The Annals of Probability
2011, Vol. 39, No. 6, 2178–2223
DOI: 10.1214/10-AOP609

DYNAMICS OF VERTEX-REINFORCED RANDOM WALKS

BY MICHEL BENAÏM[1] AND PIERRE TARRÈS[2]

Université de Neuchâtel and CNRS, Université de Toulouse

We generalize a result from Volkov [*Ann. Probab.* **29** (2001) 66–91] and prove that, on a large class of locally finite connected graphs of bounded degree $(G, \sim)$ and symmetric reinforcement matrices $a = (a_{i,j})_{i,j \in G}$, the vertex-reinforced random walk (VRRW) eventually localizes with positive probability on subsets which consist of a complete d-partite subgraph with possible loops plus its outer boundary.

We first show that, in general, any stable equilibrium of a linear symmetric *replicator* dynamics with positive payoffs on a graph G satisfies the property that its support is a complete d-partite subgraph of G with possible loops, for some $d \geq 1$. This result is used here for the study of VRRWs, but also applies to other contexts such as evolutionary models in population genetics and game theory.

Next we generalize the result of Pemantle [*Probab. Theory Related Fields* **92** (1992) 117–136] and Benaïm [*Ann. Probab.* **25** (1997) 361–392] relating the asymptotic behavior of the VRRW to *replicator* dynamics. This enables us to conclude that, given any neighborhood of a strictly stable equilibrium with support S, the following event occurs with positive probability: the walk localizes on $S \cup \partial S$ (where ∂S is the outer boundary of S) and the density of occupation of the VRRW converges, with polynomial rate, to a strictly stable equilibrium in this neighborhood.

1. General introduction. Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $(G, \sim)$ be a locally finite connected symmetric graph, and let G be its vertex set, by a slight abuse of notation. Let $a := (a_{i,j})_{i,j \in G}$ be a symmetric (i.e., $a_{i,j} = a_{j,i}$) matrix with nonnegative entries such that, for all $i, j \in G$,

$$i \sim j \quad \Leftrightarrow \quad a_{i,j} > 0.$$

Let $(X_n)_{n \in \mathbb{N}}$ be a process taking values in G. Let $\mathbb{F} = (\mathcal{F}_n)_{n \in \mathbb{N}}$ denote the filtration generated by the process, that is, $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$ for all $n \in \mathbb{N}$.

Received November 2008; revised September 2010.

[1]Supported by the Swiss National Foundation Grant 200020-120218/1.

[2]On leave from the University of Oxford. Supported in part by the Swiss National Foundation Grant 200021-1036251/1, and by a Leverhulme Prize.

MSC2010 subject classifications. Primary 60G50; secondary 60J10, 60K35, 34F05.

Key words and phrases. Reinforced random walks, martingales, random perturbations of dynamical systems, replicator dynamics, entropy function.

For any $i \in G$, let $Z_n(i)$ be the number of times that the process visits site i up through time $n \in \mathbb{N} \cup \{\infty\}$, that is,

$$Z_n(i) = Z_0(i) + \sum_{m=0}^{n} \mathbb{1}_{\{X_m = i\}}$$

with the convention that, before initial time 0, a site $i \in G$ has already been visited $Z_0(i) \in \mathbb{R}_+ \setminus \{0\}$ times.

Then $(X_n)_{n\in\mathbb{N}}$ is called a *Vertex-Reinforced Random Walk* (*VRRW*) *with starting point* $v_0 \in G$ *and reinforcement matrix* $a := (a_{i,j})_{i,j\in G}$ if $X_0 = v_0$ and, for all $n \in \mathbb{N}$,

$$\mathbb{P}(X_{n+1} = j | \mathcal{F}_n) = \mathbb{1}_{\{j \sim X_n\}} \frac{a_{X_n,j} Z_n(j)}{\sum_{k\sim X_n} a_{X_n,k} Z_n(k)}.$$

These non-Markovian random walks were introduced in 1988 by Pemantle [13] during his PhD with Diaconis, in the spirit of the model of Edge-Reinforced Random Walks by Coppersmith and Diaconis in 1987 [4], where the weights accumulate on edges rather than vertices.

Vertex-reinforced random walks were first studied in the articles of Pemantle [14] and Benaïm [2] exploring some features of their asymptotic behavior on finite graphs and, in particular, relating the behavior of the empirical occupation measure to solutions of ordinary differential equations when the graph is complete (i.e., when all vertices are related together), as explained below. On the integers $\mathbb{Z}$, Pemantle and Volkov [16] showed that the VRRW a.s. visits only finitely many vertices and, with positive probability, eventually gets stuck on five vertices, and Tarrès [18] proved that this localization on five points is the almost sure behavior.

On arbitrary graphs, Volkov [23] proved that VRRW with reinforcement coefficients $a_{i,j} = \mathbb{1}_{i\sim j}$, $i, j \in G$ (again, $i \sim j$ meaning that i and j are neighbors in the nonoriented graph G), localizes with positive probability on some specific finite subgraphs; we recall this result in Theorem 4 below, in a generalized version. More recently, Limic and Volkov [8] study VRRW with the same specific type of reinforcement on complete-like graphs (i.e., complete graphs ornamented by finitely many leaves at each vertex) and show that, almost surely, the VRRW spends positive (and equal) proportions of time on each of its nonleaf vertices.

The VRRW with polynomial reinforcement [i.e., with the probability to visit a vertex proportional to a function $W(n) = n^\rho$ of its current number of visits] has recently been studied by Volkov on $\mathbb{Z}$ [24]. In the superlinear case (i.e., $\rho > 1$), the walk a.s. visits two vertices infinitely often. In the sublinear case (i.e., $\rho < 1$), the walk a.s. either visits infinitely many sites infinitely often or is transient; it is conjectured that the latter behavior cannot occur, and that, in fact, all integers are infinitely often visited.

The similar Edge-Reinforced Random Walks and, more generally, self-interacting processes, whether in discrete or continuous time/space, have been extensively studied in recent years. They are sometimes used as models involving self-organization or learning behavior, in physics, biology or economics. We propose a short review of the subject in the introduction of [12]. For more detailed overviews, we refer the reader to surveys by Davis [5], Merkl and Rolles [10], Pemantle [15] and Tóth [19], each analyzing the subject from a different perspective.

Let us first recall a few well-known observations on the study of Vertex-Reinforced Random Walks, and, in particular, the heuristics for relating its behavior to solutions of ordinary differential equations when the graph is finite and complete (i.e., when all vertices are related together), as done in Pemantle [14] and Benaïm [2].

Let us introduce some preliminary notation, without any further assumption on $(G, \sim)$ locally finite connected symmetric graph, possibly infinite. For all $x = (x_i)_{i\in G} \in \mathbb{R}^G$, let

$$S(x) := \{i \in G / x_i \neq 0\}$$

be its support. For all $x \in \mathbb{R}^G$ such that $S(x)$ is finite, let

$$N_i(x) := \sum_{j\in G, j\sim i} a_{i,j} x_j,\ H(x) = \sum_{i,j\in G, i\sim j} a_{i,j} x_i x_j = \sum_{i\in G} x_i N_i(x) \tag{1}$$

and, if $H(x) \neq 0$, let

$$\pi(x) := \left(\frac{x_i N_i(x)}{H(x)}\right)_{i\in G}. \tag{2}$$

Let

$$\Theta := \{x \in \mathbb{R}^G \text{ s.t. } |S(x)| < \infty\},$$

and let

$$\Delta := \left\{x \in \mathbb{R}_+^G \cap \Theta \text{ s.t. } \sum_{i\in G} x_i = 1\right\}$$

be the nonnegative simplex restricted to elements x of finite support.

For all $n \in \mathbb{N}$, let

$$y(n) := \left(\frac{Z_n(i) - Z_0(i)}{n}\right)_{i\in G} \in \Theta \cap \Delta$$

and, if G is finite, let

$$x(n) := \left(\frac{Z_n(i)}{n + n_0}\right)_{i\in G},$$

where $n_0 := \sum_{j\in G} Z_0(j) > 0$: $x(n)$ [resp. $y(n)$] is the vector of density of occupation of the random walk at time n, with the convention that site i has been visited $Z_0(i)$ (resp., 0) times at time 0.

Assume, for the sake of simplicity in the following heuristic argument, that G is a finite graph. Let $L \gg 1$. For all $n \in \mathbb{N}$, the goal is to compare $x(n + L)$ to $x(n)$. If $n \gg L$, then the VRRW between these times behaves as though $x(k)$, $n \leq k \leq n + L$, were constant, and hence approximates a Markov chain which we call $M(x(n))$.

Then $\pi(x(n)) \in \Delta$ is the invariant measure of $M(x(n))$, which is reversible [trivially $H(x(n)) > 0$ since $x(n)_i > 0$ for all i, so that $\pi(x(n))$ is well defined]. If L is large enough, then, by the ergodic theorem, the local occupation density between these times will be close to $\pi(x(n))$. This means that

$$(n+L)x(n+L) \approx nx(n) + L\pi(x(n)), \tag{3}$$

hence,

$$x(n+L) - x(n) \approx \frac{L}{nH(x(n))} F(x(n)), \tag{4}$$

where

$$F(x) = (x_i[N_i(x) - H(x)])_{i \in G}. \tag{5}$$

Up to an adequate time change, $(x(k))_{k \in \mathbb{N}}$ should approximate solutions of the ordinary differential equation on Δ,

$$\frac{dx}{dt} = F(x), \tag{6}$$

also known as the linear *replicator* equation in population genetics and game theory.

However, the requirement that L be large enough so that the local occupation measure of the Markov Chain approximates the invariant measure $\pi(x(n))$ competes with the other requirement that L be small enough so that the probability transitions of this Markov Chain still match the ones of the VRRW, so that the heuristics breaks down when the relaxation time of the Markov Chain is of the order of n, which can happen in general on noncomplete graphs and is actually consistent with the fact that the walk will indeed eventually localize on a small subset. An illustration of how such a behavior can occur is given in the proof of Lemma 2.8 in Tarrès [18]. The study of the a.s. asymptotic behavior of the VRRW on an infinite graph is even more involved in general.

Let us yet study the replicator differential equation (6) associated to the random walk on Δ for general locally finite symmetric graphs $(G, \sim)$.

It is easy to check that H is a strict Lyapounov function for (6) on Δ, that is, strictly increasing on the nonconstant solutions of this equation: if $x(t) = (x_i(t))_{i \in G}$ is the solution at time t, starting at $x(0) := x_0$, then

$$\frac{d}{dt} H(x(t)) = \sum_{i \in S(x)} \frac{\partial H}{\partial x_i}(x(t)) F(x(t))_i = J(x(t)),$$

where, for all $x \in \Delta$,

$$J(x) := 2 \sum_{i \in S(x)} N_i(x) F(x)_i = 2 \sum_{i \in S(x)} x_i \bigl(N_i(x) - H(x)\bigr)^2. \tag{7}$$

Note that the restriction of H to the equilibria of (6) takes finitely many values if G is finite (see [14], e.g.).

Let us now deal with the equilibria of this differential equation: a point $x = (x_i)_{i \in G} \in \Delta$ is called an *equilibrium* if and only if $F(x) = 0$. An equilibrium is called *feasible* provided $H(x) \neq 0$.

On a finite graph G, any equilibrium point $x \in \Delta$ of $(x(n))_{n \in \mathbb{N}}$ is feasible: for all $n \in \mathbb{N}$ and $i \in G$, $Z_n(i) \leq \sum_{j \sim i} Z_n(j) + n_0$, so that x would satisfy $N_i(x) \geq (\min_{j \sim i} a_{i,j}) x_i$ for all $i \in G$, hence,

$$H(x) \geq \Bigl(\min_{\{i,j \in S(x), j \sim i\}} a_{i,j}\Bigr) \sum_{i \in S(x)} x_i^2 \geq \frac{\min_{\{i,j \in S(x), j \sim i\}} a_{i,j}}{|S(x)|} \tag{8}$$

by the Cauchy–Schwarz inequality.

By a slight abuse of notation, we let $DF(x) = (\partial F_i / \partial x_j)_{i,j \in G}$ denote both the *Jacobian matrix* of F at x, and the corresponding linear operator on Θ. Since Δ is invariant under the flow induced by F, the tangent space

$$T\Delta := \Bigl\{ y \in \Theta \Big/ \sum_{i \in G} y_i = 0 \Bigr\}$$

is invariant under $DF(x)$. We let $DF(x)|_{T\Delta}$ denote the restriction of the operator $DF(x)$ to $T\Delta$.

When x is an equilibrium, it is easily seen that $DF(x)$ has real eigenvalues (see Lemma 1). Such an equilibrium is called *hyperbolic* (resp., *a sink*) provided $DF(x)|_{T\Delta}$ has nonzero (resp., negative) eigenvalues. It is called a *stable equilibrium* if $DF(x)|_{T\Delta}$ has nonpositive eigenvalues. Note that every sink is stable. Furthermore, by Theorem 1 below, every stable equilibrium is feasible.

We will sometimes abuse notation and identify arbitrary subsets H of G to the corresponding subgraph $(H, \sim)$. Given $i \in G$ and a subset A of G, we write $i \sim A$ if there exists $j \in A$ such that $i \sim j$. Given two subsets R and S of G, we let

$$\partial R = \{j \in G \setminus R : j \sim R\}, \qquad \partial_S R = \{j \in S \setminus R : j \sim R\};$$

∂R is called the *outer boundary* of R.

Given e, $e' \in E(G)$, we write $e \sim e'$ if e and e' have at least one vertex in common.

A site $i \in G$ will be called a loop if $i \sim i$, and we will say that a subset H contains a loop iff there exists a site in it which is a loop.

We will say that x is a *strictly stable equilibrium* if it is stable and, furthermore, for all $i \in \partial S(x)$, $N_i(x) < H(x)$. We let $\mathcal{E}_s$ be the set of strictly stable equilibria of (6) in Δ. Note that x stable already implies $N_i(x) \leq H(x)$ for all $i \in \partial S(x)$, by Lemma 1.

Given $d \geq 1$, subgraph $(S, \sim)$ of $(G, \sim)$ will be called a *complete d-partite graph with possible loops*, if $(S, \sim)$ is a d-partite graph on which some loops have possibly been added. That is,

$$S = V_1 \cup \cdots \cup V_d$$

with:

(i) $\forall p \in \{1, \dots, d\}, \forall i, j \in V_p$, if $i \neq j$ then $i \nsim j$.

(ii) $\forall p, q \in \{1, \dots, d\}$, $p \neq q$, $\forall i \in V_p$, $\forall j \in V_q$, $i \sim j$.

For all $S \subseteq G$, let $(\mathrm{P})_S$ be the following predicate:

$(\mathrm{P})_S$(a) $(S, \sim)$ is a complete d-partite graph with possible loops.
$(\mathrm{P})_S$(b) If $i \sim i$ for some $i \in S$, then the partition containing i is a singleton.
$(\mathrm{P})_S$(c) If V_p, $1 \le p \le d$ are its d partitions, then for all $p, q \in \{1, \dots, d\}$ and $i, i' \in V_p$, $j, j' \in V_q$, $a_{i,j} = a_{i',j'}$.

In the following Theorems 1–4 and Propositions 2 and 3, we only assume the graph $(G, \sim)$ to be symmetric and locally finite, without any further conditions than the ones mentioned in the statements.

THEOREM 1. *If $x \in \Delta$ is a stable equilibrium of* (6), *then x is feasible and $(\mathrm{P})_{S(x)}$ holds.*

In the case $a = (a_{i,j})_{i,j \in G} = (\mathbb{1}_{i \sim j})_{i,j \in G}$ the following Theorem 2 provides a necessary and sufficient condition for $x \in \Delta$ being a stable equilibrium. Theorems 1 and 2 are proved in Section 2.2.

THEOREM 2. *Assume $a_{i,j} = \mathbb{1}_{i \sim j}$ for all $i, j \in G$, and let $x = (x_i)_{i \in G} \in \Delta$. If $(S(x), \sim)$ contains no loop, then x is a stable* (*resp.*, *strictly stable*) *equilibrium if and only if there exists $d \ge 2$ such that*:

(i) *$(S(x), \sim)$ is a complete d-partite subgraph, with partitions $=: V_1, \dots, V_d$,*
(ii) *$\sum_{i \in V_p} x_i = 1/d$ for all $p \in \{1, \dots, d\}$,*
(iii) *$\forall i \in \partial S(x)$, $N_i(x) \le$ (resp., $<$) $1 - 1/d$.*

If $(S(x), \sim)$ contains a loop, then x is a stable (*resp.*, *strictly stable*) *equilibrium if and only if $(S(x), \sim)$ is a clique of loops* [*resp.*, *with the additional assumption*: $\forall j \in \partial S(x)$, $N_j(x) < 1$ *or, equivalently*, $\partial\{j\} \not\supseteq S(x)$].

REMARK 1. Jordan [6] independently shows, in the context of preferential duplication graphs, that conditions (i)–(iii) in Theorem 2 are indeed sufficient for $x \in \Delta$ being a stable equilibrium when loops are not allowed.

REMARK 2. A connection between the number of stable rest points in the *replicator* dynamics [or of patterns of evolutionary stable sets (ESS's)] and the numbers of cliques of its graph was made by Vickers and Cannings [21, 22], Broom et al. [3] and Tyrer et al. [20], motivated by the study of evolutionary dynamics in biology.

A consequence of Theorem 1 is that supports of stable equilibria are *generically* cliques of the graph G. More precisely, assume that the coefficients $(a_{i,j})_{i,j \in G}$ are distributed according to some absolutely continuous distribution w.r.t. the Lebesgue measure on symmetric matrices. Then the supports of stable equilibria are a.s. cliques of the graph G (i.e., any two different vertices are connected), as a consequence of $(\mathrm{P})_{S(x)}$(a) and (c).

The following Theorem 3 states that, given any neighborhood $\mathcal{N}(x)$ of a strictly stable equilibrium $x \in \mathcal{E}_s$, then, with positive probability, the VRRW eventually localizes in

$$T(x) := S(x) \cup \partial S(x),$$

and the vector of density of occupation converges toward a point in $\mathcal{N}(x)$, which will not necessarily be x (there may exist a submanifold of stable equilibria in the neighborhood of x). Note that this will imply, using Remark 2, that the VRRW generically localizes with positive probability on subgraphs which consist of a clique plus its outer boundary.

More precisely, let us first introduce the following definitions. For all $R \subseteq G$, let

$$\mathcal{S}(R) := S^{-1}(R) \cap \Delta = \{x \in \Delta \text{ s.t. } S(x) = R\}.$$

For any open subset U of Δ containing $x \in \Delta$, let $\mathcal{L}(U)$ be the event

$$\mathcal{L}(U) := \Big\{ y(\infty) := \lim_{n\to\infty} y(n) \text{ exists (coordinatewise) and belongs to } \mathcal{E}_s \cap \mathcal{S}(S(x)) \cap U \Big\}.$$

Let $\mathcal{R}$ be the asymptotic range of the VRRW, that is,

$$\mathcal{R} := \{i \in G \text{ s.t. } Z_\infty(i) = \infty\}.$$

For any random variable x taking values in Δ, let

$$\mathcal{A}_\partial(x) := \left\{ \forall i \in \partial S(x), \frac{Z_n(i)}{n^{N_i(x)/H(x)}} \text{ converges to a (random) limit } \in (0, \infty) \right\}.$$

THEOREM 3. *Let $x \in \Delta$ be a strictly stable equilibrium. Then, for any open subset U of Δ containing x,*

$$\mathbb{P}\big(\{\mathcal{R} = T(x)\} \cap \mathcal{L}(U) \cap \mathcal{A}_\partial(y(\infty))\big) > 0.$$

Moreover, the rate of convergence is at least reciprocally polynomial, that is, by possibly restricting the neighborhood U of x, there exists $\nu := \mathsf{Cst}(x, a)$ such that, a.s. on $\mathcal{L}(U)$,

$$\lim_{n\to\infty} \big(y(n) - y(\infty)\big) n^\nu = 0.$$

Theorem 3 is proved in Section 2.3. It naturally leads to the following questions.

First, are all the trapping subsets always of the form $T(x)$ for some $x \in \mathcal{E}_s$? The answer is negative in general: let us consider, for instance, the graph $(\mathbb{Z}, \sim)$ of integers, to which we add a loop $0 \sim 0$ at site 0, with $a_{i,j} := \mathbb{1}_{i\sim j}$. Then $x := (\mathbb{1}_{\{i=0\}})_{i\in\mathbb{Z}}$ is a stable equilibrium, but is not strictly stable since $N_{-1}(x) = N_1(x) = 1 = H(x)$. However, Proposition 1 (proved in Appendix A.2) shows that $y(n)$ converges to x with positive probability, by combining an urn result from Athreya [1], Pemantle and Volkov [16] (Theorem 2.3) with martingale techniques from Tarrès [18] (Section 3.1).

PROPOSITION 1. *Let* $(G, \sim)$ *be the graph of integers defined above, and let* $a_{i,j} := \mathbb{1}_{i \sim j}$. *Then, with positive probability, the VRRW localizes on* $\{-2, -1, 0, 1, 2\}$, *and there exist random variables* $\alpha \in (0, 1)$, C *and* $C' > 0$ *such that*

$$\text{(i)} \quad \frac{Z_n(0)}{n} \longrightarrow_{n\to\infty} 1,$$

$$\text{(ii)} \quad \frac{(Z_n(-1), Z_n(1))}{n/\log n} \longrightarrow_{n\to\infty} (\alpha, 1-\alpha),$$

$$\text{(iii)} \quad \left(\frac{Z_n(-2)}{(\log n)^{\alpha}}, \frac{Z_n(2)}{(\log n)^{1-\alpha}}\right) \longrightarrow_{n\to\infty} (C, C').$$

We conjecture that, conditionally on a localization of the VRRW on a finite subset, its vector of density of occupation on the subset converges to a stable equilibrium x of (6), that the asymptotic range $\mathcal{R}$ is a subset of $S(x) \cup \partial S(x) \cup \partial(\partial S(x))$, and is equal to $T(x) = S(x) \cup \partial S(x)$ if $x \in \mathcal{E}_s$, which occurs generically on a (in the sense given in the paragraph after Remark 2).

A proof would require a deeper understanding of the dynamics of $(Z_.(i))_{i\in G}$ (see Lemma 4). Note that, on the integers $\mathbb{Z}$ with standard adjacency—unlike Proposition 1—and with $a_{i,j} = \mathbb{1}_{i\sim j}$, the result that the VRRW a.s. localizes on five sites [18] implies that only equilibria in $\mathcal{E}_s$ are reached with positive probability. More precisely, in this case there exist a.s. $k \in \mathbb{Z}$ and $x \in \Delta$ with $x_k = 1/2$, $x_{k-1} = \alpha/2$, $x_{k+1} = (1-\alpha)/2$, $\alpha \in (0, 1)$ (thus, $x \in \mathcal{E}_s$) such that $Z_n(i)/n \to x_i$ as $n \to \infty$ for all $i \in \mathbb{Z}$, $\mathcal{A}_\partial(x)$ holds and $\mathcal{R} = T(x)$; see [18]. Stable equilibria which are not in $\mathcal{E}_s$ correspond to cases $\alpha = 0$ or 1, which would lead to localization on six vertices if they were possible, similarly to Proposition 1. This result on $\mathbb{Z}$ can be related to the property that every neighborhood of any stable equilibrium x contains a strictly stable one.

Second, which subsets are of the form $T(x) = S(x) \cup \partial S(x)$ for some $x \in \mathcal{E}_s$? We know from Theorem 1 that subsets $S(x)$ satisfy $(\mathrm{P})_{S(x)}$ and thus always consist of a complete d-partite subgraph with possible loops and its outer boundary for some $d \geq 2$. But $(\mathrm{P})_{S(x)}$ is not sufficient, and the occurrence of such subsets also depends on the reinforcement matrix $a = (a_{i,j})_{i,j\in G}$. Even in the case $a = (a_{i,j})_{i,j\in G} = (\mathbb{1}_{i\sim j})_{i,j\in G}$ Theorem 2 provides explicit criteria for $x \in \mathcal{E}_s$, but the corresponding condition (iii) [when $(S(x), \sim)$ has no loops] is on x, thus not explicitly on the subgraph.

We introduce in the following Definition 1 the notion of *strongly trapping subsets*, which we prove in Theorem 4 to always be such subsets $T(x)$ for some $x \in \mathcal{E}_s$. As a consequence, by Theorem 3, the VRRW localizes on these subsets with positive probability. The result is thus a generalization to arbitrary reinforcement matrices of Theorem 1.1 by Volkov [23] when $a_{i,j} := \mathbb{1}_{\{i\sim j\}}$, in which case the assumptions of Definition 1 obviously reduce to (c) or (c)$'$.

DEFINITION 1. *A subset $T \subseteq G$ is called a strongly trapping subset of* $(G, \sim)$ *if* $T = S \cup \partial S$, *where*:

(a) $(i, j) \mapsto a_{i,j}$ *is constant on* $\{(i, j) \in S^2 \text{ s.t. } i \sim j\}$, *with common value* $=: a_S$,
(b) $\max_{i \in S, j \in \partial S} a_{i,j} \le a_S$, *and*

either

(c)(i) *S is a complete d-partite subgraph of G for some $d \ge 2$, with partitions* $V_1, \dots, V_d$,
(ii) $\forall j \in \partial S, \exists p \in \{1, \dots, d\}$ *and* $i \in S \setminus V_p$ *such that* $j \not\sim V_p \cup \{i\}$,

or

(c)′ *S is a clique of loops, and* $\forall j \in \partial S$, $\partial\{j\} \not\supseteq S$.

THEOREM 4. *Let T be a strongly trapping subset of* $(G, \sim)$; *then the VRRW has asymptotic range T with positive probability.*

More precisely, assume $T = S \cup \partial S$, where S satisfies conditions (a)–(c) *or* (c)′ *of Definition 1, and let us use the corresponding notation. Let*

$$\Sigma := \left\{ x \in \mathcal{S}(S) \text{ s.t. } \sum_{i \in V_q} x_i = 1/d \text{ for all } 1 \le q \le d \right\},$$
$$r_d := d/(d-1)$$

if $(S, \sim)$ *contains no loops, and* $\Sigma := \mathcal{S}(S)$, $r_d := 1$ *otherwise.*

Then, for any $x \in \Sigma$ and any neighborhood $\mathcal{N}(x)$ of x in Σ, there exist random variables $y \in \mathcal{N}(x)$ and $C_j > 0$, $j \in \partial S$ such that, with positive probability:

(i) *VRRW eventually localizes on T, that is,* $\mathcal{R} = T$,
(ii) $Z_n(i)/n \longrightarrow_{n\to\infty} y_i$ *for all* $i \in S$,
(iii) $Z_n(j) \sim_{n\to\infty} C_j n^{r_d \sum_{i \sim j} a_{i,j} y_i / a_S}$ *for all* $j \in \partial S$.

Theorem 4 is proved in Section 2.2.3. We provide in Example 1 (illustrated in Figure 1) a counterexample showing that Theorem 3 is stronger, even in the case $a = (\mathbb{1}_{i\sim j})_{i,j\in G}$.

Third, which conditions on the graph and on the reinforcement matrix a do ensure the existence of at least one strictly stable equilibrium $x \in \mathcal{E}_s$, thus implying localization with positive probability on $T(x)$? First note that, trivially, this does not always occur, for instance, on $\mathbb{Z}$ when $\phi(n) := a_{\{n,n+1\}}$ is strictly monotone, in which case we believe the walk to be transient.

In the case $a = (\mathbb{1}_{i\sim j})_{i,j\in G}$, Volkov [23] proposed the following result, using an iterative construction on subsets of the graph.

PROPOSITION 2 (Volkov [23]). *Assume that $a = (\mathbb{1}_{i\sim j})_{i,j\in G}$, and that $(G, \sim)$ does not contain loops. Then, under either of the following conditions, there exists at least one strongly trapping subset*:

(A) $(G, \sim)$ *does not contain triangles*;
(B) $(G, \sim)$ *is of bounded degree*;
(C) *the size of any complete subgraph is uniformly bounded by some number K.*

PROOF. Start, for some $d \geq 2$, with any complete d-partite subgraph $(S, \sim)$ of G with partitions $V_1, \dots, V_d$ (e.g., a pair of connected vertices, $d = 2$). Let $x \in \partial S$, $S = V_1 \cup \dots \cup V_d$:

(1) First assume that $x \sim V_p$ for all $1 \leq p \leq d$. Then, for all $1 \leq p \leq d$, let $j_p \in V_p$ be such that $x \sim j_p$; iterate the procedure with the subgraph $\bigcup_{1\leq p\leq d}\{j_p\} \cup \{x\}$, which is a clique, and thus a complete $(d + 1)$-partite subgraph.
(2) Now assume there exists p such that $x \not\sim V_p$, with $\partial\{x\} \supseteq S \setminus V_p$. Then we iterate the procedure with the complete d-partite subgraph $S \cup \{x\}$ with partitions $V_1, \dots, V_p \cup \{x\}, \dots, V_d$.
(3) Otherwise we keep the same subgraph S and try another $x \in \partial S$.

The construction eventually stops if (A), (B) or (C) holds. When it does, that is, when S has remained unchanged for all $x \in \partial S$, then $T = S \cup \partial S$ is a strongly trapping subgraph in the sense of Definition 1. $\square$

Using a similar technique, we can obtain the following necessary condition for the existence of a strongly trapping subset in the case of general reinforcement matrices a, when the graph does not contain triangles or loops. Let us first introduce some notation. Let c be the distance on $E(G)$ edges of G defined as follows: for all $e, e' \in E(G)$, let $c(e, e')$ be the minimum number of edges necessary to connect e to e' plus one (0 if $e = e'$, and 1 if $e \sim e'$). For all $e = \{i, j\}$, let $\mathcal{C}(2, e)$ be the set of maximal complete 2-partite subgraphs $S \subseteq G$ such that $i, j \in S$ and, for all $k, l \in S$ with $k \sim l$, $a_{k,l} = a_{i,j}$.

PROPOSITION 3. *Assume the graph does not contain triangles nor loops. If, for some $e \in E(G)$,*

$$\min_{S\in\mathcal{C}(2,e)} \max_{k\in S, l\in\partial S} a_{k,l} \leq a_e, \tag{9}$$

then there exists at least one strongly trapping subset.

Note that (9) *holds if*

$$\max_{c(e,e')\leq 2} a_{e'} \leq a_e.$$

REMARK 3. If, for all $e \in E(G)$, (9) does not hold, then there exists, for all $e \in E(G)$, an infinite sequence of edges $(e_n)_{n\in\mathbb{N}_0}$ such that $e_0 = e$, $e_n \sim e_{n+1}$ and, for all $n \in \mathbb{N}$, $a_{e_n} \leq a_{e_{n+1}}$ and $a_{e_n} < a_{e_{n+2}}$. However, even in this case, there can exist a strictly stable equilibrium $x \in \mathcal{E}_s$ (but no strongly trapping subset).

PROOF OF PROPOSITION 3. By assumption, there exist $e = \{i, j\}$ and a maximal complete 2-partite subgraph $S \subseteq G$ containing i and j, with partitions V_1 and V_2, and satisfying conditions (a), (b) and (c)(i) of Definition 1. For all $k \in \partial S$, k is adjacent to at most one of two partitions, say, V_1, since otherwise G would contain a triangle; if k were adjacent to all vertices in V_1, then it would be in V_2, since S is assumed maximal. Hence, (c)(ii) holds as well, and S is a strongly trapping subset. $\square$

When the graph contains triangles, the property outlined in Remark 3, that is, the existence of an infinite sequence of edges with increasing labels when there is no strongly trapping subset, does not hold anymore. The maximum of the Lyapounov function on a complete subgraph with more than two vertices takes a nontrivial form, which can lead to counterintuitive behavior.

We show, for instance, in Example 2 a case where the reinforcement matrix a has a strict global maximum at a certain edge, but where, however, there is no stable equilibrium at all. We believe the walk to be transient in this example.

EXAMPLE 1. Let us show, in the case $a = (\mathbb{1}_{i \sim j})_{i,j \in G}$, that Theorem 3 is stronger than Theorem 4. Consider a graph G on six vertices A, B, C, D, E and F, with a neighborhood relation $\sim$ defined as follows (see Figure 1): $A \sim B \sim C \sim D \sim A$, $C \sim E \sim D$ and $E \sim F$ (recall that the graph G is symmetric). Let $x = (x_A, x_B, x_C, x_D, x_E, x_F) := (3/8, 3/8, 1/8, 1/8, 0, 0)$, then $S(x) = \{A, B, C, D\}$ and $\partial S(x) = \{E\}$. Also, x is an equilibrium of (6), $(\mathrm{P})_{S(x)}$ is satisfied with $V_1 = \{A, C\}$, $V_2 = \{B, D\}$, and $N_E(x) = 1/4 < H(x) = 1/2$, which implies that x is a

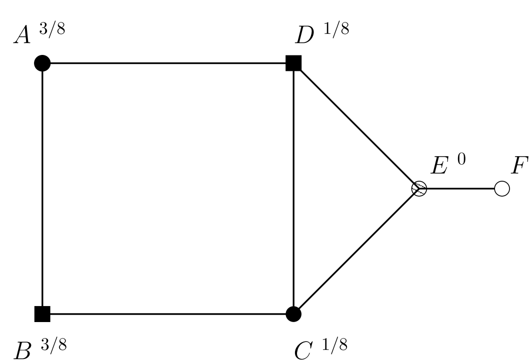

FIG. 1. *We show in Example 1 that $T := \{A, B, C, D, E\}$ does not satisfy the assumptions of Theorem 4, but is a trapping subgraph with positive probability by Theorems 2 and 3. The numbers indicated in superscript of vertices represent the limit proportions of visits to these vertices if $x(n)$ were to converge to the equilibrium x in the example. In this case the walk would asymptotically spend most of the time in the bipartite subgraph $S := V_1 \cup V_2$, where $V_1 := \{A, C\}$, $V_2 := \{B, D\}$, evenly divided between partitions V_1 and V_2, and vertex E would be seldom visited, of the order of $\sqrt{n}$ times at time n.*

strictly stable equilibrium by Theorem 2, hence subsequently by Theorem 3 that $\mathcal{R} = T(x)$ with positive probability.

Now let us prove by contradiction that $T(x)$ with such x does not satisfy the assumptions of Theorem 4 above. Indeed, if $T(x) = S \cup \partial S$, then $S \subseteq \{A, B, C, D\}$ since, otherwise, F would belong to $T(x)$. Now the condition that, for all $i \in \partial S$, $\exists p \in \{1, \dots, d\}$ and $j \in S \setminus V_p$ such that $i \not\sim V_p \cup \{j\}$ implies, in particular, that a vertex in ∂S is not connected to at least two other vertices in S, so that $i \in \partial S$ cannot be A, B, C or D, which are connected to all other but one vertex in $\{A, B, C, D\}$. Hence, $S = \{A, B, C, D\}$, but then $i := E$ is connected to both partitions of S, and does not satisfy the condition mentioned in the last sentence, bringing a contradiction.

EXAMPLE 2. Let us first study the case of a triangle $(G, \sim)$, $G := \{0, 1, 2\}$, $0 \sim 1 \sim 2 \sim 0$, with reinforcement coefficients $a := a_{0,1}$, $b := a_{1,2}$, $c := a_{0,2} > 0$.

If $a < b + c$, then the equilibrium $x = (x_0, x_1, x_2) = (1/2, 1/2, 0)$ is not stable, since $N_2(x) = (b + c)/2 > H(x) = a/2$. Hence, if we assume that

$$a < b + c, \qquad b < a + c, \qquad c < a + b, \tag{10}$$

then a stable equilibrium has to belong to the interior of the simplex Δ. A simple calculation shows that there is only one such equilibrium:

$$x = (x_0, x_1, x_2) := \left(\frac{c(a + b - c)}{\delta}, \frac{b(a + c - b)}{\delta}, \frac{a(b + c - a)}{\delta}\right),$$

where

$$\delta := (a + b + c)^2 - 2(a^2 + b^2 + c^2);$$

$\delta > 0$, which can be shown by adding up inequalities $(b - a)^2 \le c^2$, $(c - a)^2 \le b^2$ and $(c - b)^2 \le a^2$. Then $H(x) = 2abc/\delta$.

Let $\mathbb{N} := \mathbb{Z}_+$. Let us now consider the following graph $(G, \sim)$ with vertices $G := \{\underline{i}, \overline{i}, i \in \mathbb{N}\}$ and adjacency $\underline{i} \sim \underline{i+1}$, $\overline{i} \sim \overline{i+1}$, $\underline{i} \sim \overline{i}$ and $\overline{i} \sim \underline{i+1}$, for all $i \in \mathbb{N}$, as drawn in Figure 2.

Fix $\varepsilon, \eta, p, q > 0$, $\mu \in (0, 1)$, which will be chosen later. Let, for all $n \in \mathbb{N}$,

$$p_n := p \prod_{k=0}^{n-1} (1 - \mu^k \varepsilon), \qquad q_n := q \prod_{k=0}^{n-1} (1 + \mu^k \eta). \tag{11}$$

Note that, for all $n \in \mathbb{N}$,

$$p\left(1 - \frac{\varepsilon}{1 - \mu}\right) \le p_n \le p, \qquad q \le q_n \le q e^{\eta/(1-\mu)}.$$

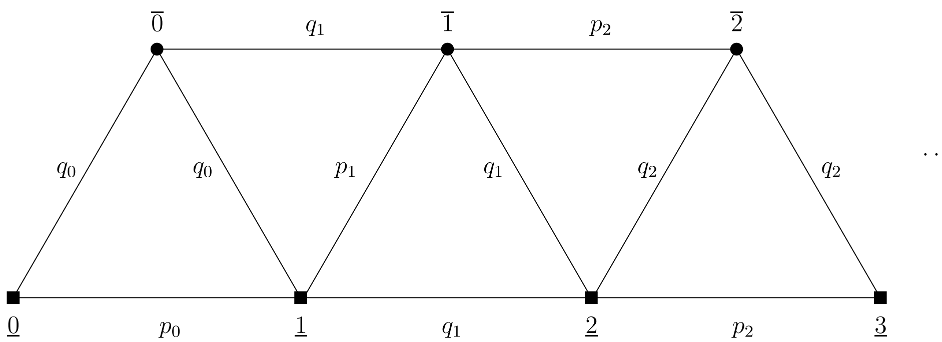

FIG. 2. *On the infinite graph on the figure, with reinforcement coefficient sequences* $(p_n)_{n\ge 0}$ *strictly decreasing and* $(q_n)_{n\ge 0}$ *strictly increasing, we show in Example* 2 *that, even if* $p_0 = \sup_{n\ge 0} p_n > \sup_{n\ge 0} q_n$, *we can choose these sequences in such a way that there is no stable equilibrium in* Δ, *and therefore no trapping subgraph.*

Now assume that the reinforcement matrix $(a_{k,l})_{k,l\in G}$ is defined as follows, depending on $(p_n)_{n\in\mathbb{N}}$ and $(q_n)_{n\in\mathbb{N}}$, for all $i \in \mathbb{N}$:

$$\begin{aligned}
a_{\underline{2i},\underline{2i+1}} &:= p_{2i}, \qquad a_{\overline{2i+1},\overline{2(i+1)}} := p_{2(i+1)},\\
a_{\overline{2i},\overline{2i+1}} &= a_{\underline{2i+1},\underline{2(i+1)}} := q_{2i+1}\\
a_{\underline{2i},\overline{2i}} &= a_{\overline{2i},\underline{2i+1}} := q_{2i}\\
a_{\underline{2i+1},\overline{2i+1}} &:= p_{2i+1}\\
a_{\overline{2i+1},\underline{2(i+1)}} &:= q_{2i+1}.
\end{aligned}$$

Let $x \in \Delta$ be a stable equilibrium of (6). Then, by Theorem 1, $(\mathrm{P})_{S(x)}$ holds, so that $S(x)$ consists of two vertices or a triangle [it cannot be made of four vertices, because of $(\mathrm{P})_{S(x)}$(c)]. Assume

$$p < 2q, \qquad \eta q e^{\eta/(1-\mu)} < p\left(1 - \frac{\varepsilon}{1-\mu}\right). \tag{12}$$

Then, for all $i \in \mathbb{N}$,

$$p_i < 2q_i, \qquad p_{i+1} < q_i + q_{i+1}, \qquad q_{i+1} < q_i + p_{i+1},$$

so that $S(x)$ has to be a triangle.

Assume $S(x) := \{\underline{2i}, \overline{2i}, \underline{2i+1}\}$ for some $i \in \mathbb{N}$; the argument is similar in other cases. Then

$$x_{\overline{2i+1}} = \frac{H(x)}{2q_i}, \qquad x_{\overline{2i}} = \frac{H(x)}{2q_i^2}(2q_i - p_i), \qquad x_{\underline{2i+1}} = \frac{H(x)}{2q_i},$$

and

$$N_{\underline{2i}}(x) = q_i x_{\overline{2i}} + p_i x_{\underline{2i+1}} = H(x),$$

and, therefore,

$$\begin{aligned}N_{\overline{2i+1}}(x) &= q_{i+1}x_{\overline{2i}} + p_{i+1}x_{\underline{2i+1}}\\ &= H(x) + \frac{H(x)}{2q_i^2}[(q_{i+1}-q_i)(2q_i-p_i)+(p_{i+1}-p_i)q_i]\\ &= H(x) + \frac{H(x)}{2q_i^2}\mu^i[\eta q_i(2q_i-p_i)-\varepsilon p_i q_i] > H(x)\end{aligned}$$

if

$$\eta > \varepsilon\frac{p}{2q-p}, \tag{13}$$

using that $p/(2q-p) > p_i/(2q_i-p_i)$ for all $i \in \mathbb{N}$.

Hence, x is not a stable equilibrium, which leads to a contradiction.

2. Introduction to the proofs.

2.1. *Notation.* We let $\mathbb{N} := \mathbb{Z}_+$, $\mathbb{N}^* := \mathbb{N} \setminus \{0\}$, $\mathbb{R}_+^* := \mathbb{R}_+ \setminus \{0\}$.

For all $y = (y_i)_{i\in G} \in \mathbb{R}^G$ and for any finite subset A of G, let

$$y_A := \sum_{i\in A} y_i.$$

Given $r \in \mathbb{N}^*$, let $(\cdot,\cdot)$ (resp., $|\cdot|$, $\|\cdot\|_\infty$) be the scalar product (resp., the canonical norm, the infinity norm) on $\mathbb{R}^r$, defined by

$$(a,b) = \sum_{i=1}^r a_i b_i, \qquad |a| = \sqrt{(a,a)}, \qquad \|a\|_\infty := \max_{1\le i\le r} |a_i|$$

if $a = (a_1,\ldots,a_r)$ and $b = (b_1,\ldots,b_r)$.

Given a real $r\times r$ matrix M with real eigenvalues, we let $\mathsf{Sp}(M)$ denote the set of eigenvalues of M. When M is symmetric we let $M[\cdot]$ denote the quadratic form associated to M, defined by $M[a] = (Ma, a)$ for all $a \in \mathbb{R}^r$.

Given $y_1,\ldots,y_r$, we let $\mathsf{Diag}(y_1,\ldots,y_r)$ be the diagonal $r\times r$ matrix of diagonal terms $y_1,\ldots,y_r$.

For all $u, v \in \mathbb{R}$, we write $u = \square(v)$ if $|u| \le v$. Given two (random) sequences $(u_n)_{n\ge k}$ and $(v_n)_{n\ge k}$ taking values in $\mathbb{R}$, we write $u_n \equiv v_n$ if $u_n - v_n$ converges a.s., and $u_n \sim_{n\to\infty} v_n$ iff $u_n/v_n \to_{n\to\infty} 1$, with the convention that $0/0 = 1$.

Let $\mathsf{Cst}(a_1, a_2, \ldots, a_p)$ denote a positive constant depending only on $a_1, a_2, \ldots, a_p$, and let Cst denote a universal positive constant.

2.2. *Proof of Theorems* 1, 2 *and* 4. Theorems 1 and 2 are a consequence of the more general three following Lemmas 1, 2 and 3 below.

2.2.1. *Lemmas 1, 2 and 3, and proof of Theorem 1.* By the following Lemma 1, if an equilibrium $x \in \Delta$ is stable, then the eigenvalues of $[a_{i,j} - 2H(x)]_{i,j \in S(x)}$, which depend only on a, $S(x)$ and $H(x)$, are nonpositive. This property will subsequently imply $(\mathrm{P})_{S(x)}$, by Lemmas 2 and 3.

LEMMA 1. *Let $x = (x_i)_{i \in G} \in \Delta$ be an equilibrium. Then*:

(a) *$DF(x)$ has real eigenvalues.*
(b) *The three following assertions are equivalent*:
  (i) *x is stable*,
  (ii) $\max \mathsf{Sp}(DF(x)) \leq 0$,
  (iii) $\max(\mathsf{Sp}([a_{i,j} - 2H(x)]_{i,j \in S(x)}) \cup \{N_i(x) - H(x), i \in \partial S(x)\}) \leq 0$.
(c) *If x is stable, then it is feasible.*

Lemma 2 yields an algebraically simpler characterization of assertion $(\mathrm{P})_S$ for $S \subseteq G$; recall that, given subsets S and R of G, $\partial_S R$, defined in Section 1, is the outer boundary of R inside S.

LEMMA 2. *The statement* $(\mathrm{P})_S$ *is equivalent to*

$(\mathrm{P})'_S$ *If $j, k \in S$ are such that $j \not\sim k$, then, for all $i \in S$, $a_{i,j} = a_{i,k}$* (*so that $\partial_S\{j\} = \partial_S\{k\}$ in particular*).

Lemma 3 states that $(\mathrm{P})_{S(x)}$ holds if the eigenvalues of $[a_{i,j} - 2H(x)]_{i,j \in S(x)}$ are nonpositive, with equivalence if $a = (\mathbb{1}_{i \sim j})_{i,j \in G}$.

LEMMA 3. *Let $x = (x_i)_{i \in G} \in \Delta$ be a feasible equilibrium. Then*

$$\max \mathsf{Sp}\big([a_{i,j} - 2H(x)]_{i,j \in S(x)}\big) \leq 0 \quad \Longrightarrow \quad (\mathrm{P})'_{S(x)}.$$

If, for some $c > 0$, $a_{i,j} = c\mathbb{1}_{i \sim j}$ for all $i, j \in S(x)$, then the above implication is an equivalence.

Lemmas 1, 2 and 3 are proved, respectively, in Sections 3.1, 3.2 and 3.3. They obviously imply Theorem 1.

2.2.2. *Proof of Theorem 2.* Suppose $a = (\mathbb{1}_{i \sim j})_{i,j \in G}$, and let $x \in \Delta$.

First assume that $(S(x), \sim)$ contains no loop. If x is a stable equilibrium, then $(\mathrm{P})_{S(x)}$ and, thus, (i) holds by Theorem 1; let V_k, $1 \leq k \leq d$ be the partitions of $S(x)$. Then $d \geq 2$ [otherwise $H(x) = 0$ and x is not feasible, thus not stable by Lemma 1] and, for all $1 \leq k \leq d$, $j \in V_k$,

$$v_k := \sum_{i \in V_k} x_i = 1 - N_j(x) = 1 - H(x),$$

so that $v_k = 1/d$ (since $\sum_k v_k = 1$) and $H(x) = 1 - 1/d$, and, subsequently, (ii)–(iii) hold by Lemma 1. Conversely, assume (i)–(iii) hold; then $N_i(x) = 1 - 1/d$ for all $i \in S(x)$, so that $H(x) = \sum_{i \in S(x)} x_i N_i(x) = 1 - 1/d$ and x is a feasible equilibrium. Now (i) implies $(\mathrm{P})_{S(x)}$ and thus $(\mathrm{P})'_{S(x)}$ by Lemma 2. Hence, using Lemmas 1 and 3, x is a stable equilibrium.

Now assume on the contrary that $(S(x), \sim)$ contains one loop $i \sim i$. If x is a stable equilibrium, then $(\mathrm{P})_{S(x)}$ again holds by Theorem 1: $(\mathrm{P})_{S(x)}$(b) implies $N_i(x) = 1 = H(x)$ (x equilibrium). Hence, for all $j \in S(x)$, $N_j(x) = 1$ and $j \sim k$ for all $k \in S(x)$, that is, $(S(x), \sim)$ is a clique of loops. Conversely, if $(S(x), \sim)$ is a clique of loops, then $(\mathrm{P})_{S(x)}$ obviously holds so that, by Lemmas 1 and 3, x is stable [since $H(x) = 1$, then $N_i(x) \le H(x)$ for all $i \in G$].

2.2.3. *Proof of Theorem* 4. First observe that

$$\Sigma = \mathcal{S}(S) \cap \mathcal{E}_s.$$

Indeed, the proof of Theorem 2 implies that $\Sigma \supseteq \mathcal{S}(S) \cap \mathcal{E}_s$ and, conversely, that if $x \in \Sigma$, then x is a equilibrium and, by (c)(ii), for all $j \in \partial S(x)$, $N_j(x) < H(x)$ [$= a_S(1-1/d)$ if $(S(x), \sim)$ contains no loops, $= a_S$ otherwise], using assumptions (a)–(b) and (c)(ii) or the second part of (c)$'$. Also, $(\mathrm{P})_{S(x)}$ holds by (c) or (c)$'$, and, therefore, x is strictly stable by Lemmas 1–3. The rest of the proof follows from Theorem 3.

2.3. *Proof of Theorem* 3. First, we provide in Lemma 4 a rigorous mathematical setting for the stochastic approximation of the density of occupation of the VRRW $x(n)$ by solutions of the ordinary differential equation (6) on a finite graph G, heuristically justified in Section 1 [see (4)]. Second, we make use of this technique and of an entropy function originally introduced in [9] to study the VRRW on the finite subgraph $T(x)$ when its density of occupation is in the neighborhood of a strictly stable equilibrium x, in Lemmas 5–10. Third, we focus again on a general graph G—possibly infinite—and prove in Proposition 4, assuming again that the density of occupation is in the neighborhood of an element $x \in \mathcal{E}_s$, that the walk eventually localizes in $T(x)$ with lower bounded probability.

In the first step, we make use of a technique originally introduced by Métivier and Priouret in 1987 [11] and adapted by Benaïm [2] in the context of vertex reinforcement when the graph is complete (Hypothesis 3.1 in [2]). In Sections 4.1–4.3, we generalize it and show that a certain quantity $z(n)$, depending only on a, $x(n)$, X_n and n and defined in (36), satisfies the recursion (37):

$$z(n+1) = z(n) + \frac{1}{n + n_0 + 1} \frac{F(z(n))}{H(x(n))} + \varepsilon_{n+1} + r_{n+1},$$

where $\mathbb{E}(\varepsilon_{n+1} | \mathcal{F}_n) = 0$. The following Lemma 4, proved in Section 4.3, provides upper bounds on the infinity norms of ε_{n+1}, r_{n+1} and $z(n) - x(n)$, and on the conditional variances of $(\varepsilon_{n+1})_i$, $i \in G$.

More precisely, let us break down the set of vertices of G as $G = S \cup \partial S$, where $(S, \sim)$ is finite, connected and not a singleton unless it is a loop. Let, for all $\alpha \in \mathbb{R}_+ \setminus \{0\}$,

$$\Lambda_\alpha := \{x = (x_j)_{j \in G} \in \Delta \text{ s.t. } x_j \geq \alpha \text{ for all } j \in S\}. \tag{14}$$

LEMMA 4. *For all $n \geq \mathsf{Cst}(\alpha)$ and $i \in G$, if $x(n) \in \Lambda_\alpha$, then*

(a) $\|\varepsilon_{n+1}\|_\infty \leq \dfrac{\mathsf{Cst}(\alpha, a, |G|)}{n + n_0}$, (b) $\mathbb{E}((\varepsilon_{n+1})_i^2 | \mathcal{F}_n) \leq \dfrac{\mathsf{Cst}(\alpha, a, |G|) x(n)_i}{(n + n_0)^2}$,

(c) $\|r_{n+1}\|_\infty \leq \dfrac{\mathsf{Cst}(\alpha, a, |G|)}{(n + n_0)^2}$, (d) $\|z(n) - x(n)\|_\infty \leq \dfrac{\mathsf{Cst}(\alpha, a, |G|)}{n + n_0}$.

Note that if G were a complete d-partite finite graph for some $d \geq 1$ or, more generally, if G were without loop and, for all $i, j \in G$ with $i \sim j$, $\{i, j\} \cup \partial\{i, j\} = G$, then the constants in the inequalities of Lemma 4 would not depend on $\alpha > 0$ and, as a consequence, the stochastic approximation of $z(n)$ by (6) would hold uniformly a.s. Indeed, for all $n \in \mathbb{N}$, by the pigeonhole principle, there exists at least one edge $\{i, j\}$ $i, j \in G$, $i \sim j$, on which the walk has spent more than $n/|G|^2$ times, so that $x(n)_i \wedge x(n)_j \geq \frac{1}{|G|^2} \frac{n}{n+n_0}$ and, under the assumption on G, Lemma 4 with $S := \{i, j\}$ would yield the claim.

In the second step, we define an entropy function $V_q(\cdot)$, measuring a "distance" between q and an arbitrary point [as can be seen by (15) below], originally introduced by Losert and Akin in 1983 in [9] in the study of the deterministic Fisher–Wright–Haldane population genetics model, and to our knowledge so far only used for the analysis of deterministic replicator dynamics. Note that it is not mathematically a distance, however, since it does not satisfy the triangle inequality in general.

In the following, until after the statement Lemma 10—and, in particular, in Lemmas 5–10—we assume that $x \in \mathcal{E}_s$ and $G = T(x) = S(x) \cup \partial S(x)$; this choice will be justified later in the proof. Note that if $q \in \mathcal{N}(x) \cap \mathcal{E}_s$, where $\mathcal{N}(x)$ is an adequately chosen neighborhood of x, then $q \in \mathcal{S}(S(x))$ since $x \in \mathcal{E}_s$, so that $T(q) = T(x)$. Set $S := S(x)$, $T := T(x)$, and $\mathcal{S} := \mathcal{S}(S(x))$ for simplicity.

Lemmas 5 and 6 below will imply that, given any stable equilibrium $q \in \mathcal{N}(x) \cap \mathcal{E}_s$ as a reference point, $V_q(z(n))$ decreases in average when $z(n)$ is close enough to x. Therefore, martingale estimates will enable us to prove in Lemma 7 that, starting in the neighborhood of x, $x(n)$ remains close to x with large probability if n is large, and converges to one of the strictly stable equilibria in this neighborhood.

For all $q = (q_i)_{i \in G} \in \mathcal{S}$ and $y \in \mathbb{R}^G$, let

$$V_q(y) := \begin{cases} -\sum_{i \in S} q_i \log(y_i/q_i) + 2y_{\partial S}, & \text{if } y_i > 0, \forall i \in S, \\ \infty, & \text{otherwise.} \end{cases}$$

Let, for all $q \in \mathcal{S}$ and $r > 0$,

$$B_{V_q}(r) := \{y \in \Delta \text{ s.t. } V_q(y) < r\}, \qquad B_\infty(q, r) := \{y \in \Delta \text{ s.t. } \|y - q\|_\infty < r\}.$$

Then, we will prove in Section 4.4 that, for all $q \in \mathcal{S}$, there exist increasing continuous functions $u_{1,q}, u_{2,q} : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ such that $u_{1,q}(0) = u_{2,q}(0) = 0$ and, for all $r > 0$,

$$B_\infty(q, u_{1,q}(r)) \subseteq B_{V_q}(r) \subseteq B_\infty(q, u_{2,q}(r)). \tag{15}$$

Let, for all $q, z \in \mathbb{R}^G$,

$$I_q(z) := -\sum_{i \in S} q_i [N_i(z) - H(z)] + 2 \sum_{i \in \partial S} z_i [N_i(z) - H(z)]. \tag{16}$$

The following Lemma 5, also proved in Section 4.4, provides the stochastic approximation equation for $V_q(z(n))$, $q \in \mathcal{S} \cap \mathcal{E}_s$; we make use of notation $u = \square(v) \iff |u| \le v$, introduced in Section 2.1.

LEMMA 5. *Let $q \in \mathcal{S} \cap \mathcal{E}_s$. There exist an adapted process $(\zeta_n)_{n \in \mathbb{N}}$ (not depending on q and a), and constants n_1 and ε (depending only on q and a) such that, if $n \ge n_1$ and $x(n) \in B_{V_q}(\varepsilon)$, then $V_q(z(n)) < \infty$, $V_q(z(n+1)) < \infty$, $\mathbb{E}(\zeta_{n+1} | \mathcal{F}_n) = 0$ and*

$$\begin{aligned} V_q\bigl(z(n+1)\bigr) = V_q(z(n)) &+ \frac{I_q(x(n))}{(n + n_0 + 1) H(x(n))} \\ &- (q, \zeta_{n+1}) + 2(\varepsilon_{n+1})_{\partial S} + \square\left(\frac{\mathsf{Cst}(q, a)}{(n + n_0)^2}\right). \end{aligned} \tag{17}$$

Lemma 6, proved in Section 3.4, provides estimates of the Lyapounov function H, and of $I_\cdot(\cdot)$, in the neighborhood of a strictly stable equilibrium. It will not only be useful in the proof of Lemma 7, stating convergence of $x(n)$ with large probability, but also for Lemma 8 on the rate of this convergence.

LEMMA 6. *There exists a neighborhood $\mathcal{N}(x)$ of x in Δ such that, for all $q \in \mathcal{N}(x) \cap \mathcal{E}_s$, $y \in \mathcal{N}(x)$,*

$$\text{(a)} \quad \mathsf{Cst}(x, a) J(y) \le H(q) - H(y) \le \mathsf{Cst}(x, a) J(y), \tag{18}$$

$$\begin{aligned} \text{(b)} \quad & -[H(q) - H(y) + \mathsf{Cst}(x, a) y_{\partial S}] \\ & \le I_q(y) \le -[H(q) - H(y) + \mathsf{Cst}(x, a) y_{\partial S}] \le 0. \end{aligned} \tag{19}$$

REMARK 4. Lemma 6 implies that $y \in \mathcal{N}(x)$ is an equilibrium iff $H(y) = H(x)$. Also note that the maximality of H at $x \in \mathcal{E}_s$ is not global in general. For instance, in the counterexample at the end of Section 1, $x := (3/8, 3/8, 1/8, 1/8, 0) \in \mathcal{E}_s$, but, letting $y := (0, 0, 1/3, 1/3, 1/3)$, $H(y) = 2/3 > H(x) = 1/2$.

The proof of Lemma 7 is shown in Section 5.1. A key point in its proof is that the martingale term $-(q, \zeta_{n+1}) + 2(\varepsilon_{n+1})_{\partial S}$, in Lemma 5, is a linear function of ζ_{n+1} and ε_{n+1} which do not depend on q, so that the two corresponding convergence results of these martingales will apply from any reference point $q \in \mathcal{E}_s \cap \mathcal{N}(x)$. It will enable us to prove that, if r is a accumulation point of $x(n)$, then $V_r(x(n))$ a.s. converges to 0 if $r \in \mathcal{N}(x)$ although r is random.

LEMMA 7. *There exist $\varepsilon_0 := \mathsf{Cst}(x, a)$ and $n_1 := \mathsf{Cst}(x, a)$ such that, if for some $\varepsilon \le \varepsilon_0$ and $n \ge n_1$, $x(n) \in B_{V_x}(\varepsilon/2)$, then*

$$\mathbb{P}\big(\mathcal{L}(B_{V_x}(\varepsilon))|\mathcal{F}_n\big) \ge 1 - \exp\big(-\varepsilon^2 \mathsf{Cst}(x, a)(n + n_0)\big).$$

Next, we provide in Lemma 8 some information on the rate of convergence of $x(n)$ to $x(\infty)$, which will be necessary for the asymptotic estimates on the frontier $\mathcal{A}_\partial(x(\infty))$ in Lemma 10.

LEMMA 8. *There exist ε, $\nu := \mathsf{Cst}(x, a)$ such that, a.s. on $\mathcal{L}(B_{V_x}(\varepsilon))$,*

$$\lim_{n\to\infty} \big(x(n) - x(\infty)\big)n^\nu = 0.$$

The proof of Lemma 8, given in Section 5.2, starts with a preliminary estimate of the rate of convergence of $H(x(n))$ to $H(x(\infty))$. To this end, we make use of Lemma 9 below, giving the stochastic approximation equation of $H(z(n))$. It implies, together with Lemma 6(a), that the expected value of $H(z(n+1)) - H(z(n))$ is at least $\mathsf{Cst}(x, a)(H(x) - H(z(n))$, so that we can then estimate the rate of $H(x(n))$ to $H(x)$ by a one-dimensional technique.

Finally, this estimate implies similar ones for the convergence of $J(x(n))$ and $I_{x(\infty)}(x(n))$ to 0 by Lemma 6, so that we conclude using entropy estimates for the rate of convergence of $V_{x(\infty)}(z(n))$, using again that only two martingales estimates are necessary, given the linearity of the perturbation in (17) with respect to the reference point $q \in \mathcal{E}_s \cap \mathcal{N}(x)$.

LEMMA 9. *For all $n \in \mathbb{N}$,*

$$H\big(z(n+1)\big) - H(z(n)) = \frac{1}{n + n_0 + 1} \frac{J(z(n))}{H(x(n))} + \xi_{n+1} + s_{n+1}, \tag{20}$$

where $\mathbb{E}(\xi_{n+1}|\mathcal{F}_n) = 0$ and, if for some $\alpha > 0$, $x(n) \in \Lambda_\alpha$ and $n \ge \mathsf{Cst}(\alpha)$, then

$$(1) \quad \|\xi_{n+1}\|_\infty \le \frac{\mathsf{Cst}(\alpha, a, |G|)}{n + n_0}, \qquad (2) \quad \|s_{n+1}\|_\infty \le \frac{\mathsf{Cst}(\alpha, a, |G|)}{(n + n_0)^2}.$$

Lemma 9 is proved in Section 4.5.

Lemma 10 yields the asymptotic behavior on the border sites ∂S. This behavior is similar to the one one would obtain without perturbation [i.e., with $(\varepsilon_n)_{n\in\mathbb{N}^*} = 0$

in (37)]. Indeed, if $i \in \partial S$, then $N_i(x) - H(x) < 0$ is the eigenvalue of the Jacobian matrix of (6) in the direction $(\delta_{i,j})_{j\in G}$ (see the proof of Lemma 1), and the renormalization in time is approximately in $H(x)^{-1}\log n$ [see equation (37)], so that the replicator equation (6) would predict that $i \in \partial S$ is visited of the order of $n^{N_i(x)/H(x)}$ times at time n. This similarity with the noiseless case is due to the fact that the perturbation $(\varepsilon_n)_{n\in\mathbb{N}^*}$ is weak near the boundary [see Lemma 4(b)].

LEMMA 10. *There exists* $\varepsilon := \mathsf{Cst}(x, a)$ *such that, a.s. on* $\mathcal{L}(B_{V_x}(\varepsilon))$, $\mathcal{A}_\partial(x(\infty))$ *occurs a.s.*

The proof of Lemma 10, given in Section 5.3, makes use of a martingale technique developed in [18], Section 3.1, and in [7] in the context of strong edge reinforcement. We could have shown Lemma 10 by a thorough study of the border sites coordinates of the stochastic approximation equation (37), but it would lead to a significantly longer—and less intuitive—proof.

Now we do not assume anymore that $G = T(x)$ for some $x \in \Delta$, in other words, we let the graph $(G, \sim)$ be arbitrary, possibly infinite.

Let, for all $n, k \in \mathbb{N} \cup \{\infty\}$, $n \geq k$, $\mathcal{R}_{n,k}$ be the range of the vertex-reinforced random walk between times n and k, that is,

$$\mathcal{R}_{n,k} := \{i \in G \text{ s.t. } X_j = i \text{ for some } j \in [n, k]\};$$

note that, for all $n \in \mathbb{N}$, $\mathcal{R} \subseteq \mathcal{R}_{n,\infty}$.

PROPOSITION 4. *Let* $x \in \mathcal{E}_s$. *There exists* $\varepsilon := \mathsf{Cst}(x, a)$ *such that, for all* $n \geq \mathsf{Cst}(x, a)$, *if* $X_n \in T(x)$ *and* $x(n) \in B_{V_x}(\varepsilon/2)$, *then*

$$\mathbb{P}\big(\{\mathcal{R}_{n,\infty} = T(x)\} \cap \mathcal{L}(B_{V_x}(\varepsilon)) \cap \mathcal{A}_\partial(y(\infty)) | \mathcal{F}_n\big) > 0.$$

Moreover, the rate of convergence is at least reciprocally polynomial, that is, there exists $\nu := \mathsf{Cst}(x, a)$ *such that, a.s. on* $\mathcal{L}(B_{V_x}(\varepsilon))$,

$$\lim_{k\to\infty} \big(y(k) - y(\infty)\big)k^\nu = 0.$$

Proposition 4 is proved in Section 5.4. It obviously implies Theorem 3: indeed, given U a neighborhood of x, there exists $\varepsilon > 0$ such that $B_{V_x}(\varepsilon) \subseteq U$, and $X_n \in T(x)$ and $x(n) \in B_{V_x}(\varepsilon/2)$ occurs with positive probability if n is large enough.

Observe that, if $G = T(x)$, then this Proposition 4 is a direct consequence of Lemmas 7, 8 and 10. The localization with positive probability in this subgraph $T(x)$ results from a Borel–Cantelli type argument: the probability to visit $\partial T(x)$ at time n starting from $S(x)$ is, by Lemma 10, upper bounded by a term smaller than $n^{\alpha-2}$, where $\alpha \approx \max_{i\in\partial S} N_i(x)/H(x) < 1$, and $\sum_{n\in\mathbb{N}} n^{\alpha-2} < \infty$. Technically, the proof is based on a comparison of the probability of arbitrary paths remaining in $T(x)$ for the VRRWs defined, respectively, on the graphs $T(x)$ and G.

2.4. *Contents*. Section 3 concerns the results on the deterministic replicator dynamics: Lemmas 1–3 and Lemma 6 are proved, respectively, in Sections 3.1–3.3 and 3.4.

Section 4 develops the framework relating the behavior of the vector of density of occupation $x(n)$ to the replicator equation (6): we write the stochastic approximation equation (37) in Section 4.1, establish in Section 4.2 some preliminary estimates on the underlying Markov Chain $M(x)$, prove Lemma 4 in Section 4.3, prove Lemmas 5 and 9 [stochastic approximation equations for $V_q(z(n))$ and $H(z(n))$] and inclusions (15) in Sections 4.4 and 4.5.

Section 5 is devoted to the proofs of the asymptotic results for the VRRW: Lemma 7 in Section 5.1 on the convergence of $x(n)$ with positive probability, Lemma 8 in Section 5.2 on the corresponding speed of convergence, Lemma 10 in Section 5.3 on the asymptotic behavior of the number of visits on the frontier of the trapping subset, and Proposition 4 in Section 5.4 on localization with positive probability in the trapping subsets.

Finally, we show in Appendix A.1 a lemma on the remainder of square-bounded martingales, which is useful in the proofs of Lemma 8 and Proposition 1, whereas Appendix A.2 is devoted to the proof of Proposition 1.

3. Results on the replicator dynamics.

3.1. *Proof of Lemma* 1. Note that $DF(x)v = -H(x)v = 0$ if $S(v) \cap T(x) = \varnothing$, so that it is sufficient to study the eigenvalues of $DF(x)$ on $\{v \in \mathbb{R}^G \text{ s.t. } S(v) \subseteq T(x)\}$; hence, we can assume that G is finite [equal to $T(x)$] w.l.o.g.

Let $S := S(x)$ for convenience. For all $i, j \in G$,

$$\frac{\partial F_i}{\partial x_j} = \begin{cases} N_i(x) - H(x), & \text{if } x_i = 0 \text{ and } j = i, \\ 0, & \text{if } x_i = 0 \text{ and } j \neq i, \\ x_i[a_{i,j} - 2H(x)], & \text{if } x_i \neq 0 \text{ and } x_j \neq 0, \\ x_i[a_{i,j} - 2N_j(x)], & \text{if } x_i \neq 0 \text{ and } x_j = 0. \end{cases}$$

Let us now consider matrix $DF(x)$ by taking the following order on the indices: we take first the indices $i, j \in G \setminus S$, and second the indices $i, j \in S$,

$$\begin{pmatrix} \mathsf{Diag}\big(N_i(x) - H(x)\big)_{i \in G\setminus S} & (0) \\ (*) & DB \end{pmatrix},$$

where

$$B = [a_{i,j} - 2H(x)]_{i,j \in S}, \qquad D = \mathsf{Diag}(x_i)_{i \in S}.$$

The matrix DB is easily seen to be self-adjoint with respect to the scalar product $(u, v)_{D^{-1}} := (D^{-1}u, v)$. Hence, DB has real eigenvalues. This proves the first statement of the lemma.

Note that if we consider (6) as a differential equation on $\mathbb{R}^G$, then

$$(F(x), \mathbb{1}) = \frac{d(x(t), \mathbb{1})}{dt}\bigg|_{t=0, x(0)=x} = -\big((x, \mathbb{1}) - 1\big)H(x).$$

Therefore, if $x \in \Delta$ [which implies $(x, \mathbb{1}) = 1$], for all vector $u \in \mathbb{R}^G$,

$$(DF(x)u, \mathbb{1}) = -H(x)(u, \mathbb{1}). \tag{21}$$

Hence, $p : u \mapsto (u, \mathbb{1})$ is an eigenvector of ${}^t DF(x)$ with eigenvalue $-H(x)$. This makes $-H(x)$ an eigenvalue of $DF(x)$ and, more precisely,

$$\mathsf{Sp}(DF(x)) = \{-H(x)\} \cup \mathsf{Sp}(DF(x)|_{T\Delta});$$

indeed, by (21), an eigenvector u of $DF(x)$ with eigenvalue $\lambda \neq -H(x)$ belongs to $\mathsf{Ker} p = T\Delta$. Therefore, the stability of an equilibrium x of (6) on $\mathbb{R}^G$ is equivalent to the stability restricted on Δ, which completes the proof of the first equivalence in statement (b).

CLAIM. Let $M = \mathsf{Diag}(y_1, \dots, y_r)$ be a diagonal $r \times r$ matrix, with $y_1, \dots, y_r \in \mathbb{R}_+^*$, and let N be a symmetric $r \times r$ matrix. Then $\min \mathsf{Sp}(N) \geq 0 \iff \min \mathsf{Sp}(MN) \geq 0$ and, under this assumption,

$$\min \mathsf{Sp}(MN) \geq \min \mathsf{Sp}(N) \min\{y_i\}_{1 \leq i \leq r}.$$

PROOF. It suffices to prove that $\min \mathsf{Sp}(N) \geq 0$ implies $\min \mathsf{Sp}(MN) \geq 0$ and the corresponding inequality, since the coinverse statement is symmetrical.

Recall that, for any $r \times r$ symmetric matrix R with nonnegative eigenvalues, there exist a diagonal matrix D and an orthogonal matrix Q such that $R = Q^T DQ$, hence,

$$\min \mathsf{Sp}(R) = \inf_{|t| \geq 1} (Dt, t) = \inf_{|t| \geq 1} (DQt, Qt) = \inf_{|t| \geq 1} (Rt, t).$$

Let us define $L = \mathsf{Diag}(\sqrt{y_1}, \dots, \sqrt{y_r})$. Observe that $L^2 = M$. Now $MN = L(LNL)L^{-1}$ implies $\mathsf{Sp}(MN) = \mathsf{Sp}(LNL)$.

LNL is symmetric; therefore,

$$\begin{aligned}
\min \mathsf{Sp}(MN) = \min \mathsf{Sp}(LNL) &= \inf_{|t| \geq 1} (LNLt, t) \\
&= \inf_{|t| \geq 1} (NLt, Lt) \geq \inf_{|u| \geq \min_{1 \leq i \leq r} \sqrt{y_i}} (Nu, u) \\
&= \min_{1 \leq i \leq r} y_i \inf_{|u| \geq 1} (Nu, u) = \min_{1 \leq i \leq r} y_i \,\mathsf{Sp}(N).
\end{aligned}$$

□

To complete the proof of statement (b), we apply the claim to $M := D$ and $N := -B$.

It remains to prove that a stable equilibrium in Δ is feasible. Let $x \in \Delta$ be such an equilibrium. Assume that $H(x) = 0$. If $x_i = 0$ for some i then, by Lemma 1(b), $N_i(x) = 0$, so that $x_j = 0$ for all $j \sim i$. Hence, $x = 0$, which is contradictory. Now, if $x_i \neq 0$ for all i, then G is necessarily finite (by definition of Δ), and $a = (a_{i,j})_{i,j \in G} = 0$ since its eigenvalues are nonpositive [Lemma 1(b) again] and its trace is nonnegative. This is again contradictory.

3.2. *Proof of Lemma* 2. Let $\partial := \partial_S$, (P) := $(\mathrm{P})_S$ and $(\mathrm{P})' := (\mathrm{P})'_S$ for simplicity.

Assume (P) holds for some $d \geq 1$. Let us prove that, if $i, j, k \in S$ are such that $i \sim j \not\sim k$, then $a_{i,j} = a_{i,k}$.

If $i = j$, then $i = j \not\sim k$ implies, by (P)(a)–(b), that $k \notin S$—and therefore a contradiction—since if k were in S, it would be in the partition of i, which is a singleton. If $i \neq j \not\sim k$, then j and k are in the same partition of S. Hence, $a_{i,j} = a_{i,k}$ by (P)(c), which completes the proof of $(\mathrm{P})'$.

Assume now $(\mathrm{P})'$. Let us prove that the relation R defined on S by

$$iRj \quad \Longleftrightarrow \quad i \not\sim j \text{ or } i = j$$

is an equivalence relation on S. It is clearly symmetric and reflexive. Let us prove that it is transitive: let $i, j, k \in S$ be such that iRj and jRk, and prove iRk. This is immediate if $i = j$ or $j = k$; hence, assume that $i \neq j$ and $j \neq k$; then $(\mathrm{P})'$ implies $\partial_S\{i\} = \partial_S\{j\} = \partial_S\{k\}$. If we had $i \sim k$, then it would imply $k \in \partial_S\{i\} = \partial_S\{j\}$, and, therefore, $j \sim k$, which leads to a contradiction.

Now let us prove that there is only one element in the partition of a loop. Assume that iRj, $i \sim i$ and $j \neq i$ for $i, j \in S$; $(\mathrm{P})'$ implies in this case that $a_{i,i} = a_{i,j} > 0$, so that $i \sim j$, hence, $i = j$ since iRj holds, which leads to a contradiction.

Let V_p, $p = 1, \ldots, d$ be the partitions of R: elements of different partitions are connected, by definition, and (P)(a)–(b) holds for some $d \geq 1$. Let us prove (P)(c): let $p, q \in \{1, \ldots, d\}$ be such that $p \neq q$, and assume $i \in V_p$, $j \in V_q$. Let

$$W_{i,j} := \{(i', j') \in S^2 \text{ s.t. } a_{i',j'} = a_{i,j}\}.$$

By applying $(\mathrm{P})'$ twice, we first obtain that $W_{i,j} \supseteq \{i\} \times V_q$, and second that $W_{i,j} \supseteq V_p \times V_q$, which enables us to conclude.

3.3. *Proof of Lemma* 3. Let $S := S(x)$ and $(\mathrm{P})' := (\mathrm{P})'_{S(x)}$ for simplicity. Let

$$B = [a_{i,j} - 2H(x)]_{i,j \in S}.$$

Now $\max \mathsf{Sp}(B) \leq 0 \iff \forall t \in \mathbb{R}^S$, $B[t] \leq 0$. Observe that, for all $t = (t_i)_{i \in S} \in \mathbb{R}^S$,

$$B[t] = \sum_{i,j \in S} (a_{i,j} - 2H(x)) t_i t_j = H(t) - 2H(x) \left(\sum_{i \in S} t_i \right)^2.$$

Let us assume that $(\mathrm{P})'$ does not hold, and deduce that $B[t] > 0$ for some $t \in \mathbb{R}^S$, which will prove the first statement.

There exist $i, j, k \in S$ such that $j \not\sim k$ and $a_{i,j} \neq a_{i,k}$ [otherwise $(\mathrm{P})'$ would be satisfied]. Let, for all $\lambda \in \mathbb{R}$,

$$t_\lambda := \big(\mathbb{1}_{\{v=i\}} + \lambda \mathbb{1}_{\{v=j\}} - (1+\lambda)\mathbb{1}_{\{v=k\}}\big)_{v \in S} \in \mathbb{R}^S,$$

then

$$B[t_\lambda] \geq 2\lambda(a_{i,j} - a_{i,k}) - 2a_{i,k},$$

so that $B[t_\lambda] > 0$ for some $\lambda \in \mathbb{R}$, which yields the contradiction.

Let us now assume that (P)′ holds, and that $a_{i,j} = c\mathbb{1}_{i\sim j}$, with $c = 1$ for simplicity. First assume S contains no loop. Then, by Lemma 2, S is a d-partite subgraph for some $d \geq 1$ [$(\mathrm{P})_S$(a) holds]; let $V_1, \ldots, V_d$ be its partitions, then

$$\begin{aligned} B[t] &= \sum_{i,j\in S} (\mathbb{1}_{i\sim j} - 2H(x))t_i t_j = -2H(x)\Bigl(\sum_{i\in S} t_i\Bigr)^2 + \sum_{i,j\in S} \mathbb{1}_{i\sim j} t_i t_j \\ &= -2H(x)\Bigl(\sum_{k=1}^d v_k\Bigr)^2 + \Bigl(\sum_{k=1}^d v_k\Bigr)^2 - \sum_{k=1}^d v_k^2, \end{aligned}$$

where, for all $i \in \{1, \ldots, d\}$, $v_k = \sum_{i\in V_k} t_i$. Therefore,

$$B[t] = -\bigl(2H(x) - 1\bigr)\Bigl(\sum_{k=1}^d v_k\Bigr)^2 - \sum_{k=1}^d v_k^2 \leq 0,$$

where we use the fact that $H(x) \geq 1/2$, since $H(x) = 1 - 1/d$ and $d \geq 2$ (see proof of Theorem 2, Section 2.2.2).

Now assume that S contains one loop; then, again by the proof of Theorem 2, Section 2.2.2, it is a clique of loops and $H(x) = 1$; thus,

$$B[t] = -2\Bigl(\sum_{i\in S} t_i\Bigr)^2 + \Bigl(\sum_{i\in S} t_i\Bigr)^2 = -\Bigl(\sum_{i\in S} t_i\Bigr)^2 \leq 0.$$

3.4. *Proof of Lemma* 6. Let us first prove (a) in the case $q := x$, which will imply $H(q) = H(x)$ for any equilibrium $q \in \mathcal{N}(x)$ and therefore imply (a) in the general case. Let $x \in \mathcal{E}_s$, and let $y \in T\Delta$ be such that $x + y \in \Delta$. Let $S := S(x)$ for simplicity.

Recall that $G = S \cup \partial S$. We have

$$\begin{aligned} H(x+y) &= \sum_{i,j\in G} a_{i,j}(x_i + y_i)(x_j + y_j) = H(x) + 2\sum_{i\in G} N_i(x)y_i + H(y) &(22)\\ &= H(x) + 2\sum_{i\in G}\bigl(N_i(x) - H(x)\bigr)y_i + \sum_{i,j\in G}\bigl(a_{i,j} - 2H(x)\bigr)y_i y_j \\ &= H(x) + 2\sum_{i\in\partial S}\bigl(N_i(x) - H(x)\bigr)y_i \\ &\quad + \sum_{i,j\in S}\bigl(a_{i,j} - 2H(x)\bigr)y_i y_j + \sum_{i\in\partial S} w_i(y) &(23)\\ &\leq H(x) + 2\sum_{i\in\partial S}\bigl(N_i(x) - H(x)\bigr)y_i + \sum_{i\in\partial S} w_i(y). \end{aligned}$$

In the third equality, we make use of the identity $\sum_{i\in G} y_i = 0$, whereas in the fourth equality we notice that $N_i(x) = H(x)$ for all $i \in S$ and that the reinforcement matrix $a := (a_{i,j})_{i,j\in G}$ is symmetric, and let

$$w_i(y) := y_i\left(2\sum_{j\in S}(a_{i,j} - 2H(x))y_j + \sum_{j\in\partial S}(a_{i,j} - 2H(x))y_j\right)$$
$$= o_{|y|\to 0}(y_i) = o_{|y|\to 0}(y_{\partial S}),$$

using that, for all $j \in \partial S$, $y_j \geq 0$. Finally, we apply in the inequality that $B := (a_{i,j} - 2H(x))_{i,j\in S}$ is a negative semidefinite matrix by Lemma 1.

Using that, for all $i \in \partial S$, $N_i(x) < H(x)$ (and $y_i \geq 0$), we deduce that there exists a neighborhood $\mathcal{N}(x)$ of x in Δ such that, if $x + y \in \mathcal{N}(x)$, then $H(x + y) \leq H(x)$.

In order to obtain the required estimate of $H(x + y) - H(x)$, we observe that, if $z := (y_i)_{i\in S}$, then, by semidefiniteness of the symmetric matrix B,

$$-\mathsf{Cst}(x,a)|Bz|^2 \leq (Bz, z) = \sum_{i,j\in S}(a_{i,j} - 2H(x))y_i y_j \leq -\mathsf{Cst}(x,a)|Bz|^2. \tag{24}$$

But

$$Bz = \left(N_i(y) - 2H(x)\sum_{i\in S} y_i\right)_{i\in S} = (N_i(y) + 2H(x)y_{\partial S})_{i\in S},$$

where we use that $y_{\partial S} = -y_S$ in the second equality, since $y \in T\Delta$. Hence,

$$|Bz|^2 = \sum_{i\in S}(N_i(y) + 2H(x)y_{\partial S})^2 = \sum_{i\in S} N_i(y)^2 + o_{|y|\to 0}(y_{\partial S}) \tag{25}$$

and, if we let

$$K(y) := \sum_{i\in S} N_i(y)^2 + y_{\partial S},$$

then, by combining identities (23), (24) and (25) [and using that $w_i(y) = o_{|y|\to 0}(y_{\partial S})$ for all $i \in \partial S$], restricting $\mathcal{N}(x)$ if necessary,

$$-\mathsf{Cst}(x,a)K(y) \leq H(x + y) - H(x) \leq -\mathsf{Cst}(x,a)K(y). \tag{26}$$

On the other hand, let

$$L(y) := \sum_{i\in S}(N_i(x + y) - H(x + y))^2 + y_{\partial S}.$$

Then, again by restricting $\mathcal{N}(x)$ if necessary,

$$\mathsf{Cst}(x,a)L(y) \leq J(x + y) \leq \mathsf{Cst}(x,a)L(y), \tag{27}$$

where we use again that $N_i(x) < H(x)$ for all $i \in \partial S$. But

$$
\begin{aligned}
L(y) &= \sum_{i\in S}\bigl[N_i(y) - \bigl(H(x+y) - H(x)\bigr)\bigr]^2 + y_{\partial S} \\
&= K(y) + o_{|y|\to 0}\bigl(|H(x+y) - H(x)|\bigr).
\end{aligned}
\tag{28}
$$

Combining inequalities (26), (27) and (28), and further restricting $\mathcal{N}(x)$ if necessary, we obtain inequality (18) as required.

Let us now prove (b). If $q \in \mathcal{S}(S(x))$ and $y \in \Delta$, then

$$-\sum_{i\in S} q_i[N_i(y) - H(y)] = H(y) - \sum_{i\in S} q_i N_i(y)$$

and

$$\sum_{i\in S} q_i N_i(y) = \sum_{i\in G} q_i N_i(y) = \sum_{i\in G} y_i N_i(q) = H(q) + \sum_{i\in\partial S} y_i[N_i(q) - H(q)],$$

where we use that $(a_{i,j})_{i,j\in G}$ is symmetric in the second equality, and that q is an equilibrium in the third equality. Therefore,

$$I_q(y) = H(y) - H(q) + \sum_{i\in\partial S} y_i\bigl[2\bigl(N_i(y) - H(y)\bigr) - \bigl(N_i(q) - H(q)\bigr)\bigr]. \tag{29}$$

If $q, y \in \mathcal{N}(x)$, then [by restricting $\mathcal{N}(x)$ if necessary] $x \in \mathcal{E}_s$ implies that, for all $i \in \partial S$,

$$-\mathsf{Cst}(x,a) \le 2\bigl(N_i(y) - H(y)\bigr) - \bigl(N_i(q) - H(q)\bigr) \le -\mathsf{Cst}(x,a).$$

Inequality (19) follows.

4. Stochastic approximation results for the VRRW.

4.1. *The stochastic approximation equation.* We assume in this section that G is finite. The main idea is to modify the density of occupation measure

$$x(n) = \left(\frac{Z_n(i)}{n+n_0}\right)_{i\in G}$$

into a vector $z(n)$ that takes into account the position of the random walk, so that the conditional expectation of $z(n+1) - z(n)$ roughly only depends on $z(n)$ and not on the position X_n. This expectation will actually approximately be $F(z(n))/(n+n_0)$, where F is the map involved in the ordinary differential equation (6).

For all $x \in \Delta$, let $M(x)$ be the following matrix of transition probabilities of the reversible Markov chain:

$$M(x)(i,j) : \mathbb{1}_{i\sim j}\frac{a_{i,j}x_j}{\sum_{k\sim i} a_{i,k}x_k}; \tag{30}$$

$M(x(n))$ provides the transition probabilities from the VRRW at time n. Recall that $\pi(x)$ in (2) is the invariant probability measure for $M(x)$.

Let us denote by $\mathcal{G}$ (resp., $\mathcal{H}$) the set of functions on G taking values in $\mathbb{R}$ (resp., in $\mathbb{R}^G$). Let $\mathbb{1}$ be the function identically equal to 1. Let $M(x)$ and $\Pi(x)$ denote the linear transformations on $\mathcal{G}$ defined by

$$(M(x)f)(i) := \sum_{j\in G} M(x)(i,j)f(j), \tag{31}$$

$$\Pi(x)(f) := \Bigl(\sum_{i\in G} \pi(x)(i)f(i)\Bigr)\mathbb{1}. \tag{32}$$

Note that, by a slight abuse of notation, $M(x)$ equally denotes the Markov chain defined in (30) and its transfer operator in (31); $\Pi(x)$ is the linear transformation of $\mathcal{G}$ that maps f to the linear form identically equal to the mean of f under the invariant probability measure $\pi(x)$.

Any linear transformation P of $\mathcal{G}$ [and, in particular, $M(x)$ and $\Pi(x)$] also defines a linear transformation of $\mathcal{H}$: for all $f = (f_i)_{i\in G} \in \mathcal{H}$,

$$Pf := (Pf_i)_{i\in G}. \tag{33}$$

Let us now introduce a solution of the Poisson equation for the Markov chain $M(x)$. Let us define, for all $t \in \mathbb{R}_+$,

$$G_t(x) := e^{-t(I-M(x))} = e^{-t}\sum_0^\infty \frac{t^i M(x)^i}{i!},$$

which is the Markov operator of the continuous time Markov chain associated with $M(x)$. For all $x \in Int(\Delta)$, $M(x)$ is indecomposable so that $G_t(x)$ converges toward $\Pi(x)$ at an exponential rate, hence,

$$Q(x) := \int_0^\infty \bigl(G_t(x) - \Pi(x)\bigr)\,dt$$

is well defined. Note that

$$Q(x)\mathbb{1} = 0,$$

and that $Q(x)$ is the solution of the Poisson equation

$$\bigl(I - M(x)\bigr)Q(x) = Q(x)\bigl(I - M(x)\bigr) = I - \Pi(x), \tag{34}$$

using that $M(x)\Pi(x)f = \Pi(x)f = \Pi(x)M(x)f$ for all $f \in \mathcal{G}$ (or $f \in \mathcal{H}$).

Let us now expand $x(n+1) - x(n)$, using (34). Let $(e_i)_{i\in G}$ be the canonical basis of $\mathbb{R}^G$, that is, $e_i := (\mathbb{1}_{j=i})_{j\in G}$ for all $i \in G$. Let $\iota \in \mathcal{H}$ be defined by

$$\iota : G \longrightarrow \mathbb{R}^G,$$
$$i \longmapsto e_i.$$

First note that, for all $x \in \Delta$, $\Pi(x)\iota = \pi(x)\mathbb{1}$ since, for all $j \in G$,

$$\Pi(x)\iota(j) = ((\Pi(x)\iota_k)(j))_{k\in G} = ((\pi(x)(k)\mathbb{1})(j))_{k\in G} = \pi(x).$$

Therefore,

$$\begin{aligned}(n+n_0+1)\bigl(x(n+1)-x(n)\bigr) &= \bigl(\mathbb{1}_{X_{n+1}=i} - x(n)_i\bigr)_{i\in G} = \iota(X_{n+1}) - x(n)\\ &= \iota(X_{n+1}) - \pi(x(n)) + F(x(n))\\ &= [I - \Pi(x(n))]\iota(X_{n+1}) + F(x(n)),\end{aligned}$$

where F is the function defined in (5).

Now,

$$\begin{aligned}\frac{[I-\Pi(x(n))]\iota(X_{n+1})}{n+n_0+1} &= \frac{(Q(x(n)) - M(x(n)Q(x(n)))\iota(X_{n+1})}{n+n_0+1}\\ &= \varepsilon_{n+1} + \eta_{n+1} + r_{n+1,1} + r_{n+1,2},\end{aligned} \tag{35}$$

where

$$\begin{aligned}\varepsilon_{n+1} &:= \frac{Q(x(n))\iota(X_{n+1}) - M(x(n))Q(x(n))\iota(X_n)}{n+n_0+1},\\ r_{n+1,1} &:= \left(\frac{1}{n+n_0+1} - \frac{1}{n+n_0}\right)M(x(n))Q(x(n))\iota(X_n)\\ &= -\frac{M(x(n))Q(x(n))\iota(X_n)}{(n+n_0)(n+n_0+1)},\\ \eta_{n+1} &:= \frac{M(x(n))Q(x(n))\iota(X_n)}{n+n_0} - \frac{M(x(n+1))Q(x(n+1))\iota(X_{n+1})}{n+n_0+1},\\ r_{n+1,2} &:= \frac{[M(x(n+1))Q(x(n+1)) - M(x(n))Q(x(n))]\iota(X_{n+1})}{n+n_0+1}.\end{aligned}$$

Let, for all $n \in \mathbb{N}$,

$$z(n) := x(n) + \frac{M(x(n))Q(x(n))\iota(X_n)}{n+n_0} \tag{36}$$

and

$$\begin{aligned}r_{n+1,3} &:= \frac{1}{n+n_0+1}\frac{F(x(n)) - F(z(n))}{H(x(n))},\\ r_{n+1} &:= r_{n+1,1} + r_{n+1,2} + r_{n+1,3}.\end{aligned}$$

Then, for all $n \in \mathbb{N}$, it follows from equation (35) that

$$z(n+1) = z(n) + \frac{1}{n+n_0+1}\frac{F(z(n))}{H(x(n))} + \varepsilon_{n+1} + r_{n+1}. \tag{37}$$

Note that $\mathbb{E}(\varepsilon_{n+1}|\mathcal{F}_n) = 0$, since

$$\mathbb{E}(Q(x(n))\iota(X_{n+1})|\mathcal{F}_n) = M(x(n))Q(x(n))\iota(X_n);$$

also observe that

$$\sum_{i\in G} z(n)_i = \sum_{i\in G} x(n)_i + \frac{(M(x(n))Q(x(n))\mathbb{1})(X_n)}{n+n_0} = 1.$$

We provide in Section 4.2 estimates of the conditional variance of ε_{n+1} and of r_{n+1}, which will be sufficient to prove localization of the vertex-reinforced random walk with positive probability.

4.2. *Estimates on the underlying Markov chain $M(x)$.* For convenience we assume here that $G = S \cup \partial S$, where $(S, \sim)$ is finite, connected and not a singleton unless it is a loop. Let $\overline{a} := \max_{i,j\in G, i\sim j} a_{i,j}$, $\underline{a} := \min_{i,j\in G, i\sim j} a_{i,j}$.

Let us first introduce some general notation on Markov chains. Let K be a reversible Markov chain on the graph $(G, \sim)$, with invariant measure μ. Let $\langle\cdot,\cdot\rangle_\mu$ be the scalar product defined by, for all $f, g \in \mathcal{G}$,

$$\langle f, g\rangle_\mu := \sum_{x\in G} f(x)g(x)\mu(x).$$

On $\mathcal{G}$, we define the $\ell^p(\mu)$ norm, $1 \le p < \infty$ by

$$\|f\|_{\ell^p(\mu)} := \left(\sum_{x\in G} |f(x)|^p\mu(x)\right)^{1/p},$$

and the infinity norm

$$\|f\|_\infty := \max_{x\in G} |f(x)|.$$

We also define the infinity norm on $\mathcal{H}$: if $f = (f_i)_{i\in G} \in \mathcal{H}$,

$$\|f\|_\infty = \max_{i\in G} \|f_i\|_\infty = \max_{i,x\in G} |f_i(x)|. \tag{38}$$

Let $\mathbb{E}_\mu$ be the expectation operator

$$\mathbb{E}_\mu f := \sum_{x\in G} f(x)\mu(x) = \langle f, \mathbb{1}\rangle_\mu,$$

where $\mathbb{1}$ is the constant function equal to 1.

We let $\mathcal{E}_K$ be the Dirichlet form of K,

$$\mathcal{E}_K(f, g) = \langle (I - K)f, g\rangle_\mu,$$

and let Var_μ be the variance operator,

$$\mathsf{Var}_\mu(f) := \|f - \mathbb{E}_\mu f\|^2_{\ell^2(\mu)} = \|f\|^2_{\ell^2(\mu)} - (\mathbb{E}_\mu f)^2.$$

Simple calculations yield that

$$\mathcal{E}_K(f, f) = \frac{1}{2}\sum_{i\sim j}\big(f(i) - f(j)\big)^2 K(i, j)\mu(i),$$

and

$$\mathsf{Var}_\mu(f) = \frac{1}{2} \sum_{i,j \in G} \big(f(i) - f(j)\big)^2 \mu(i)\mu(j).$$

Let $\lambda(K)$ be the spectral gap of the Markov chain K,

$$\lambda(K) := \min\left\{ \frac{\mathcal{E}_K(f,f)}{\mathsf{Var}_\mu(f)} \text{ s.t. } \mathsf{Var}_\mu(f) \neq 0 \right\}.$$

The following Lemma 11 states that the spectral gap of the Markov chain $M(x)$ is lower bounded on Λ_α [defined in (14)].

LEMMA 11. *For all $x \in \Lambda_\alpha$, $\lambda(M(x)) \geq \mathsf{Cst}(\alpha, a, |G|)$.*

PROOF. Let $M := M(x)$ and $\pi := \pi(x)$ for simplicity. Let us first observe that, for all $i \in G$, $j \in S$ such that $i \sim j$,

$$\begin{aligned} &M(i,j) \geq \underline{a} x_j / \overline{a} \geq \alpha \underline{a} / \overline{a} \quad \text{and} \\ &M(i,j)\pi(i) = \pi(j) M(j,i) \geq \underline{a} \alpha^2 \mathbb{1}_{i \in S} / \overline{a}, \end{aligned} \tag{39}$$

where the second inequality comes from

$$M(i,j)\pi(i) = \frac{a_{i,j} x_j}{N_i(x)} \frac{x_i N_i(x)}{H(x)} = \frac{a_{i,j} x_i x_j}{H(x)} \geq \frac{\underline{a}\alpha^2}{\overline{a}} \mathbb{1}_{i \in S}.$$

Now, by connectedness of $(S, \sim)$, for all $i, j \in G$, there exists $l \leq |G|$ and a path $(n_k)_{1 \leq k \leq l} \in G \times S^{l-2} \times G$ such that $i = n_1$, $j = n_l$, $n_k \sim n_{k+1}$ for all $k \in \{1, \dots, l-1\}$.

Hence, for all $k \in \{1, \dots, l\}$, using inequalities (39),

$$\begin{aligned} &\pi(i)\pi(j)\big(f(i) - f(j)\big)^2 \\ &\quad \leq l\pi(i)\pi(j) \sum_{k \in \{1,\dots,l-1\}} \big(f(n_k) - f(n_{k+1})\big)^2 \\ &\quad \leq l\pi(i)\big(f(i) - f(n_2)\big)^2 + l\pi(j)\big(f(j) - f(n_{l-1})\big)^2 \\ &\qquad + l \sum_{k \in \{2,\dots,l-2\}} \big(f(n_k) - f(n_{k+1})\big)^2 \\ &\quad \leq \frac{\overline{a} l}{\underline{a}\alpha} \big[M(i, n_2)\pi(i)\big(f(i) - f(n_2)\big)^2 \\ &\qquad\qquad + M(j, n_{l-1})\pi(j)\big(f(j) - f(n_{l-1})\big)^2 \big] \\ &\qquad + \frac{\overline{a} l}{\underline{a}\alpha^2} \sum_{k \in \{2,\dots,l-2\}} \big(f(n_k) - f(n_{k+1})\big)^2 M(n_k, n_{k+1})\pi(n_k) \end{aligned}$$

$$\leq \frac{\overline{a}l}{\underline{a}\alpha^2} \sum_{k\in\{1,\dots,l-1\}} \bigl(f(n_k) - f(n_{k+1})\bigr)^2 M(n_k, n_{k+1})\pi(n_k)$$
$$\leq \frac{2\overline{a}|G|}{\underline{a}\alpha^2} \mathcal{E}_M(f, f).$$

Therefore,

$$\mathsf{Var}_\pi(f) = \frac{1}{2} \sum_{i,j\in G} \pi(i)\pi(j)\bigl(f(i) - f(j)\bigr)^2 \leq \frac{\overline{a}|G|^3}{\underline{a}\alpha^2} \mathcal{E}_M(f, f). \qquad \square$$

Lemma 12 provides upper bounds on the norms of $Q(x)$, $M(x)Q(x)$ and their partial derivatives on Λ_α, which will be needed in the estimates of r_{n+1} and of the conditional variance of ε_{n+1} in Lemma 4.

The norm on linear transformations of $\mathcal{G}$ will be the infinity norm

$$\|A\|_\infty := \sup_{f\in\mathcal{G}, f\neq 0} \frac{\|Af\|_\infty}{\|f\|_\infty}.$$

Note that, for any linear transformation A of $\mathcal{G}$, the corresponding linear transformation of $\mathcal{H}$ (still called A) defined in (33) still has the same infinity norm [the $\|\cdot\|_\infty$ on $\mathcal{H}$ is defined by (38)],

$$\|A\|_\infty = \sup_{f\in\mathcal{H}, f\neq 0} \frac{\|Af\|_\infty}{\|f\|_\infty}.$$

LEMMA 12. *For all $x \in \Lambda_\alpha$, $i, j \in G$, $f \in \mathcal{G}$:*

(a) $M(x)(i, j) \leq \left(\dfrac{\overline{a}}{\underline{a}}\right)^2 \dfrac{\pi(x)(j)}{\alpha^2}$,

(b) $\|Q(x)f\|_{\ell^2(\pi(x))} \leq \dfrac{\sqrt{\mathsf{Var}_{\pi(x)}(f)}}{\lambda(M(x))} \leq \dfrac{\|f\|_{\ell^2(\pi(x))}}{\lambda(M(x))}$,

(c) $\|Q(x)\|_\infty \leq \mathsf{Cst}(\alpha, a, |G|)$, $\qquad \|M(x)Q(x)\|_\infty \leq \mathsf{Cst}(\alpha, a, |G|)$,

(d) $\left\|\dfrac{\partial Q(x)}{\partial x_i}\right\|_\infty \leq \mathsf{Cst}(\alpha, a, |G|)$, $\qquad \left\|\dfrac{\partial(M(x)Q(x))}{\partial x_i}\right\|_\infty \leq \mathsf{Cst}(\alpha, a, |G|)$.

PROOF. Let $M := M(x)$, $Q := Q(x)$, $\pi := \pi(x)$, $\lambda := \lambda(M(x))$ for simplicity.

Inequality (a) is obvious: for all $j \in G$,

$$M(i, j) = \frac{a_{i,j}x_j}{N_i(x)} = \frac{x_j N_j(x)}{H(x)} \frac{a_{i,j} H(x)}{N_i(x)N_j(x)} \leq \left(\frac{\overline{a}}{\underline{a}}\right)^2 \frac{\pi(j)}{\alpha^2}.$$

Let us now prove (b). For all $f \in \mathcal{G}$,

$$\|G_t f - \pi(f)\|^2_{\ell^2(\pi)} \leq e^{-2\lambda t} \mathsf{Var}_\pi(f),$$

by definition of the spectral gap (see, e.g., Lemma 2.1.4, [17]), so that

$$\begin{aligned}\|Q(x)f\|_{\ell^2(\pi)} &\le \left\|\int_0^\infty \big(G_t(x)f - \Pi(x)f\big)\,dt\right\|_{\ell^2(\pi)} \\ &\le \int_0^\infty \big\|\big(G_t(x)f - \Pi(x)f\big)\big\|_{\ell^2(\pi)}\,dt \\ &\le \sqrt{\mathsf{Var}_\pi(f)}\int_0^\infty e^{-\lambda t}\,dt = \frac{\sqrt{\mathsf{Var}_\pi(f)}}{\lambda} \le \frac{\|f\|_{\ell^2(\pi)}}{\lambda}.\end{aligned} \tag{40}$$

Inequality (c) translates this upper bound of the $\ell^2(\pi) \to \ell^2(\pi)$-norm of $Q(x)$ into one involving the infinity norm for MQ, using (a):

$$\begin{aligned}|MQf(i)| &= \left|\sum_{j\in G} M(i,j)Qf(j)\right| \\ &\le \frac{1}{\alpha^2}\left(\frac{\overline{a}}{\underline{a}}\right)^2 \sum_{j\in G}\pi(j)|Qf(j)| = \left(\frac{\overline{a}}{\underline{a}}\right)^2 \frac{\|Qf\|_{\ell^1(\pi)}}{\alpha^2} \\ &\le \left(\frac{\overline{a}}{\underline{a}}\right)^2 \frac{\|Qf\|_{\ell^2(\pi)}}{\alpha^2} \le \left(\frac{\overline{a}}{\underline{a}}\right)^2 \frac{\|f\|_{\ell^2(\pi)}}{\lambda\alpha^2}.\end{aligned}$$

Hence, using Lemma 11,

$$\|MQf\|_\infty \le \left(\frac{\overline{a}}{\underline{a}}\right)^2 \frac{\|f\|_{\ell^2(\pi)}}{\lambda\alpha^2} \le \left(\frac{\overline{a}}{\underline{a}}\right)^2 \frac{\|f\|_\infty}{\lambda\alpha^2} \le \mathsf{Cst}(\alpha, a, |G|)\|f\|_\infty.$$

Then the same upper bound for $\|Q(x)f\|_\infty$ follows from the Poisson equation (34):

$$Q(x) = M(x)Q(x) + I - \Pi(x).$$

Let us now prove (d). Given $i \in G$, let us take the derivative of the Poisson equation $Q(x)(I - M(x)) = I - \Pi(x)$ with respect to x_i:

$$\frac{\partial Q(x)}{\partial x_i}\big(I - M(x)\big) = Q(x)\frac{\partial M(x)}{\partial x_i} - \frac{\partial \Pi(x)}{\partial x_i}.$$

This equality, multiplied on the right by $Q(x)$, yields, using now the Poisson equation $(I - M(x))Q(x) = I - \Pi(x)$,

$$\frac{\partial Q(x)}{\partial x_i} = \frac{\partial Q(x)}{\partial x_i}\big(I - \Pi(x)\big) = \left(Q(x)\frac{\partial M(x)}{\partial x_i} - \frac{\partial \Pi(x)}{\partial x_i}\right)Q(x), \tag{41}$$

where we use that, for all $f \in \mathcal{G}$,

$$\frac{\partial Q(x)}{\partial x_i}\Pi(x)f = \langle f, \mathbb{1}\rangle_{\pi(x)}\frac{\partial Q(x)}{\partial x_i}\mathbb{1} = 0,$$

since $Q(x)\mathbb{1} = 0$ for all $x \in \Delta$.

Equality (41) implies the required upper bound of $\|\frac{\partial Q(x)}{\partial x_i}\|_\infty$. Indeed, the following estimates hold: for all $i, j, k \in G$, $j \sim k$,

$$\begin{aligned}\left|\frac{\partial[M(x)(j,k)]}{\partial x_i}\right| &= \left|\frac{\partial}{\partial x_i}\left(\frac{a_{j,k}x_k}{N_j(x)}\right)\right| \\ &= \left|\frac{\partial x_k}{\partial x_i}\frac{a_{j,k}}{N_j(x)} - \frac{a_{j,k}x_k}{N_j(x)^2}\frac{\partial N_j(x)}{\partial x_i}\right| \\ &\leq \frac{2\overline{a}}{N_j(x)} \leq \frac{2\overline{a}}{\underline{a}\alpha},\end{aligned}$$

where we use that $a_{j,k}x_k \leq N_j(x)$ and $\partial N_j/\partial x_i(x) = a_{j,i}$, and that there exists $l \in S$ with $l \sim j$, given the assumptions on S. Also,

$$\begin{aligned}\left|\frac{\partial \pi(x)(j)}{\partial x_i}\right| &= \left|\frac{\partial}{\partial x_i}\left(\frac{x_j N_j(x)}{H(x)}\right)\right| \\ &= \left|\frac{\partial(x_j N_j(x))}{\partial x_i}\frac{1}{H(x)} - \frac{x_j N_j(x)}{H(x)^2}\frac{\partial H(x)}{\partial x_i}\right| \\ &\leq \frac{4\overline{a}}{H(x)} \leq \frac{4\overline{a}}{\underline{a}\alpha^2},\end{aligned}$$

where we note that $|\frac{\partial H(x)}{\partial x_i}| = 2N_i(x) \leq 2\overline{a}$. The upper bound of $\|\frac{\partial(M(x)Q(x))}{\partial x_i}\|_\infty$ follows directly. $\square$

4.3. *Proof of Lemma* 4. The estimates (a) and (d) readily follow from the definitions of ε_{n+1} and $z(n)$, and from Lemma 12(c).

Let $M := M(x(n))$, $Q := Q(x(n))$, $\pi := \pi(x(n))$, $\lambda := \lambda(M(x(n)))$ for simplicity. Let us prove (b):

$$\begin{aligned}(n+n_0)^2\mathbb{E}((\varepsilon_{n+1})_i^2|\mathcal{F}_n) &\leq \mathbb{E}([Qe_i(X_{n+1})]^2|\mathcal{F}_n) \\ &= \sum_{j\sim X_n} M(X_n, j)[Qe_i(j)]^2 \\ &\leq \frac{1}{\alpha^2}\left(\frac{\overline{a}}{\underline{a}}\right)^2 \sum_{j\in G}\pi(j)[Qe_i(j)]^2 \\ &= \left(\frac{\overline{a}}{\underline{a}}\right)^2 \frac{1}{\alpha^2}\|Qe_i\|^2_{\ell^2(\pi(x(n)))} \\ &\leq \mathsf{Cst}(\alpha, a, |G|)\|e_i\|^2_{\ell^2(\pi(x(n)))} \leq \mathsf{Cst}(\alpha, a, |G|)x(n)_i,\end{aligned}$$

where we use Lemma 12(a) and (b), respectively, in the second and in the third inequality.

In order to prove (c), let us first upper bound $\|r_{n+1,1}\|_\infty$ using Lemma 12(c):

$$\|r_{n+1,1}\|_\infty \le \frac{\|M(x(n))Q(x(n))\iota(X_n)\|_\infty}{(n+n_0)^2} \le \frac{\mathsf{Cst}(\alpha, a, |G|)}{(n+n_0)^2}.$$

Let us now bound $\|r_{n+1,2}\|_\infty$:

$$\begin{aligned}(n+n_0)\|r_{n+1,2}\|_\infty &\le \sup_{\theta\in[0,1]}\left\|\frac{\partial(MQ)(\theta x(n)+(1-\theta)x(n+1))}{\partial\theta}\right\|_\infty \\ &\le \sum_{i\in G}\left|\bigl(x(n+1)-x(n)\bigr)_i\right| \\ &\qquad \times \sup_{i\in G,\theta\in[0,1]}\left\|\frac{\partial(MQ)(\theta x(n)+(1-\theta)x(n+1))}{\partial x_i}\right\|_\infty \\ &\le \frac{\mathsf{Cst}(\alpha, a, |G|)}{n+n_0},\end{aligned}$$

where we use Lemma 12(d) in the last inequality.

It remains to upper bound $\|r_{n+1,3}\|_\infty$. First observe that, for all $y = (y_i)_{i\in G}$, $z = (z_i)_{i\in G} \in \Delta$, $i \in G$,

$$|F_i(z) - F_i(y)| \le \sum_{j\in G}|z_j - y_j| \sup_{k\in G, x\in\Delta}\left|\frac{\partial F_i(x)}{\partial x_k}\right| \le 2\overline{a}\sum_{i\in G}|z_i - y_i|,$$

where we use the explicit computations of $\partial F_i/\partial x_j$ in the proof of Lemma 1. Hence,

$$\|F(z) - F(y)\|_\infty \le 2\overline{a}|G|\|z - y\|_\infty,$$

which implies

$$\|r_{n+1,3}\|_\infty \le \frac{1}{n+n_0}\frac{|G|}{\underline{a}}2\overline{a}|G|\|x(n) - z(n)\|_\infty \le \frac{\mathsf{Cst}(\alpha, a, |G|)}{(n+n_0)^2},$$

where we use that, by inequality (8), $H(x) \ge \underline{a}/|G|$ for all $x \in \Delta$.

4.4. *Proof of Lemma 5 and inclusions* (15). Let us first prove inclusions (15). If we let $g : \mathbb{R}_+ \setminus \{0\} \longrightarrow \mathbb{R}_+$ be the function defined by $g(u) := u - \log(u+1)$, nonnegative by concavity of the log function, then, for all $y \in \Delta$ such that $y_i > 0$ for all $i \in S$,

$$V_q(y) = -\sum_{i\in S} q_i \log\left(1 + \frac{y_i - q_i}{q_i}\right) + 2y_{\partial S} = \sum_{i\in S} q_i g\left(\frac{y_i - q_i}{q_i}\right) + 3y_{\partial S}, \tag{42}$$

which implies the inclusions.

Let us now prove Lemma 5; let, for all $n \in \mathbb{N}$,

$$\zeta_{n+1} := \left(\frac{(\varepsilon_{n+1})_i}{z(n)_i}\mathbb{1}_{i\in S}\right)_{i\in G},$$

with the convention that $\zeta_{n+1} = 0$ if $z(n)_i = 0$ for some $i \in S$. Fix $\varepsilon > 0$ such that $B_{V_q}(2\varepsilon) \subseteq \Lambda_\alpha$ for some $\alpha = \mathsf{Cst}(q) > 0$, and assume $x(n) \in B_{V_q}(\varepsilon)$ for some $n \geq n_1$. Thus, $\|z(n) - x(n)\|_\infty \leq \mathsf{Cst}(q, a)/(n + n_0)$ by Lemma 4(d); we assume in the rest of the proof that $\varepsilon < \mathsf{Cst}(q)$ and $n_0 \geq \mathsf{Cst}(q, a)$ so that, using (42), $z(n) \in B_{V_q}(2\varepsilon) \subseteq \Lambda_\alpha$.

Note that $\|x(n) - x(n+1)\|_\infty \leq (n + n_0)^{-1}$, which implies, using Lemma 4, that $\|z(n) - z(n+1)\|_\infty \leq \mathsf{Cst}(q, a)(n + n_0)^{-1}$. Hence, using that $z(n) \in \Lambda_\alpha$,

$$\begin{aligned} V_q\big(z(n+1)\big) - V_q(z(n)) &= -\sum_{i \in S} q_i \log\left(\frac{z(n+1)_i}{z(n)_i}\right) + 2[z(n+1)_{\partial S} - z(n)_{\partial S}] \\ &= -\sum_{i \in S} q_i \frac{z(n+1)_i - z(n)_i}{z(n)_i} + 2[z(n+1)_{\partial S} - z(n)_{\partial S}] \\ &\quad + \square\left(\frac{\mathsf{Cst}(q, a)}{(n + n_0)^2}\right), \end{aligned}$$

where we again make use of notation $u = \square(v) \iff |u| \leq v$ from Section 2.1.

Hence, using identity (37) and Lemma 4(c)–(d), we obtain subsequently [recall that $I_q(\cdot)$ is defined in (16)]

$$\begin{aligned} &V_q\big(z(n+1)\big) - V_q(z(n)) \\ &\quad = \frac{1}{n + n_0 + 1} \frac{I_q(z(n))}{H(x(n))} - (q, \zeta_{n+1}) + 2(\varepsilon_{n+1})_{\partial S} + \square\left(\frac{\mathsf{Cst}(q, a)}{(n + n_0)^2}\right) \\ &\quad = \frac{1}{n + n_0 + 1} \frac{I_q(x(n))}{H(x(n))} - (q, \zeta_{n+1}) + 2(\varepsilon_{n+1})_{\partial S} + \square\left(\frac{\mathsf{Cst}(q, a)}{(n + n_0)^2}\right). \end{aligned}$$

4.5. *Proof of Lemma* 9. Using identities (22) and (37) [recall that J is defined in (7)],

$$\begin{aligned} H\big(z(n+1)\big) - H(z(n)) &= 2 \sum_{i \in G} N_i(z(n)) \cdot \big(z(n+1) - z(n)\big)_i \\ &\quad + H\big(z(n+1) - z(n)\big) \\ &= \frac{1}{n + n_0 + 1} \frac{J(z(n))}{H(x(n))} + \xi_{n+1} + s_{n+1}, \end{aligned}$$

where

$$\begin{aligned} \xi_{n+1} &:= 2 \sum_{i \in G} N_i(z(n))(\varepsilon_{n+1})_i, \\ s_{n+1} &:= 2 \sum_{i \in G} N_i(z(n))(r_{n+1})_i + H\big(z(n+1) - z(n)\big). \end{aligned}$$

Let $\alpha > 0$, and assume $x(n) \in \Lambda_\alpha$. Inequalities (1) and (2) of our lemma follow from Lemma 4(a)–(c), and from $\|z(n+1) - z(n)\|_\infty \leq \mathsf{Cst}(\alpha, a, |G|)/(n + n_0)$ (see, e.g., the beginning of the proof of Lemma 5).

5. Asymptotic results for the VRRW.

5.1. *Proof of Lemma 7*. Fix $\varepsilon > 0$ such that $B_{V_x}(\varepsilon) \subseteq \Lambda_\alpha$ for some $\alpha > 0$ depending on x, and assume $x(n) \in B_{V_x}(\varepsilon/2)$ for some $n \geq n_1$.

Let $(\zeta_k)_{k\geq 2}$ be defined as in Section 4.4, and let us define the martingales $(A_k)_{k\geq n}$, $(B_k)_{k\geq n}$ and $(\kappa_k)_{k\geq n}$ by

$$A_k := \sum_{j=n+1}^{k} \zeta_j \mathbb{1}_{\{V_x(x(j-1))\leq\varepsilon\}}, \qquad B_k := \sum_{j=n+1}^{k} (\varepsilon_j)_{\partial S} \mathbb{1}_{\{V_x(x(j-1))\leq\varepsilon\}},$$

$$\kappa_k := -(q, A_k) + 2B_k,$$

with the convention that $A_n := 0$ and $B_n = \kappa_n := 0$. Using Lemma 4(a), it follows from Doob's convergence theorem that $(A_k)_{k\geq n}$, $(B_k)_{k\geq n}$ and $(\kappa_k)_{k\geq n}$ converge a.s. and in $\mathcal{L}^2$.

Let us briefly outline the proof: we first show that, on an event of large probability Υ, where κ_k, $k \geq n$, remains small, $x(k)$ remains in the neighborhood of x and the stochastic approximation (17) remains valid. This implies, together with (19), the existence of a subsequence j_k such that $(x(j_k))_{k\geq 0}$ converges to a random $r \in \mathcal{E}_s$ [see (44)]. Using the linearity of the martingale part of (17) in ζ and ε, we can conclude from the a.s. convergence of $(A_k)_{k\geq n}$ and $(B_k)_{k\geq n}$ that $x(k) \longrightarrow_{k\to\infty} r$ a.s. [see (45) and (46)].

The upper bound $|\kappa_k - \kappa_{k-1}| \leq \Gamma/(k+n_0)$ a.s., for some $\Gamma := \mathsf{Cst}(x, a)$, implies that, for all $k \geq n+1$ and $\theta \in \mathbb{R}$,

$$\mathbb{E}\big(\exp(\theta(\kappa_k - \kappa_{k-1}))|\mathcal{F}_{k-1}\big) \leq \exp\left(\frac{\Gamma^2}{2}\frac{\theta^2}{(k+n_0)^2}\right).$$

On the other hand, $(\exp(\theta\kappa_k))_{k\geq n}$ is a submartingale since $(\kappa_k)_{k\geq n}$ is a martingale, so that Doob's submartingale inequality implies, for all $\theta > 0$,

$$\mathbb{P}\Big(\sup_{k\geq n}\kappa_k \geq c|\mathcal{F}_n\Big) = \mathbb{P}\Big(\sup_{k\geq n} e^{\theta\kappa_k} \geq e^{\theta c}|\mathcal{F}_n\Big) \leq e^{-\theta c}\mathbb{E}(e^{\theta\kappa_\infty}|\mathcal{F}_n)$$
$$\leq \exp\left(-\theta c + \frac{\theta^2\Gamma^2}{2(n+n_0)}\right).$$

Choosing $\theta := c(n+n_0)/\Gamma^2$ yields

$$\mathbb{P}\Big(\sup_{k\geq n}\kappa_k \geq c\big|\mathcal{F}_n\Big) \leq \exp\left(-\frac{c^2}{2\Gamma^2}(n+n_0)\right). \tag{43}$$

Let

$$\Upsilon := \left\{\sup_{k\geq n}\kappa_k < \frac{\varepsilon}{12}\right\};$$

inequality (43) implies that

$$\mathbb{P}(\Upsilon|\mathcal{F}_n) \geq 1 - \exp\bigl(-\varepsilon^2 \mathsf{Cst}(x,a)(n+n_0)\bigr).$$

Now assume that Υ holds, and let T be the stopping time

$$T := \inf\{k \geq n \text{ s.t. } V_x(z(k)) \geq 2\varepsilon/3\}.$$

Note that, using Lemma 4(d), if $n \geq \mathsf{Cst}(x,a)$, then for all $k \in [n,T)$, $V_x(x(k)) < \varepsilon$. We upper bound $V_x(x(T)) - V_x(x(k))$ by adding up identity (17) in Lemma 5 with $q := x$, from time n to $T-1$: this yields, together with Lemma 6, that $V_x(z(T)) < 2\varepsilon/3$ if $T < \infty$, if we assume $n \geq n_1 := \mathsf{Cst}(x,a)$ large enough and $\varepsilon < \varepsilon_0 := \mathsf{Cst}(x,a)$ small enough.

Therefore, $V_x(x(k)) < \varepsilon$ for all $k \geq n$. Using again identity (17) [and Lemma 6(b)], we obtain subsequently that

$$\liminf_{k\to\infty}[H(x) - H(x(k)) + x(k)_{\partial S}] = 0 \qquad \text{a.s.}$$

since, otherwise, the convergence of (κ_k) as $k \to \infty$ would imply $\lim_{k\to\infty} V_x(z(k)) = \lim_{k\to\infty} V_x(x(k)) = -\infty$, which is in contradiction with $V_x(x(k)) \geq 0$.

Hence, there exists a (random) increasing sequence $(j_k)_{k\geq 0}$ such that

$$\lim_{k\to\infty} H(x(j_k)) = H(x), \qquad \lim_{k\to\infty} x(j_k)_{\partial S} = 0. \tag{44}$$

Let r be an accumulation point of $(x(j_k))_{k\geq 0}$. Then $H(r) = H(x)$ and $r_{\partial S} = 0$.

Note that $V_x(r) = \lim_{k\to\infty} V_x(z(j_k)) \leq \varepsilon$. By possibly choosing a smaller $\varepsilon_0 := \mathsf{Cst}(x,a)$, we obtain by Lemma 6 that r is an equilibrium, and by Lemma 1 that it is strictly stable.

Let, for all $j \in \mathbb{N}$,

$$\Lambda_j := \Bigl\{\sup_{k\geq j} |A_k - A_j| < \frac{\varepsilon}{24}\Bigr\} \cap \Bigl\{\sup_{k\geq j} |B_k - B_j| < \frac{\varepsilon}{24}\Bigr\}.$$

There exists a.s. $j \in \mathbb{N}$ such that Λ_j holds; let l_0 be such a j (l_0 is random, and is not a stopping time).

Let $k \in \mathbb{N}$ be such that $j_k \geq l_0$ and $V_r(z(j_k)) < \varepsilon/2$. Then Lemma 5 applies to $r \in \mathcal{S} \cap \mathcal{E}_s$ and a similar argument as previously shows that, for all $j' \geq j \geq j_k$, $V_r(x(j)) \leq \varepsilon$ and

$$V_r(z(j')) \leq V_r(z(j)) + \sup_{k\geq j} |A_k - A_j| + 2\sup_{k\geq j} |B_k - B_j| + \frac{\mathsf{Cst}(q,a)}{j+n_0}, \tag{45}$$

if $n_1 := \mathsf{Cst}(x,a)$ was chosen sufficiently large.

Now, $\liminf_{j\to\infty} V_r(z(j)) = 0$ and

$$\lim_{j\to\infty} \sup_{k\geq j} |A_k - A_j| = \lim_{j\to\infty} \sup_{k\geq j} |B_k - B_j| = \lim_{j\to\infty} \frac{\mathsf{Cst}(q)}{j+n_0} = 0, \tag{46}$$

hence, $\lim_{j\to\infty} V_r(x(j)) = 0$ which implies $\lim_{j\to\infty} x(j) = r$ and completes the proof.

5.2. *Proof of Lemma* 8. Let us start with an estimate of the rate of convergence of $H(z(n))$ to $H(x)$. Let, for all $n \in \mathbb{N}$,

$$\chi_n := H(x) - H(z(n)), \nu_n := \frac{J(z(n))}{H(x(n))\chi_n},$$

with the convention that $\nu_n := 0$ if $\chi_n = 0$.

By Lemma 6 there exist ε, λ, $\mu := \mathsf{Cst}(x, a)$ such that, for all $n \in \mathbb{N}$ such that $x(n) \in B_{V_x}(2\varepsilon)$, $\nu_n \in [\lambda, \mu]$. On the other hand, for all $n \in \mathbb{N}$, using Lemma 9 and the observation that $J(z(n)) = 0$ if $\chi_n = 0$ by Lemma 6,

$$\begin{aligned}\chi_{n+1} &= \left(1 - \frac{\nu_n}{n + n_0 + 1}\right)\chi_n - \xi_{n+1} - s_{n+1}\\ &\le \left(1 - \frac{\lambda}{n + n_0 + 1}\right)\chi_n - \xi_{n+1} + s'_{n+1},\end{aligned} \tag{47}$$

where

$$s'_{n+1} := -s_{n+1} + (\nu_n - \lambda)\max(-\chi_n, 0)/(n + n_0 + 1).$$

If $x(n) \in B_{V_x}(2\varepsilon)$ for sufficiently small $\varepsilon := \mathsf{Cst}(x, a)$, then, by Lemma 9,

$$\|\xi_{n+1}\|_\infty \le \frac{\mathsf{Cst}(x, a)}{n + n_0}, \qquad \|s'_{n+1}\|_\infty \le \frac{\mathsf{Cst}(x, a)}{(n + n_0)^2}, \tag{48}$$

where we use in the second inequality that $\max(-\chi_n, 0) \le \mathsf{Cst}(x, a)/(n + n_0 + 1)$, since $\|x(n) - z(n)\|_\infty \le \mathsf{Cst}(x, a)/(n + n_0 + 1)$ by Lemma 4(d), and $H(x(n)) \le H(x)$ by Lemma 6.

Let, for all $n \in \mathbb{N}$,

$$\beta_n := \prod_{k=1}^{n}\left(1 - \frac{\lambda}{k + n_0}\right).$$

Note that $\beta_n n^\lambda$ converges to a positive limit. Inequality (47) implies by induction that, for all $n \in \mathbb{N}$,

$$\chi_n \le \beta_n\left(\chi_0 - \sum_{j=1}^{n}\frac{\xi_j}{\beta_j} + \sum_{j=1}^{n}\frac{s'_j}{\beta_j}\right).$$

Assume $\mathcal{L}(B_{V_x}(\varepsilon))$ holds so that, in particular, $x(n) \in \mathcal{L}(B_{V_x}(2\varepsilon))$ for large $n \in \mathbb{N}$. The upper bounds (48) yield, assuming w.l.o.g. $\lambda < 1/2$, that $\sum_{j=1}^n s'_j/\beta_j < \infty$ and $\sum_{j=1}^n \mathbb{E}(\xi_j^2)/\beta_j^2 < \infty$; the latter implies, by the Doob convergence theorem in $\mathcal{L}^2$, that $\sum_{j=1}^n \xi_j/\beta_j$ converges a.s. Therefore, $\chi_n n^\lambda$ is bounded a.s.

We deduce subsequently, by Lemma 6(a), that for all $\lambda \le \mathsf{Cst}(x, a)$, $J(x(n))n^\lambda$ converges a.s. to 0, so that $\lim_{n\to\infty} x(n)_{\partial S} n^\lambda = 0$ in particular. This implies that $\lim_{n\to\infty} I_{x(\infty)}(x(n))n^\lambda = 0$ by Lemma 6(b).

Now apply Lemma 5 with $q := x(\infty)$: for large $n \in \mathbb{N}$,

$$\begin{aligned}V_{x(\infty)}(z(n)) &= -\sum_{k=n}^{\infty} \frac{I_{x(\infty)}(x(k))}{k+n_0+1} + \left(x(\infty), \sum_{k=n+1}^{\infty} \zeta_k\right) - 2\sum_{k=n+1}^{\infty} (\varepsilon_k)_{\partial S} \\ &\qquad + \mathsf{Cst}(x,a)\Box\left(\sum_{k=n}^{\infty} \frac{1}{(k+n_0)^2}\right) \\ &= o(n^{-\lambda}) \qquad \text{a.s.,}\end{aligned}$$

if we still assume w.l.o.g. $\lambda < 1/2$, so that $\sum_{k=n+1}^{\infty} (\varepsilon_k)_{\partial S} = o(n^{-\lambda})$ a.s by Lemmas 4(a) and A.1. This completes the proof of the lemma, using (42).

5.3. *Proof of Lemma* 10. Let, for all $n \in \mathbb{N}$ and $i, j \in G$, $i \sim j$,

$$Y_n^{i,j} := \sum_{k=1}^{n} \frac{\mathbb{1}_{\{X_{k-1}=i, X_k=j\}}}{Z_{k-1}(j)}, \qquad Y_n^i := \sum_{k=1}^{n} \frac{\mathbb{1}_{\{X_{k-1}=i\}}}{\sum_{j\sim i} a_{j,i} Z_{k-1}(j)}.$$

Then, by definition of the vertex-reinforced random walk,

$$M_n^{i,j} := Y_n^{i,j} - a_{i,j} Y_n^i$$

is a martingale, and

$$\begin{aligned}&\sum_{k=1}^{\infty} \mathbb{E}\big((M_k^{i,j} - M_{k-1}^{i,j})^2\big) \\ &\quad = \mathbb{E}\left(\sum_{k=1}^{\infty} \frac{\mathbb{1}_{\{X_{k-1}=i\}}}{Z_{k-1}(j)^2} \frac{a_{i,j} Z_{k-1}(j)}{\sum_{j\sim i} a_{j,i} Z_{k-1}(j)}\left(1 - \frac{a_{i,j} Z_{k-1}(j)}{\sum_{j\sim i} a_{j,i} Z_{k-1}(j)}\right)\right) \qquad (49) \\ &\quad \le \mathbb{E}\left(\sum_{k=1}^{\infty} \frac{\mathbb{1}_{\{X_{k-1}=i, X_k=j\}}}{Z_{k-1}(j)^2}\right) < \infty\end{aligned}$$

so that, by the Doob convergence theorem in $\mathcal{L}^2$, $M_n^{i,j}$ converges a.s.

Hence, for all $i \in \partial S$,

$$\begin{aligned}\log Z_n(i) &\equiv \sum_{k=1}^{n} \frac{\mathbb{1}_{\{X_k=i\}}}{Z_{k-1}(i)} = \sum_{j\sim i} Y_n^{j,i} \equiv \sum_{j\sim i} a_{j,i} Y_n^j \\ &= \sum_{j\sim i} a_{j,i} \sum_{k=1}^{n} \frac{\mathbb{1}_{\{X_{k-1}=j\}}}{Z_{k-1}(j)} \frac{x(k-1)_j}{N_j(x(k-1))} \\ &\equiv \sum_{j\sim i, j\notin\partial S} a_{i,j} \frac{x(\infty)_j}{N_j(x(\infty))} \sum_{k=1}^{n} \frac{\mathbb{1}_{\{X_{k-1}=j\}}}{Z_{k-1}(j)} \equiv \frac{N_i(x(\infty))}{H(x(\infty))} \log n,\end{aligned}$$

using Lemma 8, the symmetry of a and $N_j(x(\infty)) \neq 0$ for all $j \in G = T(x)$ in the third equivalence, and $H(x(\infty)) = N_j(x(\infty))$ for all $j \in S$ in the fourth equivalence [$x(\infty)$ being an equilibrium].

5.4. *Proof of Proposition 4*. We will compare the probability of arbitrary paths remaining in $T(x)$ for the VRRWs defined, respectively, on the graphs $T(x)$ and G. Let $x(n)$ [and its limit $x(\infty)$] denote the vector of occupation density defined in the Introduction, on the (finite) subgraph $T(x)$.

Let us introduce some notation. For all $k \in \mathbb{N}$ and $A \subseteq G$, let $\mathcal{P}^A := A^{\mathbb{N}}$ be the set of infinite sequences taking values in A, and let $\mathcal{T}_k^A$ be the smallest σ-field on $\mathcal{P}^A$ that contains the cylinders

$$\mathcal{C}_{v,k}^A := \{w \in \mathcal{P}^A \text{ s.t. } w_0 = v_0, \dots, w_k = v_k\}, \qquad v \in A^k.$$

Let $\mathcal{T}^A := \bigvee_{k \in \mathbb{N}} \mathcal{T}_k^A$. Finally, let $(X_j^A)_{j \geq n}$ be the VRRW on A after time n, conditionally to $X_n \in A$ (and be constant equal to X_n otherwise).

For all $k \geq n$ and $v \in T(x)^k$,

$$\mathbb{P}\bigl((X_{n+1}, \dots, X_k) = v | \mathcal{F}_n\bigr) = \mathbb{P}\bigl((X_{n+1}^{T(x)}, \dots, X_k^{T(x)}) = v | \mathcal{F}_n\bigr) Y_{n,k}^{(v)},$$

where

$$Y_{n,k} := \prod_{j=n}^{k-1} \prod_{\alpha \in \partial S(x)} \left(1 - \mathbb{1}_{\{X_j = \alpha\}} \frac{\sum_{\gamma \sim \alpha, \gamma \in G \setminus T(x)} a_{\alpha,\gamma} Z_n(\gamma)}{\sum_{\beta \sim \alpha} a_{\alpha,\beta} Z_j(\beta)}\right) \in (0, 1), \tag{50}$$

and $Y_{n,k}^{(v)}$ denotes the value of $Y_{n,k}$ at $(X_{n+1}, \dots, X_k) := v$, where $Z_j(w)$, $w \in G$, $n \leq j \leq k-1$, assumes the corresponding number of visits of $X.$ to w.

We easily deduce that, for all $E \in \mathcal{T}^{T(x)}$,

$$\mathbb{P}\bigl((X_{j+n})_{j \in \mathbb{N}} \in E | \mathcal{F}_n\bigr) = \mathbb{E}\bigl(\mathbb{1}_{(X_{j+n}^{T(x)})_{j \in \mathbb{N}} \in E} Y_{n,\infty} | \mathcal{F}_n\bigr).$$

Let us now apply this equality with $E := \{\mathcal{R}_{n,\infty} = T(x)\} \cap \mathcal{L}(B_{V_x}(\varepsilon)) \cap \mathcal{A}_\partial(x(\infty))$ and prove that, a.s. on E, $Y_{n,\infty} > 0$, which will complete the proof of the proposition: for all $\alpha \in \partial S(x)$, a.s. on E, if ε is sufficiently small, then

$$\begin{aligned}
\sum_{j=k}^{\infty} \frac{\mathbb{1}_{\{X_j = \alpha\}}}{\sum_{\beta \sim \alpha} a_{\alpha,\beta} Z_j(\beta)} &= \sum_{j=k}^{\infty} \frac{Z_j(\alpha) - Z_{j-1}(\alpha)}{\sum_{\beta \sim \alpha} a_{\alpha,\beta} Z_j(\beta)} \\
&\leq \sum_{j=k}^{\infty} Z_j(\alpha) \left(\frac{1}{\sum_{\beta \sim \alpha} a_{\alpha,\beta} Z_j(\beta)} - \frac{1}{\sum_{\beta \sim \alpha} a_{\alpha,\beta} Z_{j+1}(\beta)}\right) \\
&\leq \overline{a} \sum_{j=k}^{\infty} \frac{Z_j(\alpha)}{(\sum_{\beta \sim \alpha} a_{\alpha,\beta} Z_j(\beta))^2} \mathbb{1}_{\{X_{j+1} \sim \alpha\}} \\
&\leq \overline{a} \sum_{j=k}^{\infty} \frac{x_j(\alpha)}{j (N_\alpha(x(j)))^2} < \infty,
\end{aligned}$$

where we use that, since $\mathcal{A}_\partial(x(\infty))$ holds, $x(j)_\alpha \sim_{j \to \infty} C j^{N_\alpha(x(\infty))/H(x) - 1}$ for some random $C > 0$, so that $\frac{x(j)_\alpha}{j (N_\alpha(x(j))^2} \sim_{j \to \infty} C \frac{j^{N_\alpha(x(\infty))/H(x) - 2}}{N_\alpha(x(\infty))}$, and $N_\alpha(x(\infty)) < H(x(\infty)) = H(x)$ is ε is sufficiently small.

APPENDIX

A.1. Remainder of square-bounded martingales. The following lemma provides an almost sure estimate of $M_n - M_\infty$ for large n, when M_n is a martingale bounded in $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

LEMMA A.1. *Let $(M_n)_{n\geq 0}$ be a bounded martingale in L^2, and let $f:\mathbb{R}_+ \to \mathbb{R}_+$ be a nondecreasing function such that $\int_0^1 (f(x))^{-2}\,dx < \infty$. Then*

$$M_n - M_\infty = o\big(f\big(\mathbb{E}((M_n - M_\infty)^2)\big)\big) \qquad a.s.$$

PROOF. For all $n \geq 0$, let $s_n := \mathbb{E}((M_n - M_\infty)^2)$ and let

$$N_n := \sum_{k=1}^{n} \frac{M_k - M_{k-1}}{f(s_{k-1})}, \qquad N_0 := 0.$$

Then, for all $n \geq 0$,

$$\mathbb{E}[N_n^2] = \sum_{k=1}^{n} \frac{s_{k-1} - s_k}{f(s_{k-1})^2} \leq \int_0^{s_0} \frac{dx}{(f(x))^2} < \infty.$$

Therefore, $(M_n)_{n\geq 0}$ and $(N_n)_{n\geq 0}$ are martingales bounded in L^2, and thus converge a.s.

Now, letting $O_n := N_n - N_\infty$ for all $n \geq 0$,

$$\begin{aligned} M_n - M_\infty &= \sum_{k=n}^{\infty} f(s_k)(O_k - O_{k+1}) = f(s_n)O_n + \sum_{k=n+1}^{\infty} \big(f(s_k) - f(s_{k-1})\big)O_k \\ &= o(f(s_n)) \text{ a.s.} \end{aligned}$$

$\square$

A.2. Proof of Proposition 1. Assume $X_0 := 0$ for simplicity. Let, for all $n \in \mathbb{N}$,

$$\begin{aligned} &A_n := Z_n(-1) + Z_n(1), && \alpha_n^{\pm} := Z_n(\pm 1)/A_n, \\ &R_n := Z_n(0)/A_n - \log A_n, && S_n := \log\left(\frac{Z_n(-1)}{Z_n(1)}\right) = \log\left(\frac{\alpha_n^-}{1 - \alpha_n^-}\right). \end{aligned}$$

Let $a \in (0,1)$, $\varepsilon < [a \wedge (1-a)]/2$. Given $n_0 \in \mathbb{N}$ with $Z_{n_0}(0)$ sufficiently large and $X_{n_0} = 0$, assume that $Z_{n_0}(-2)/\log Z_{n_0}(-1)$, $Z_{n_0}(2)/\log Z_{n_0}(1) \in (1/3, 1/2)$, $Z_{n_0}(\pm 3) \leq \mathsf{Cst}$, $\alpha_{n_0}^- \in (a - \varepsilon/3, a + \varepsilon/3)$ and $R_{n_0} \in (-\varepsilon/3, \varepsilon/3)$, which trivially occurs with positive probability.

Let us define the following stopping times:

$$\begin{aligned} T_0 &:= \inf\{n \geq n_0 \text{ s.t. } X_n \in \{-3, 3\} \text{ or } X_n = X_{n-2} \in \{-2, 2\}\}, \\ T_1 &:= \inf\{n \geq n_0 \text{ s.t. } Z_n(2) \vee Z_n(-2) > \log Z_n(0)\}, \\ T_2 &:= \inf\{n \geq n_0 \text{ s.t. } \alpha_n^- \notin (a - \varepsilon/2, a + \varepsilon/2) \text{ or } R_n \notin (-\varepsilon/2, \varepsilon/2)\}, \\ T &:= T_0 \wedge T_1 \wedge T_2. \end{aligned}$$

For all $n \in \mathbb{N}$, let t_n be the nth return time to 0, and let $t'_n := t_n \wedge T$.

As long as $n_0 \le t_n < T$, $Z_n(0) = A_n(\log A_n + \Box(\varepsilon/2))$, which implies, for sufficiently large $Z_{n_0}(0)$, $A_{t_n} \le n$ by contradiction, hence, $A_{t_n} \ge n/(\log n + \varepsilon/2)$ and, subsequently, $A_{t_n} \le n/(\log(\frac{n}{\log n + \varepsilon/2}) - \varepsilon/2)$. Therefore, $Z_{t_n}(-1) \in ((a-\varepsilon)n/\log n, (a+\varepsilon)n/\log n)$ and $Z_{t_n}(1) \in ((1-a-\varepsilon)n/\log n, (1-a+\varepsilon)n/\log n)$, if $Z_{n_0}(0) \ge \mathsf{Cst}(a, \varepsilon)$.

We successively upper bound $\mathbb{P}(T_0 < T_1 \wedge T_2|\mathcal{F}_{n_0})$, $\mathbb{P}(T_1 < T_0 \wedge T_2|\mathcal{F}_{n_0})$ and $\mathbb{P}(T_2 < T_0 \wedge T_1|\mathcal{F}_{n_0})$, which will enable us to conclude that $\mathbb{P}(T = \infty|\mathcal{F}_{n_0}) > 0$ for large $Z_{n_0}(0)$.

First, for sufficiently large $Z_{n_0}(0)$,

$$\begin{aligned}
&\mathbb{P}(T_0 < T_1 \wedge T_2|\mathcal{F}_{n_0})\\
&\qquad \le \sum_{n \ge Z_{n_0}(0): t_n < T} \mathbb{P}(X_{t_n+2} = X_{t_n+3} \mp 1 = X_{t_n+4} = \pm 2|\mathcal{F}_{n_0})\\
&\qquad\qquad + \mathbb{P}(X_{t_n+3} = \pm 3|\mathcal{F}_{n_0}) \qquad (51)\\
&\qquad \le \mathsf{Cst}(a, \varepsilon) \sum_{n \ge Z_{n_0}(0)} \left[\frac{1}{\log n}\left(\frac{\log n}{n}\right)^2 + \frac{1}{\log n}\frac{\log n}{n}\frac{Z_{n_0}(3) + Z_{n_0}(-3)}{n/\log n}\right]\\
&\qquad \le \mathsf{Cst}(a, \varepsilon) \sum_{n \ge Z_{n_0}(0)} \frac{\log n}{n^2} < \frac{1}{3}.
\end{aligned}$$

Let $\mathbb{G} := (\mathcal{F}_{t'_n})_{n \ge Z_{n_0}(0)}$, and let us consider the Doob decompositions of the $\mathbb{G}$-adapted processes $R_{t'_n}$ and $S_{t'_n}$, $n \ge Z_{n_0}(0)$:

$$R_{t'_n} = R_{n_0} + \Delta_n + \Psi_n, \qquad S_{t'_n} := S_{n_0} + \Phi_n + \Xi_n,$$

where $\Delta_{Z_{n_0}(0)} = \Phi_{Z_{n_0}(0)} = \Psi_{Z_{n_0}(0)} = \Xi_{Z_{n_0}(0)} := 0$ and, for all $n > Z_{n_0}(0)$,

$$\Delta_n - \Delta_{n-1} := \mathbb{E}(R_{t'_n} - R_{t'_{n-1}}|\mathcal{F}_{t'_{n-1}}), \qquad \Phi_n - \Phi_{n-1} := \mathbb{E}(S_{t'_n} - S_{t'_{n-1}}|\mathcal{F}_{t'_{n-1}}),$$

and $(\Psi_n)_{n \ge Z_{n_0}(0)}$ and $(\Xi_n)_{n \ge Z_{n_0}(0)}$ are $\mathbb{G}$-adapted martingales.

Let us now estimate the expectation and variance of the increments of the processes $(R_{t'_n})_{n \in \mathbb{N}}$: if $n \ge \mathsf{Cst}(\varepsilon)$,

$$\begin{aligned}
\mathbb{E}(R_{t'_n+1} - R_{t'_n}|\mathcal{F}_{t'_n}) &= \frac{1}{A_{t'_n}} + \frac{A_{t'_n}}{A_{t'_n} + n}\left(-\frac{n+1}{A_{t'_n}(A_{t'_n}+1)} - \frac{1}{A_{t'_n}} + \Box\left(\frac{1}{(A_{t'_n})^2}\right)\right)\\
&= \frac{1}{A_{t'_n}} + \frac{A_{t'_n}}{A_{t'_n} + n}\left(-\frac{n}{A_{t'_n}(A_{t'_n}+1)} - \frac{1}{A_{t'_n}}\right) + \Box\left(\mathsf{Cst}\frac{\log n}{n^2}\right)\\
&= -\frac{n}{A_{t'_n}(A_{t'_n}+1)(A_{t'_n}+n)} + \Box\left(\mathsf{Cst}\frac{\log n}{n^2}\right)\\
&= \Box\left(\mathsf{Cst}\frac{(\log n)^3}{n^2}\right),
\end{aligned}$$

$$\mathbb{E}(R_{t'_{n+1}} - R_{t'_n+1}|\mathcal{F}_{t'_n}) = \square\biggl(\mathsf{Cst}\frac{\log n}{n}\biggl(\frac{1}{A_{t'_n}} + \frac{n+1}{(A_{t'_n}+1)(A_{t'_n}+2)}\biggr)\biggr)$$
$$= \square\biggl(\mathsf{Cst}\frac{(\log n)^3}{n^2}\biggr)$$

and

$$|R_{t'_{n+1}} - R_{t'_n}| \le \frac{2}{A_{t'_n}} + \frac{2(n+1)}{A_{t'_n}(A_{t'_n}+1)} \le \mathsf{Cst}\frac{(\log n)^2}{n}$$

so that, in summary,

$$(52)\qquad |\Delta_n - \Delta_{n-1}| \le \mathsf{Cst}\frac{(\log n)^3}{n^2}, \qquad \mathbb{E}\bigl((\Psi_{n+1}-\Psi_n)^2|\mathcal{F}_{t'_n}\bigr) \le \mathsf{Cst}\frac{(\log n)^4}{n^2}.$$

Let us do similar computations for $(S_{t'_n})_{n\in\mathbb{N}}$: if $n \ge \mathsf{Cst}(a,\varepsilon)$,

$$\Phi_n - \Phi_{n-1} = \biggl(\frac{1}{Z_{t_n}(-1)} + \frac{1}{Z_{t_n}(-1)}\square\biggl(\frac{\log n}{n}\biggr) + \square\biggl(\frac{1}{(Z_{t_n}(-1))^2}\biggr)\biggr)$$
$$\times \frac{Z_{t_n}(-1)}{n + Z_{t_n}(-1) + Z_{t_n}(1)}$$
$$- \biggl(\frac{1}{Z_{t_n}(1)} + \frac{1}{Z_{t_n}(1)}\square\biggl(\frac{\log n}{n}\biggr) + \square\biggl(\frac{1}{(Z_{t_n}(1))^2}\biggr)\biggr)$$
$$\times \frac{Z_{t_n}(1)}{n + Z_{t_n}(-1) + Z_{t_n}(1)}$$
$$= \frac{\log n}{n^2}\square(\mathsf{Cst}(a,\varepsilon)),$$

and

$$|S_{t'_{n+1}} - S_{t'_n}| \le \log\biggl(1 + \frac{2}{Z_{t_n}(1)}\biggr) \vee \log\biggl(1 + \frac{2}{Z_{t_n}(-1)}\biggr),$$

so that

$$(53)\qquad |\Phi_n - \Phi_{n-1}| \le \mathsf{Cst}(a,\varepsilon)\frac{\log n}{n^2},$$
$$\mathbb{E}\bigl((\Xi_{n+1}-\Xi_n)^2|\mathcal{F}_{t'_n}\bigr) \le \mathsf{Cst}(a,\varepsilon)\biggl(\frac{\log n}{n}\biggr)^2.$$

Hence, by Chebyshev's and Doob's martingale inequalities, for all $\delta > 0$,

$$\mathbb{P}\Bigl(\max_{k\ge Z_{n_0}(0)} |\Psi_k| > \delta|\mathcal{F}_{n_0}\Bigr) \le \frac{\mathsf{Cst}}{\delta^2}\sum_{j=Z_{n_0}(0)}^{\infty}\frac{(\log n)^4}{n^2} \le \frac{\mathsf{Cst}}{\delta^2}\frac{(\log Z_{n_0}(0))^4}{Z_{n_0}(0)}$$

and a similar inequality holds on the maximum of $|\Xi_k|$, $k \ge Z_{n_0}(0)$, so that, for sufficiently large $Z_{n_0}(0)$, $\mathbb{P}(T_2 < T_0 \wedge T_1|\mathcal{F}_{n_0}) < 1/3$.

Let us now make use of notation $Y_n^{i,j}$, Y_n^i and $M_n^{i,j}$ from Section 5.3 (with $a_{i,j} = \mathbb{1}_{i\sim j}$), and let $U_n^{\pm} := Y_n^{\pm 1,\pm 2}$, $V_n^{\pm} := Y_n^{\pm 1}$ and $W_n^{\pm} := M_n^{\pm 1,\pm 2} = U_n^{\pm} - V_n^{\pm}$. Then the processes $(U_n^{\pm})_{n\ge 0}$ are martingales and, using (49), for all $n \ge n_0$,

$$\mathbb{E}\big((W_n^{\pm} - W_{n_0}^{\pm})^2 | \mathcal{F}_{n_0}\big) \le \mathbb{E}\left(\sum_{k=n_0+1}^{n} \frac{\mathbb{1}_{\{X_{k-1}=\pm 1, X_k=\pm 2\}}}{Z_{k-1}(\pm 2)^2} \Big| \mathcal{F}_{n_0}\right) \le \sum_{j\ge Z_{n_0}(\pm 2)} \frac{1}{j^2} \tag{54}$$

so that, if $\Upsilon := \{\max_{k\ge n_0} |W_k^i - W_{n_0}^i| \le \delta, i \in \{+,-\}\}$, then, for all $\delta > 0$,

$$\mathbb{P}(\Upsilon^c | \mathcal{F}_{n_0}) \le \frac{1}{\delta^2}\left(\frac{1}{Z_{n_0}(2) - 1} + \frac{1}{Z_{n_0}(-2) - 1}\right) < \frac{1}{3}$$

for sufficiently large $Z_{n_0}(0)$.

Now, on Υ, for all $n < T$, choosing $\delta = (\log 2)/3$, and again for sufficiently large $Z_{n_0}(0)$,

$$\begin{aligned}
\log Z_n(\pm 2) &\le \log Z_{n_0}(\pm 2) + U_n^{\pm} - U_{n_0}^{\pm} + \delta \le 2\delta + \log Z_{n_0}(\pm 2) + V_n^{\pm} - V_{n_0}^{\pm} \\
&\le 2\delta + \log Z_{n_0}(\pm 2) + \sum_{k=n_0+1}^{n} \frac{\mathbb{1}_{\{X_{k-1}=\pm 1\}}}{Z_{k-1}(0)} \\
&\le 2\delta + \log Z_{n_0}(\pm 2) + \sum_{k=Z_{n_0}(\pm 1)}^{Z_{n-1}(\pm 1)} \frac{1}{k \log k} \\
&\le 3\delta + \log\left(\frac{Z_{n_0}(\pm 2)}{\log Z_{n_0}(\pm 1)}\right) + \log(\log Z_n(\pm 1)) \\
&\le \log(\log Z_n(\pm 1)) \le \log(\log Z_n(0)),
\end{aligned}$$

where we use in the fourth inequality that, if $n < T$, then $T_n \ge -\varepsilon/2$ and $\alpha_n^- \in (a - \varepsilon/2, a + \varepsilon/2)$ so that $Z_n(0) \ge Z_n(\pm 1)\log Z_n(\pm 1)$ if $Z_{n_0} \ge \mathsf{Cst}(a, \varepsilon)$, and in the sixth inequality that $Z_{n_0}(\pm 2)/\log Z_{n_0}(\pm 1) \le 1/2$. This completes the proof, as $\mathbb{P}(T_1 < T_0 \wedge T_2 | \mathcal{F}_{n_0}) \le \mathbb{P}(\Upsilon^c | \mathcal{F}_{n_0}) < 1/3$ for large $Z_{n_0}(0)$.

The estimates (52)–(53) [resp., (54)] imply that the $\mathbb{G}$ (resp., $\mathbb{F}$)-adapted martingales $(\Psi_n)_{n\ge Z_{n_0}(0)}$ and $(\Xi_n)_{n\ge Z_{n_0}(0)}$ (resp., $W_n^{\pm}$) are bounded in L^2 and hence converge a.s.

Therefore, on $\{T = \infty\}$, (i)–(ii) hold, and $(\alpha_n)_{n\ge 0}$ and $(R_n)_{n\ge 0}$ converge a.s. Note that Lemma A.1 implies more precisely, for all $\nu < 1/2$, $\Xi_n - \Xi_\infty = o(n^{-\nu})$, hence, $\alpha_n - \alpha_\infty = o(Z_n(0)^{-\nu})$. Thus, on $\{T = \infty\}$,

$$\begin{aligned}
\log Z_n(\pm 2) \equiv U_n^{\pm} \equiv V_n^{\pm} &= \sum_{k=0}^{n-1} \frac{\mathbb{1}_{\{X_k=\pm 1\}}}{Z_k(\pm 2) + Z_k(0)} \\
&\equiv \alpha_\infty^{\pm} \sum_{k=0}^{n-1} \frac{\mathbb{1}_{\{X_k=\pm 1\}}}{Z_k(\pm 1)\log Z_k(\pm 1)}\left(1 + O\left(\frac{1}{\log Z_k(\pm 1)}\right)\right) \\
&\equiv \alpha_\infty^{\pm} \log(\log Z_n(\pm 1)) \equiv \alpha_\infty^{\pm} \log(\log n),
\end{aligned}$$

which proves (iii).

Acknowledgments. We would like to thank the referees for very helpful comments.

REFERENCES

[1] ATHREYA, K. (1967). Limit theorems for multitype continuous time Markov branching processes and some classical urn schemes. Ph.D. dissertation. Stanford Univ. MR2616252

[2] BENAÏM, M. (1997). Vertex-reinforced random walks and a conjecture of Pemantle. *Ann. Probab.* **25** 361–392. MR1428513

[3] BROOM, M., CANNINGS, C. and VICKERS, G. T. (1993). On the number of local maxima of a constrained quadratic form. *Proc. Roy. Soc. London Ser. A* **443** 573–584. MR1252602

[4] COPPERSMITH, D. and DIACONIS, P. (1986). Random walks with reinforcement. Unpublished manuscript.

[5] DAVIS, B. (1999). Reinforced and perturbed random walks. In *Random Walks* (*Budapest*, 1998). *Bolyai Soc. Math. Stud.* **9** 113–126. János Bolyai Math. Soc., Budapest. MR1752892

[6] JORDAN, J. (2008). Degree sequences of geometric preferential duplication graphs. Preprint.

[7] LIMIC, V. and TARRÈS, P. (2007). Attracting edge and strongly edge reinforced walks. *Ann. Probab.* **35** 1783–1806. MR2349575

[8] LIMIC, V. and VOLKOV, S. (2010). VRRW on complete-like graphs: Almost sure behavior. *Ann. Appl. Probab.* **20** 2346–2388. MR2759737

[9] LOSERT, V. and AKIN, E. (1983). Dynamics of games and genes: Discrete versus continuous time. *J. Math. Biol.* **17** 241–251. MR0714271

[10] MERKL, F. and ROLLES, S. W. W. (2006). Linearly edge-reinforced random walks. In *Dynamics & Stochastics*. *IMS Lecture Notes Monogr. Ser.* **48** 66–77. IMS, Beachwood, OH. MR2306189

[11] MÉTIVIER, M. and PRIOURET, P. (1987). Théorèmes de convergence presque sure pour une classe d'algorithmes stochastiques à pas décroissant. *Probab. Theory Related Fields* **74** 403–428. MR0873887

[12] MOUNTFORD, T. and TARRÈS, P. (2008). An asymptotic result for Brownian polymers. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 29–46. MR2451570

[13] PEMANTLE, R. (1988). Random processes with reinforcement. Ph.D. dissertation, Massachussets Institute of Technology.

[14] PEMANTLE, R. (1992). Vertex-reinforced random walk. *Probab. Theory Related Fields* **92** 117–136. MR1156453

[15] PEMANTLE, R. (2007). A survey of random processes with reinforcement. *Probab. Surv.* **4** 1–79 (electronic). MR2282181

[16] PEMANTLE, R. and VOLKOV, S. (1999). Vertex-reinforced random walk on **Z** has finite range. *Ann. Probab.* **27** 1368–1388. MR1733153

[17] SALOFF-COSTE, L. (1997). Lectures on finite Markov chains. In *Lectures on Probability Theory and Statistics* (*Saint-Flour*, 1996). *Lecture Notes in Math.* **1665** 301–413. Springer, Berlin. MR1490046

[18] TARRÈS, P. (2004). Vertex-reinforced random walk on $\mathbb{Z}$ eventually gets stuck on five points. *Ann. Probab.* **32** 2650–2701. MR2078554

[19] TÓTH, B. (1999). Self-interacting random motions—a survey. In *Random Walks* (*Budapest*, 1998). *Bolyai Soc. Math. Stud.* **9** 349–384. János Bolyai Math. Soc., Budapest. MR1752900

[20] TYRER, J. P., CANNINGS, C. and VICKERS, G. T. (1993). Routes to polymorphism. *J. Theor. Biol.* **165** 213–223.

[21] VICKERS, G. T. and CANNINGS, C. (1988). On the number of stable equilibria in a one-locus, multi-allelic system. *J. Theoret. Biol.* **131** 273–277. MR0936311

[22] VICKERS, G. T. and CANNINGS, C. (1988). Patterns of ESSs. I, II. *J. Theoret. Biol.* **132** 387–408, 409–420. MR0949814

[23] VOLKOV, S. (2001). Vertex-reinforced random walk on arbitrary graphs. *Ann. Probab.* **29** 66–91. MR1825142

[24] VOLKOV, S. (2006). Phase transition in vertex-reinforced random walks on $\mathbb{Z}$ with non-linear reinforcement. *J. Theoret. Probab.* **19** 691–700. MR2280515

INSTITUT DE MATHÉMATIQUES
UNIVERSITÉ DE NEUCHÂTEL
RUE EMILE-ARGAND 11, CASE POSTALE 2
2007 NEUCHÂTEL
SWITZERLAND
E-MAIL: michel.benaim@unine.ch

INSTITUT DE MATHÉMATIQUES
CNRS, UNIVERSITÉ DE TOULOUSE
118 ROUTE DE NARBONNE
31062 TOULOUSE CEDEX 9
FRANCE
E-MAIL: tarres@math.univ-toulouse.fr