M_n^{\text{errw}} - P[A|\mathcal{F}_n]\right) \\ &+ \frac{w_{e^*}(T_n)}{w_{v^*}(T_0)+2n+2}\left(1 - M_n^{\text{errw}} - P[A|\mathcal{F}_n]\right) \\ =& M_n^{\text{errw}}. \end{aligned} \tag{4.11}$$

Hence, $(M_n^{\text{errw}})_{n\in\mathbb{N}_0}$ is a martingale. ■

Now consider the star subgraph $G^{\text{polya}} = (V^{\text{polya}}, E^{\text{polya}})$ of G consisting of v^* and its immediate neighbors and the edges connecting v^* with its immediate neighbors. We consider edge-reinforced random walk $(X_t^{\text{polya}})_{t\in\mathbb{N}_0}$ on G^{polya} starting in v^* with *random* initial weights $(a_e^{\text{polya}})_{e\in E^{\text{polya}}}$ having the same joint distribution as $(w_e(T_0))_{e\in E^{\text{polya}}}$, the weights of the edges $e \in E^{\text{polya}}$ for the edge-reinforced random walk on G at time T_0. Let us explain the difference between a_e and a_e^{polya}: If $v^* = \mathbf{0}$, then $(a_e^{\text{polya}})_{e\in E^{\text{polya}}}$ just equals $(a_e)_{e\in E^{\text{polya}}}$. But if $v^* \neq \mathbf{0}$, then v^* is entered the first time via a random edge $\tilde{e} = \{X_{T_0-1}, X_{T_0}\} \in E^{\text{polya}}$; then one has $w_{\tilde{e}}(T_0) = a_{\tilde{e}} + 1$, while $w_e(T_0) = a_e$ for all other edges $e \in E^{\text{polya}} \setminus \{\tilde{e}\}$.

For this reinforced random walk on the subgraph G^{polya}, we introduce similar notation as for the reinforced random walk on the full graph G: Let $w_e^{\text{polya}}(t)$, $e \in E^{\text{polya}}$, denote the weight of the edge e at time t, and $T_n^{\text{polya}} = 2n$ is the time of the $n+1$-st visit of the vertex v^*. We set

$$a_{v^*}^{\text{polya}} = \sum_{e\in E^{\text{polya}}} a_e^{\text{polya}} = \begin{cases} a_{v^*} & \text{if } \mathbf{0} = v^*, \\ a_{v^*} + 1 & \text{if } \mathbf{0} \neq v^*, \end{cases} \tag{4.12}$$

$$w_{v^*}^{\text{polya}}(t) = \sum_{e\in E^{\text{polya}}} w_e^{\text{polya}}(t), \tag{4.13}$$

$$M_n^{\text{polya}} := \frac{w_{e^*}^{\text{polya}}(T_n^{\text{polya}})}{w_{v^*}^{\text{polya}}(T_n^{\text{polya}})} = \frac{w_{e^*}^{\text{polya}}(2n)}{a_{v^*}^{\text{polya}} + 2n}. \tag{4.14}$$

Note that the denominator $a_{v^*}^{\text{polya}} + 2n$ is not random.

Remark 4.2 *In down-to-earth terms, $(M_n^{\text{polya}})_{n\in\mathbb{N}_0}$ is a Polya urn model, at least if the initial weights are natural numbers: Consider an urn containing initially $a_{e^*}^{\text{polya}}$ red and $a_{v^*}^{\text{polya}} - a_{e^*}^{\text{polya}}$ blue balls. In each discrete time step, draw a ball at random and put it back together with two balls of the same color. Then, the probability to draw a red ball in the $n+1$-st drawing equals M_n^{polya}. It is well-known that $(M_n^{\text{polya}})_{n\in\mathbb{N}_0}$ is a martingale.*

Furthermore, $(M_n^{\text{polya}})_{n\in\mathbb{N}_0}$ is a time-inhomogeneous Markov chain with random initial state $M_0^{\text{polya}} = a_{e^*}^{\text{polya}}/a_{v^*}^{\text{polya}}$ and transitions

$$M_{n+1}^{\text{polya}} = \begin{cases} \alpha_n M_n^{\text{polya}} + 1 - \alpha_n & \text{with probability } M_n^{\text{polya}}, \\ \\ \alpha_n M_n^{\text{polya}} & \text{with probability } 1 - M_n^{\text{polya}}, \end{cases} \tag{4.15}$$

where we have set

$$\alpha_n = \frac{a_{v^*}^{\text{polya}} + 2n}{a_{v^*}^{\text{polya}} + 2n + 2}. \tag{4.16}$$

The following lemma plays a central role in the whole article: It compares reinforced random walk on any finite graph with a much simpler Polya urn model.

Lemma 4.3 *For all $n \in \mathbb{N}$, one has $M_n^{\text{errw}} \lhd M_n^{\text{polya}}$.*

Proof. The proof is by induction. The claim is obvious for $n = 0$, since M_0^{errw} has the same distribution as M_0^{polya}. For the step $n \rightsquigarrow n+1$, we introduce the Polya urn transition kernel

$$K_n : [0,1] \times \mathcal{B}([0,1]) \to [0,1], \quad K_n(x,\cdot) := x\delta_{\alpha_n x + 1 - \alpha_n} + (1-x)\delta_{\alpha_n x}, \tag{4.17}$$

which was already described in (4.15). We define an auxiliary random variable M_{n+1}^{aux} by

$$M_n^{\text{aux}} = \begin{cases} \alpha_n M_n^{\text{errw}} + 1 - \alpha_n & \text{for } \{X_{T_n}, X_{T_n+1}\} = e^*, \\ \alpha_n M_n^{\text{errw}} & \text{otherwise.} \end{cases} \tag{4.18}$$

It has the law

$$\mathcal{L}_{M_n^{\text{aux}}} = \mathcal{L}_{M_n^{\text{errw}}} K_n. \tag{4.19}$$

For an interpretation of M_n^{aux} in the case of integer weights, consider an urn containing $w_{e^*}(T_n)$ red and $w_{v^*}(T_n) - w_{e^*}(T_n)$ blue balls. Draw a ball at random and put it back together with two balls of the same color. Then, the probability to draw a red ball in the next drawing equals M_n^{aux}.

We claim that

$$M_{n+1}^{\text{errw}} \lhd M_{n+1}^{\text{aux}} \quad \text{and} \quad M_{n+1}^{\text{aux}} \lhd M_{n+1}^{\text{polya}} \tag{4.20}$$

hold. Since the relation $\lhd$ is transitive by Lemma 3.2, these claims imply our goal $M_{n+1}^{\text{errw}} \lhd M_{n+1}^{\text{polya}}$.

To prove the first claim $M^{\text{errw}}_{n+1} \lhd M^{\text{aux}}_{n+1}$, we introduce the following σ-field:

$$\mathcal{G}_{n+1} = \sigma(w_e(T_k) : e \in E; k = 0, \dots, n+1). \tag{4.21}$$

(Note that $\mathcal{G}_{n+1}$ contains less information than $\mathcal{F}_{n+1}$. In particular, the orientation of the $n+1$st excursion from v^* of the random walker is not measurable with respect to $\mathcal{G}_{n+1}$.)

We now show that

$$M^{\text{errw}}_{n+1} = E[M^{\text{aux}}_{n+1} | \mathcal{G}_{n+1}] \tag{4.22}$$

holds, which implies the first claim $M^{\text{errw}}_{n+1} \lhd M^{\text{aux}}_{n+1}$. Indeed, M^{errw}_{n+1} is $\mathcal{G}_{n+1}$-measurable, and $M^{\text{errw}}_{n+1} = M^{\text{aux}}_{n+1}$ holds on the two $\mathcal{G}_{n+1}$-measurable events $\{w_{e^*}(T_{n+1}) = w_{e^*}(T_n)\}$ and $\{w_{e^*}(T_{n+1}) = w_{e^*}(T_n) + 2\}$. The complement of the union of these two events equals

$$\{w_{e^*}(T_{n+1}) = w_{e^*}(T_n) + 1\} = A \cup B, \tag{4.23}$$

with the two disjoint events A and B defined in (4.4) and (4.5). Note that M^{aux}_{n+1} equals $\alpha_n M^{\text{errw}}_n + 1 - \alpha_n$ on A and $\alpha_n M^{\text{errw}}_n$ on B. Consider again the transformation $\Pi_{n+1} : \Omega_{\mathbf{0}} \to \Omega_{\mathbf{0}}$ which reverts the orientation of the excursion between time T_n and time T_{n+1}, while keeping the rest of the path unchanged; Π_{n+1} was introduced after (4.8). For every random variable X taking values in $[0, 1]$, one has $E[X \circ \Pi_{n+1} | \mathcal{G}_{n+1}] = E[X | \mathcal{G}_{n+1}]$, i.e. $\mathcal{G}_{n+1}$ cannot distinguish between any path $\omega \in \Omega_{\mathbf{0}}$ and its partially time-reversed variant $\Pi_{n+1}(\omega)$. Furthermore,

$$\frac{1}{2}\left(M^{\text{aux}}_{n+1} + M^{\text{aux}}_{n+1} \circ \Pi_{n+1}\right) = M^{\text{errw}}_{n+1} \tag{4.24}$$

holds true. Using that M^{errw}_{n+1} is measurable with respect to $\mathcal{G}_{n+1}$, we conclude:

$$E[M^{\text{aux}}_{n+1} | \mathcal{G}_{n+1}] 1_{A \cup B} = \frac{1}{2} E[M^{\text{aux}}_{n+1} + M^{\text{aux}}_{n+1} \circ \Pi_{n+1} | \mathcal{G}_{n+1}] 1_{A \cup B} = M^{\text{errw}}_{n+1} 1_{A \cup B}, \tag{4.25}$$

which finishes the proof of the claim (4.22).

The second claim $M^{\text{aux}}_{n+1} \lhd M^{\text{polya}}_{n+1}$ in (4.20), i.e. $\mathcal{L}_{M^{\text{errw}}_n} K_n \lhd \mathcal{L}_{M^{\text{polya}}_n} K_n$, is a consequence of the induction hypothesis $M^{\text{errw}}_n \lhd M^{\text{polya}}_n$ and the following lemma. ■

Lemma 4.4 *If μ is a discrete distribution on $[0, 1]$ and ν is any distribution on $[0, 1]$ with $\mu \lhd \nu$, then $\mu K_n \lhd \nu K_n$ also holds.*

Proof. We prove the lemma first in the special case $\mu = \delta_x$, $x \in [0, 1]$. Thus we claim: If the distribution ν has the expectation x, then one has

$$x \delta_{\alpha_n x + 1 - \alpha_n} + (1 - x) \delta_{\alpha_n x} \lhd \nu K_n. \tag{4.26}$$

To prove this, let X and Y be $[0, 1]$-valued random variables with joint distribution $\nu \otimes K_n$, i.e. with joint distribution $\nu(dx)\, K_n(x, dy)$. We abbreviate the variance of ν by σ^2. Note that $Y \geq X$ is the same as $Y = \alpha_n X + 1 - \alpha_n$ and $Y < X$ is the same as $Y = \alpha_n X$ a.s. We get:

$$E[Y, Y \geq X] = \int_{[0,1]} (\alpha_n y + 1 - \alpha_n) y \, \nu(dy) = (\alpha_n x + 1 - \alpha_n) x + \alpha_n \sigma^2, \tag{4.27}$$

$$P[Y \geq X] = \int_{[0,1]} y \, \nu(dy) = x, \tag{4.28}$$

$$E[Y, Y < X] = \int_{[0,1]} \alpha_n y (1 - y) \, \nu(dy) = \alpha_n x (1 - x) - \alpha_n \sigma^2, \tag{4.29}$$

$$P[Y < X] = \int_{[0,1]} (1 - y) \, \nu(dy) = 1 - x. \tag{4.30}$$

Thus,

$$x \delta_{\alpha_n x + 1 - \alpha_n + \alpha_n \sigma^2 / x} + (1 - x) \delta_{\alpha_n x - \alpha_n \sigma^2 / (1 - x)} \sim E[Y | \sigma(\{Y \geq X\})] \triangleleft Y \sim \nu K_n. \tag{4.31}$$

Finally, we show that

$$x \delta_{\alpha_n x + 1 - \alpha_n} + (1 - x) \delta_{\alpha_n x} \triangleleft x \delta_{\alpha_n x + 1 - \alpha_n + \alpha_n \sigma^2 / x} + (1 - x) \delta_{\alpha_n x - \alpha_n \sigma^2 / (1 - x)}. \tag{4.32}$$

Since $\triangleleft$ is transitive, (4.31) and (4.32) imply the claim (4.26).

To verify (4.32), we show more generally that for any numbers $A \geq C \geq D \geq B$ with $xA + (1 - x)B = xC + (1 - x)D$, the relation

$$x \delta_C + (1 - x) \delta_D \triangleleft x \delta_A + (1 - x) \delta_B \tag{4.33}$$

holds. We apply this then with

$$A = \alpha_n x + 1 - \alpha_n + \alpha_n \frac{\sigma^2}{x} \geq C = \alpha_n x + 1 - \alpha_n \geq D = \alpha_n x \geq B = \alpha_n x - \alpha_n \frac{\sigma^2}{1 - x}. \tag{4.34}$$

Before proving (4.33) formally, let us interpret it intuitively: Suppose we have two bottles of vinegar, having volumes x and $1 - x$ and concentrations A and B, respectively. By partially mixing the content of the bottles, we can obtain two other bottles of vinegar, also having volumes x and $1 - x$, but concentrations C and D, respectively.

For $A = B$, there is nothing to show. Otherwise, take $\Omega = \{C, D\} \times \{A, B\}$ with the probability measure

$$\frac{x(C - B)}{A - B} \delta_{(C,A)} + \frac{(1 - x)}{A - B} \left\{ (D - B) \left[\delta_{(C,B)} + \delta_{(D,A)} \right] + (A - D) \delta_{(D,B)} \right\}, \tag{4.35}$$

and let ω_1 and ω_2 denote the projections on the first and second coordinate, respectively. Then, $E[\omega_2 | \omega_1] = \omega_1 \sim x \delta_C + (1 - x) \delta_D$ and $\omega_2 \sim x \delta_A + (1 - x) \delta_B$.

To prove the lemma in the general case, we write

$$\mu = \sum_{i=1}^{n} p_i \delta_{x_i}. \tag{4.36}$$

Let $X \sim \mu$ and $Y \sim \nu$ be random variables with $X = E[Y|X]$. For $i = 1, \dots, n$, we define

$$\nu_i = P[Y \in \cdot \mid X = x_i] \tag{4.37}$$

to denote the conditional distribution of Y given $X = x_i$. Then $\delta_{x_i} \triangleleft \nu_i$. By the already proven special case, we infer

$$\delta_{x_i} K_n \triangleleft \nu_i K_n. \tag{4.38}$$

The claim

$$\mu K_n = \sum_{i=1}^{n} p_i \delta_{x_i} K_n \triangleleft \sum_{i=1}^{n} p_i \nu_i K_n = \nu K_n \tag{4.39}$$

now follows from Lemma 3.2(c). ■

Let $\beta(a, b)$ denote the beta distribution with parameters $a, b > 0$.

Lemma 4.5 *As $n \to \infty$, the sequence $(M_n^{\text{polya}})_{n\in\mathbb{N}}$ converges almost surely to a random limit with distribution*

$$\begin{aligned}\mathbb{Q}^{\text{polya}} =& \mathbb{Q}^{\text{polya},G}_{\mathbf{0},a,v^*,e^*} \\ =& \begin{cases} \beta\left(\dfrac{a_{e^*}}{2}, \dfrac{a_{v^*} - a_{e^*}}{2}\right) & \text{if } v^* = \mathbf{0}, \\ P^G_{\mathbf{0},a}[w_{e^*}(T_0) = a_{e^*} + 1]\beta\left(\dfrac{a_{e^*}+1}{2}, \dfrac{a_{v^*} - a_{e^*}}{2}\right) & \\ +P^G_{\mathbf{0},a}[w_{e^*}(T_0) = a_{e^*}]\beta\left(\dfrac{a_{e^*}}{2}, \dfrac{a_{v^*} - a_{e^*}+1}{2}\right) & \text{if } v^* \neq \mathbf{0}. \end{cases}\end{aligned} \tag{4.40}$$

Proof. Recall Remark 4.2. To make the proof more intuitive, we use the language of the Polya urn model, by abuse of notation also in the case where the initial weights are not natural numbers. We consider an urn with a random initial composition consisting of $a_{e^*}^{\text{polya}}$ red and $a_{v^*}^{\text{polya}} - a_{e^*}^{\text{polya}}$ blue balls. The dynamics consists of drawing a ball at random from the urn and putting it back into the urn together with two balls of the same color. Conditioned on the initial number of balls $a_{e^*}^{\text{polya}}$ and $a_{v^*}^{\text{polya}} - a_{e^*}^{\text{polya}}$, the fraction of red balls in the urn after n drawings, namely M_n^{polya}, converges almost surely to a beta distribution with parameters $a_{e^*}^{\text{polya}}/2$ and $(a_{v^*}^{\text{polya}} - a_{e^*}^{\text{polya}})/2$; see e.g. [Dur04], Section 4.3b, page 238. The initial composition is distributed according to

$$\begin{aligned}&(a_{e^*}^{\text{polya}}, a_{v^*}^{\text{polya}} - a_{e^*}^{\text{polya}}) \\ =& \begin{cases} (a_{e^*}, a_{v^*} - a_{e^*}) & \text{if } v^* = \mathbf{0}, \\ (a_{e^*}+1, a_{v^*} - a_{e^*}) & \text{with probability } P^G_{\mathbf{0},a}[w_{e^*}(T_0) = a_{e^*}+1] \text{ if } v^* \neq \mathbf{0}, \\ (a_{e^*}, a_{v^*} - a_{e^*}+1) & \text{with probability } P^G_{\mathbf{0},a}[w_{e^*}(T_0) = a_{e^*}] \text{ if } v^* \neq \mathbf{0}. \end{cases}\end{aligned} \tag{4.41}$$

The claim follows. ■

It is well-known (see e.g. Theorem 3.1 of [Rol03]) that the edge-reinforced random walk on a *finite* graph G has the same distribution as a random walk in a random environment given by random weights on the edges. Let

$$\mathbb{Q}^{\mathrm{errw}} = \mathbb{Q}^G_{\mathbf{0},a} \tag{4.42}$$

denote the unique mixing measure on the simplex $\Delta := \left\{(x_e)_{e\in E} \in (0,1)^E : \sum_{e\in E} x_e = 1\right\}$.

Theorem 4.6 *For any finite graph G, any starting point $\mathbf{0}$, any initial weights a, any vertex v^*, and any edge e^* incident to v^*, the distribution $\mathbb{Q}^G_{\mathbf{0},a}[x_{e^*}/x_{v^*} \in \cdot]$ of x_{e^*}/x_{v^*} with respect to $\mathbb{Q}^G_{\mathbf{0},a}$ fulfills*

$$\mathbb{Q}^G_{\mathbf{0},a}\left[\frac{x_{e^*}}{x_{v^*}} \in \cdot\right] \lhd \mathbb{Q}^{\mathrm{polya},G}_{\mathbf{0},a,v^*,e^*}. \tag{4.43}$$

Proof. Recall the abbreviations (4.40) and (4.42). Both, $(M_n^{\mathrm{errw}})_{n\in\mathbb{N}}$ and $(M_n^{\mathrm{polya}})_{n\in\mathbb{N}}$ are bounded and hence uniformly integrable martingales. By Lemma 4.5, the limit of $(M_n^{\mathrm{polya}})_{n\in\mathbb{N}}$ has the distribution $\mathbb{Q}^{\mathrm{polya}}$. Since the edge-reinforced random walk on the finite graph G is a mixture of recurrent Markov chains, we know that $(w_t(e)/t)_{e\in E}$ converges almost surely as $t \to \infty$ to a limit with distribution $\mathbb{Q}^{\mathrm{errw}}$. Consequently, $M_n^{\mathrm{errw}} = w_{e^*}(T_n)/w_{v^*}(T_n)$ converges almost surely as $n \to \infty$ to a random limit with distribution $\mathbb{Q}^{\mathrm{errw}}[x_{e^*}/x_{v^*} \in \cdot]$. Using that $M_n^{\mathrm{errw}} \lhd M_n^{\mathrm{polya}}$ for all n by Lemma 4.3, the claim follows from Lemma 3.2(b). ■

5 Proofs of the main results

First, we prove the following uniform tail estimates for x_e/x_v:

Lemma 5.1 *Theorem 2.3 holds for all* finite *graphs G.*

Proof. Recall the abbreviation (4.42). Since $\mathbb{Q}^{\mathrm{errw}}[x_e/x_v \le \varepsilon] \le 1$ and $\mathbb{Q}^{\mathrm{errw}}[x_e/x_v \ge 1-\varepsilon] \le 1$, it suffices to prove the estimates (2.3) for all $\varepsilon \in (0, 1/3)$. The claim follows then with possibly larger constants c_1 and c_2.

Let G be a finite graph, let $\mathbf{0}$ be any vertex, and let $a \in \mathbb{R}_+^E$ be any initial weights. In this case, it is known that the measure $\mathbb{Q}^{\mathrm{errw}}$ satisfies the assertions of Theorem 2.2 (see e.g. Theorem 3.1 of [Rol03]).

Let $v \in V$ and let $e \in E$ be an edge incident to v. If $\mathrm{degree}(v) = 1$, then $x_e/x_v = 1$ and $a_v = a_e$. Hence, the left estimate in (2.3) always holds and the right estimate in (2.3) holds for any $c_2 \ge 1$. In the following, we assume $\mathrm{degree}(v) \ge 2$. Combining Theorem 4.6 with Lemma 3.2(d) yields for any convex bounded function $f : [0,1] \to \mathbb{R}$ the inequality

$$E_{\mathbb{Q}^{\mathrm{errw}}}\left[f\left(\frac{x_e}{x_v}\right)\right] \le \int_0^1 f(x)\,\mathbb{Q}^{\mathrm{polya}}(dx). \tag{5.1}$$

Let $\varepsilon \in (0, 1/3)$. We apply the last inequality with

$$f(x) = \begin{cases} 1 - \dfrac{x}{2\varepsilon} & \text{for } 0 \le x \le 2\varepsilon, \\ 0 & \text{for } 2\varepsilon \le x \le 1 \end{cases} \tag{5.2}$$

to obtain

$$\begin{aligned} \frac{1}{2}\mathbb{Q}^{\text{errw}}\left[\frac{x_e}{x_v} \le \varepsilon\right] &\le E_{\mathbb{Q}^{\text{errw}}}\left[f\left(\frac{x_e}{x_v}\right)\right] \\ &\le \int_0^1 f(x)\,\mathbb{Q}^{\text{polya}}(dx) \le \mathbb{Q}^{\text{polya}}[[0, 2\varepsilon]]. \end{aligned} \tag{5.3}$$

Recall the definition (4.40) of $\mathbb{Q}^{\text{polya}}$. In the case $v \neq \mathbf{0}$, we first estimate

$$I_1 := P^G_{\mathbf{0},a}[w_e(T_0) = a_e + 1] \cdot \frac{1}{\mathrm{B}\left(\frac{a_e+1}{2}, \frac{a_v - a_e}{2}\right)} \int_0^{2\varepsilon} x^{(a_e+1)/2-1}(1-x)^{(a_v-a_e)/2-1}\,dx, \tag{5.4}$$

where B denotes the Beta function. Clearly, $P^G_{\mathbf{0},a}[w_e(T_0) = a_e + 1] \le 1$. Since the Beta function is continuous, the normalizing constant of the beta distribution in (5.4) depends continuously on a_v and a_e. To bound the last integral, we use $1/3 \le 1 - 2\varepsilon \le 1 - x \le 1$ for all $x \in (0, 2\varepsilon)$. This yields the estimate

$$I_1 \le c_4(a_v, a_e)\varepsilon^{(a_e+1)/2} \le c_4(a_v, a_e)\varepsilon^{a_e/2} \tag{5.5}$$

with a constant $c_4(a_v, a_e) > 0$ depending continuously on a_v and a_e and not depending on any other quantity, in particular, not depending on G, $\mathbf{0}$, v and e. An analogous argument yields

$$\begin{aligned} I_2 :=& P^G_{\mathbf{0},a}[w_e(T_0) = a_e] \cdot \frac{1}{\mathrm{B}\left(\frac{a_e}{2}, \frac{a_v - a_e + 1}{2}\right)} \int_0^{2\varepsilon} x^{a_e/2-1}(1-x)^{(a_v-a_e+1)/2-1}\,dx \\ \le& c_5(a_v, a_e)\varepsilon^{a_e/2} \end{aligned} \tag{5.6}$$

with a constant $c_5(a_v, a_e) > 0$ depending continuously on a_e and a_v. Combining (5.5) and (5.6), we conclude

$$\mathbb{Q}^{\text{polya}}[[0, 2\varepsilon]] \le [c_4(a_v, a_e) + c_5(a_v, a_e)]\varepsilon^{a_e/2}. \tag{5.7}$$

In the case $v = \mathbf{0}$, a similar argument yields the bound (5.7) with a different constant. This proves the estimates for the lower tail probabilities in Theorem 2.3.

To bound the upper tail probabilities, we apply (5.1) with

$$f(x) = \begin{cases} 0 & \text{for } 0 \le x \le 1 - 2\varepsilon, \\ 1 - \dfrac{1-x}{2\varepsilon} & \text{for } 1 - 2\varepsilon \le x \le 1. \end{cases} \tag{5.8}$$

This yields

$$\frac{1}{2}\mathbb{Q}^{\text{errw}}\left[\frac{x_e}{x_v}\geq 1-\varepsilon\right]\leq E_{\mathbb{Q}^{\text{errw}}}\left[f\left(\frac{x_e}{x_v}\right)\right]$$
$$\leq\int_0^1 f(x)\,\mathbb{Q}^{\text{polya}}(dx)\leq \mathbb{Q}^{\text{polya}}[[1-2\varepsilon,1]]. \tag{5.9}$$

We indicate only the estimate in the case $v=\mathbf{0}$. In this case, one gets

$$\mathbb{Q}^{\text{polya}}[[1-2\varepsilon,1]] = \frac{1}{\mathrm{B}\left(\frac{a_e}{2},\frac{a_v-a_e}{2}\right)}\int_{1-2\varepsilon}^1 x^{a_e/2-1}(1-x)^{(a_v-a_e)/2-1}\,dx$$
$$= \frac{1}{\mathrm{B}\left(\frac{a_e}{2},\frac{a_v-a_e}{2}\right)}\int_0^{2\varepsilon} x^{(a_v-a_e)/2-1}(1-x)^{a_e/2-1}\,dx. \tag{5.10}$$

Again, similar arguments as above imply

$$\mathbb{Q}^{\text{polya}}[[1-2\varepsilon,1]]\leq c_6(a_v,a_e)\varepsilon^{(a_v-a_e)/2} \tag{5.11}$$

with a constant $c_6(a_v,a_e)>0$ depending continuously on a_v and a_e. ■

Lemma 5.2 *Theorem 2.4 holds for all* finite *graphs G.*

Proof. Given a compact set $K\subset(0,\infty)$, we set

$$c_3(K)=\sup_{\substack{a,a'\in K\\ a\geq a'}} c_1(a,a')<\infty, \tag{5.12}$$

where $c_1(a,a')$ is taken from Lemma 5.1, the version of Theorem 2.3 for finite graphs. Now take a finite graph $G=(V,E)$, a starting point $\mathbf{0}\in V$, and initial weights $a\in\mathbb{R}_+^E$. By Lemma 5.1, $\mathbb{Q}^{\text{errw}}$ has the properties specified in Theorems 2.2 and 2.3. Given any edges $e,f\in E$, any path

$$\xrightarrow{e=e_0} v_1 \xrightarrow{e_1} v_2 \xrightarrow{e_2} \dots \xrightarrow{e_{l-1}} v_l \xrightarrow{e_l=f} \tag{5.13}$$

from e to f, any $M>0$, and assuming that the initial weights $(a_{e'})_{e'\in E}$ satisfy $a_{e_i}\in K$ and $a_{v_j}=\sum_{e'\ni v_j}a_{e'}\in K$ ($i=0,\dots,l$; $j=1,\dots,l$), we estimate, using Lemma 5.1 in the last but one step:

$$\mathbb{Q}^{\text{errw}}[x_e\geq Mx_f]\leq\mathbb{Q}^{\text{errw}}[x_{e_{i-1}}\geq M^{1/l}x_{e_i}\text{ for at least one }i=1,\dots,l]$$
$$\leq\sum_{i=1}^l\mathbb{Q}^{\text{errw}}[x_{e_{i-1}}\geq M^{1/l}x_{e_i}]\leq\sum_{i=1}^l\mathbb{Q}^{\text{errw}}[x_{v_i}\geq M^{1/l}x_{e_i}]$$
$$\leq\sum_{i=1}^l c_1(a_{v_i},a_{e_i})M^{-a_{e_i}/2l}\leq c_3(K)lM^{-\gamma}, \tag{5.14}$$

where γ is defined in (2.6). ■

Let $G = (V, E)$ be a locally finite graph, and let $G_n = (V_n, E_n)$ be finite connected subgraphs of G with $V_n \uparrow V$ and $E_n \uparrow E$, and $\mathbf{0} \in V_n$ for all n. We denote by $\mathbb{Q}_n^{\mathrm{errw}} := \mathbb{Q}_{\mathbf{0},a}^{G_n}$ the unique mixing measure as defined in (4.42) for the edge-reinforced random walk on the graph G_n. By its definition, $\mathbb{Q}_n^{\mathrm{errw}}$ is a measure on the simplex $\Delta_n = \left\{(x_e)_{e\in E_n} \in (0,1)^{E_n} : \sum_{e\in E_n} x_e = 1\right\}$. We introduce the canonical projections $x_e : \Delta_n \to (0,1)$, $e \in E_n$, suppressing the dependence on n in the notation.

Fix $e_0 \in E$. We assume without loss of generality that $e_0 \in E_n$ for all n. We set $\tilde{x}_e = x_e/x_{e_0}$ for $e \in E_n$. In particular, $\tilde{x}_{e_0} = 1$ holds. The weights $(x_e)_{e\in E_n}$ and $(\tilde{x}_e)_{e\in E_n}$ are multiples of each other; hence they induce the same Markov chain ($Q_{\mathbf{0},x}^{G_n} = Q_{\mathbf{0},\tilde{x}}^{G_n}$). Thus, we can use the law of $(\tilde{x}_e)_{e\in E_n}$ with respect to $\mathbb{Q}_n^{\mathrm{errw}}$ as mixing measure for the edge-reinforced random walk on G_n.

Lemma 5.3 *There exists a strictly increasing sequence $(n(k))_{k\in\mathbb{N}}$ in $\mathbb{N}$ such that for all finite $F \subseteq E$, the law of $(\tilde{x}_e)_{e\in F}$ with respect to $\mathbb{Q}_{n(k)}^{\mathrm{errw}}$ converges weakly as $k \to \infty$ to a distribution, supported on $(0,\infty)^F$.*

Proof. We prove first that for any fixed $k \in \mathbb{N}$, there exists $n_0 = n_0(k) \geq k$ such that the distributions $\mathbb{Q}_n^{\mathrm{errw}}[(\ln \tilde{x}_e)_{e\in E_k} \in \cdot]$, $n \geq n_0$, are tight. Fix $e \in E_k$ and a path π:

$$\xrightarrow{e_0} v_1 \xrightarrow{e_1} v_2 \xrightarrow{e_2} \ldots \xrightarrow{e_{l-1}} v_l \xrightarrow{e_l = e} \tag{5.15}$$

from e_0 to e in E_k. Choose $n_0 \in \mathbb{N}$ so large that for any v_j in the path π, any edge $e \in E$ incident to v_j belongs to E_{n_0}. Let $K := \{a_{e_i}, a_{v_j} : \ i = 0, \ldots, l, \ j = 1, \ldots, l\}$ and let γ be as in (2.6). Then, by Lemma 5.2 (the finite graph version of Theorem 2.4), taking $c_3(K)$ as in that lemma, we obtain for all $n \geq n_0$ and all $M > 0$:

$$\mathbb{Q}_n^{\mathrm{errw}}[\tilde{x}_e \geq M] = \mathbb{Q}_n^{\mathrm{errw}}[x_e \geq M x_{e_0}] \leq c_3(K) l M^{-\gamma} \tag{5.16}$$

and

$$\mathbb{Q}_n^{\mathrm{errw}}[\tilde{x}_e \leq M^{-1}] = \mathbb{Q}_n^{\mathrm{errw}}[x_{e_0} \geq M x_e] \leq c_3(K) l M^{-\gamma}. \tag{5.17}$$

This implies tightness of $\mathbb{Q}_n^{\mathrm{errw}}[(\ln \tilde{x}_e)_{e\in E_k} \in \cdot]$, $n \geq n_0$.

Using a compactness argument, the tightness proven above allows us to construct, by recursion over k, a sequence of strictly increasing sequences $m_k = (m_k(i))_{i\in\mathbb{N}}$, $k \in \mathbb{N}$, with values in $\mathbb{N}$, such that for all $k \in \mathbb{N}$, m_{k+1} is a subsequence of m_k and the sequence $\mathbb{Q}_{m_k(i)}^{\mathrm{errw}}[(\tilde{x}_e)_{e\in E_k} \in \cdot]$ converges weakly as $i \to \infty$. By (5.17), the limiting distribution is supported on $(0,\infty)^{E_k}$. Then, the diagonal sequence $n(k) := m_k(k)$, $k \in \mathbb{N}$, fulfills the claim of the lemma. ■

Proof of Theorem 2.2. Let $\mathbb{Q}_{\mathbf{0},a}^{G}$ be the limit of a weakly convergent subsequence $\left(\mathbb{Q}_{n(k)}^{\mathrm{errw}}\left[(\tilde{x}_e)_{e\in E_k} \in \cdot\right]\right)_{k\in\mathbb{N}}$ as in Lemma 5.3. By Lemma 5.3, $\mathbb{Q}_{\mathbf{0},a}^{G}$ is supported on $(0,\infty)^E$.

Let $\pi = (\mathbf{0}, v_1, \ldots, v_l)$ be a finite path in G. Then, for all k sufficiently large,

$$\begin{aligned} P^G_{\mathbf{0},a}\left[(X_s)_{s=0\ldots l} = \pi\right] =& P^{G_{n(k)}}_{\mathbf{0},a}\left[(X_s)_{s=0\ldots l} = \pi\right] \\ =& \int Q^{G_{n(k)}}_{\mathbf{0},\tilde{x}}\left[(X_s)_{s=0\ldots l} = \pi\right] \mathbb{Q}^{\text{errw}}_{n(k)}(dx). \end{aligned} \tag{5.18}$$

Note that for all k large enough and all $x \in (0,\infty)^E$, we have $Q^{G_{n(k)}}_{\mathbf{0},\tilde{x}}\left[(X_s)_{s=0\ldots l} = \pi\right] = Q^G_{\mathbf{0},\tilde{x}}\left[(X_s)_{s=0\ldots l} = \pi\right]$. Since $x \mapsto Q^G_{\mathbf{0},\tilde{x}}\left[(X_s)_{s=0\ldots l} = \pi\right]$ is a bounded and continuous cylinder function, taking the limit as $k \to \infty$ in (5.18) yields

$$P^G_{\mathbf{0},a}\left[(X_s)_{s=0\ldots k} = \pi\right] = \int Q^G_{\mathbf{0},\tilde{x}}\left[(X_s)_{s=0\ldots l} = \pi\right] \mathbb{Q}^G_{\mathbf{0},a}(dx). \tag{5.19}$$

The events of the form $\{(X_s)_{s=0\ldots l} = \pi\}$ form a closed system with respect to intersection and generate the canonical σ-algebra on $\Omega_{\mathbf{0}}$. Thus, the claim (2.2) follows from (5.19). ■

Proof of Theorem 2.3 and 2.4. By Lemmas 5.1 and 5.2, the claimed estimates hold for the measures $\mathbb{Q}^{\text{errw}}_{n(k)}\left[(\tilde{x}_e)_{e\in E_k} \in \cdot\right]$, uniformly in $k \in \mathbb{N}$. Lemma 5.3 allows us to take the limit as $k \to \infty$ to obtain

$$\mathbb{Q}^G_{\mathbf{0},a}\left[\frac{x_e}{x_v} < \varepsilon\right] \le c_1 \varepsilon^{a_e/2} \qquad \text{and} \qquad \mathbb{Q}^G_{\mathbf{0},a}\left[\frac{x_e}{x_v} > 1-\varepsilon\right] \le c_2 \varepsilon^{(a_v - a_e)/2} \tag{5.20}$$

for all $\varepsilon > 0$. Taking the limit $\varepsilon \downarrow \tilde{\varepsilon}$, the claims (2.3) follow with ε replaced by $\tilde{\varepsilon}$. An analogous argument yields (2.5). ■

Proof of Theorem 2.1. By Theorem 2.2, the edge-reinforced random walk is a mixture of irreducible Markov chains. Since for every irreducible Markov chain the statements (a)–(d) are equivalent, the claim follows. ■

References

[CD86] D. Coppersmith and P. Diaconis. Random walk with reinforcement. Unpublished manuscript, 1986.

[DF80] P. Diaconis and D. Freedman. de Finetti's theorem for Markov chains. *Ann. Probab.*, 8(1):115–130, 1980.

[Dur04] R. Durrett. *Probability: Theory and Examples.* Duxbury advanced series, Third edition, 2004.

[KR00] M.S. Keane and S.W.W. Rolles. Edge-reinforced random walk on finite graphs. In *Infinite dimensional stochastic analysis (Amsterdam, 1999)*, pages 217–234. R. Neth. Acad. Arts Sci., Amsterdam, 2000.

[MR05a] F. Merkl and S.W.W. Rolles. Asymptotic behavior of edge-reinforced random walks. Preprint. Available from http://www.win.tue.nl/~srolles, 2005.

[MR05b] F. Merkl and S.W.W. Rolles. Edge-reinforced random walk on a ladder. To appear in *The Annals of Probability*, 2005.

[MR05c] F. Merkl and S.W.W. Rolles. Linearly edge-reinforced random walks. To appear in *Dynamics and Stochastics: Festschrift in the honor of Michael Keane*, 2005.

[Rol03] S.W.W. Rolles. How edge-reinforced random walk arises naturally. *Probab. Theory Related Fields*, 126(2):243–260, 2003.

[Rol05] S.W.W. Rolles. On the recurrence of edge-reinforced random walk on $\mathbb{Z} \times G$. To appear in *Probability Theory and Related Fields*, 2005.

[Str65] V. Strassen. The existence of probability measures with given marginals. *Ann. Math. Statist*, 36:423–439, 1965.