The Annals of Probability
1997, Vol. 25, No. 1, 361–392

VERTEX-REINFORCED RANDOM WALKS AND A CONJECTURE OF PEMANTLE

BY MICHEL BENAÏM

Université Paul Sabatier

We discuss and disprove a conjecture of Pemantle concerning vertex-reinforced random walks.

The setting is a general theory of non-Markovian discrete-time random processes on a finite space $E = \{1, \dots, d\}$, for which the transition probabilities at each step are influenced by the proportion of times each state has been visited. It is shown that, under mild conditions, the asymptotic behavior of the empirical occupation measure of the process is precisely related to the asymptotic behavior of some deterministic dynamical system induced by a vector field on the $d - 1$ unit simplex. In particular, any minimal attractor of this vector field has a positive probability to be the limit set of the sequence of empirical occupation measures. These properties are used to disprove a conjecture and to extend some results due to Pemantle. Some applications to edge-reinforced random walks are also considered.

1. Introduction. This paper considers a class of random processes with reinforcement introduced by Pemantle (1992) as well as some generalizations.

Let $\{Y_n\}_{n \geq 0}$ be a discrete time stochastic process living in a finite state space $E = \{1, \dots, d\}$ representing the vertices of a graph. Initially each edge is given a weight but each time Y visits a vertex all the weights of the edges leading to this vertex are increased by a positive amount. At the next step an edge leading out from the current vertex is chosen with a probability proportional to its weight.

Formally, the process is defined as follows. Let R be a real $d \times d$ matrix with positive entries. Let $S_i(0)$ denote a positive integer which represents the initial importance given to vertex i and let $S_i(n) - S_i(0)$ denote the number of times Y has occupied the state i between times 1 and n. $S_i(0)$ can be seen as the number of times Y has occupied vertex i before the initial time $t = 0$. That is,

$$S_i(n) = S_i(0) + \sum_{k=1}^{n} \delta_{Y_k, i}.$$

Suppose that

$$P(Y_{n+1} = j | F_n) = \frac{R_{Y_n, j} S_j(n)}{\sum_{k=1}^{d} R_{Y_n, k} S_k(n)}$$

Received October 1995; revised June 1996.

AMS 1991 *subject classifications.* Primary 60J10; secondary 34F05, 34C35.

Key words and phrases. Reinforced random walks, random perturbations of dynamical systems, chain recurrence, attractors.

where F_n denotes the σ-field generated by $\{Y_j: 0 \le j \le n\}$. Set $n_0 = \sum_{i=1}^d S_i(0)$ and let

$$v(n) = \frac{S(n)}{n + n_0}$$

denote the empirical occupation measure.

Let $\Delta \subset \mathbf{R}^d$ denote the $d-1$ unit simplex. For $v \in \Delta$, let $M(v)$ denote the Markov transition matrix defined by

$$M_{i,j}(v) = \frac{R_{i,j} v_j}{\sum_k R_{i,k} v_k}, \qquad i, j = 1, \dots, d$$

and let $\pi(v)$ denote the invariant probability for $M(v)$. Pemantle (1992) proved the following theorem.

THEOREM 1.1 (Pemantle, 1992). *Suppose that the matrix R is symmetric. Then the following hold.*

(i) *The sequence $\{v(n)\}_{n \ge 0}$ converges, with probability* 1, *toward the critical set*

$$C = \{v: v_i \ge 0, v = \pi(v)\}.$$

(ii) *In the (generic) situation where C is finite $\lim_{n\to\infty} v(n)$ exists almost surely.*

He also proposed the two following conjectures.

CONJECTURE 1.2 (Pemantle, 1992). *Part* (ii) *of Theorem* 1.1 *holds without any nondegeneracy assumption.*

CONJECTURE 1.3 (Pemantle, 1992). *Part* (i) *of Theorem* 1.1 *holds whether or not R is symmetric.*

The principal and initial motivation of this paper was to disprove Conjecture 1.3. To achieve this goal we will introduce a general class of processes with reinforcement and show how the asymptotic behavior of these processes can be precisely related to the asymptotic behavior of some vector field on the unit $(d-1)$-dimensional simplex.

The organization of the paper is as follows: Section 2 introduces the general class of Vertex-reinforced random walks (VRRW) to be considered. Section 3 reviews earlier results on which the paper is based and states the main result: with probability 1, the limit set of the sequence of empirical occupation measures is a continuum (i.e., a compact connected set) L which is invariant under the flow Φ of some vector field defined on the unit $(d-1)$-dimensional simplex and such that the restricted flow $\Phi|L$ enjoys a technical property, akin to ergodicity, called chain recurrence. As an application of this result, we give in Section 4 a short proof of Pemantle's main result and we discuss the behavior of a class of edge-reinforced processes inspired by the work of Coppersmith and Diaconis (1986). Section 6 gives some conditions which ensure that a given attractor of the vector field has a positive probability to be the limit set of the sequence $\{v(n)\}$. Section 7 makes precise the behavior of $\{v_n\}$ near an attractor and states some shadowing properties which are used to give a partial answer to Conjecture 1.2. All these results are used in Section 9 to construct a simple counterexample to Pemantle's conjecture. Finally, Sections 5 and 8 contain proofs of the more technical results.

2. Generalized VRRW.

Notation. Throughout the paper we use the following notation: the *unit* $(d-1)$ *simplex* $\Delta \subset \mathbf{R}^d$ is the set

$$\Delta = \left\{ v \in \mathbf{R}^d\colon v_i \geq 0,\ \sum_i v_i = 1 \right\}.$$

The *affine hull* of Δ is the set

$$\operatorname{aff}(\Delta) = \left\{ v \in \mathbf{R}^d\colon \sum_i v_i = 1 \right\}$$

and its *tangent space* is

$$T\Delta = \left\{ v \in \mathbf{R}^d\colon \sum_i v_i = 0 \right\}.$$

The *boundary* of Δ is the set

$$\partial\Delta = \left\{ v \in \Delta\colon \prod_i v_i = 0 \right\}$$

and its ith *face* is

$$\partial_i \Delta = \{ v \in \partial\Delta\colon v_i = 0 \}.$$

Let M be a $d \times d$ real matrix. M is called a *transition* matrix if it has nonnegative entries and satisfies the normalization condition

$$\sum_j M_{i,j} = 1, \qquad i = 1, \dots, d.$$

Let $M_d(\mathbf{R})$ denotes the space of real $d \times d$ matrices. We let $M \subset M_d(\mathbf{R})$ denote the set of $d \times d$ transition matrices and we set

$$TM = \left\{ M \in M_d(\mathbf{R})\colon \forall\ i,\ \sum_j M_{i,j} = 0 \right\}.$$

We let $\|\cdot\|$ denote the L^1 norm on $\mathbf{R}^d$ defined by $\|v\| = \sum_{i=1}^d |v_i|$. We may denote by d the induced distance.

Definition. This paper considers a generalized version of Pemantle's vertex-reinforced random walk which is defined as follows.

Let

$$M: \Delta \to M$$
$$v \to M(v) = \{M_{i,j}(v)\}$$

be a C^k ($k \geq 1$) map. Given a point $v \in \mathrm{Int}(\Delta)$, a vertex $y \in E$ and a positive integer $n_0 \in \mathbf{N}$, consider a stochastic process $\{Y_n, (S_1(n), \ldots, S_d(n))\}_{n \geq 0}$ defined on $E \times \mathbf{R}_+^E$ by

$$S_i(0) = n_0 v_i, \qquad Y_0 = y;$$
$$S_i(n) = S_i(0) + \sum_{k=1}^{n} \delta_{Y_k, i}, \qquad n \geq 0;$$
$$P(Y_{n+1} = j | F_n) = M_{Y_n, j}(v(n)),$$

where F_n denotes the σ-field generated by $\{Y_j: 0 \leq j \leq n\}$ and $v(n) = S(n)/n + n_0$. Hereafter the parameters $v(0) = v$, $Y_0 = y$ and n_0 will be referred to as the *initial condition*, the *initial state* and the *initial mass*, respectively.

EXAMPLES. *Urn processes*: Let $\pi: \Delta \to \Delta$ be a C^k map and suppose $M_{i,j}(v) = \pi_j(v)$. In this case, the VRRW associated to M is a generalized Pólya urn process as it has been considered by Hill, Lane and Sudderth (1980), Arthur, Ermol'ev and Kaniovskii, (1983), Pemantle (1990) and Benaïm and Hirsch (1995), among others.

Edge-reinforced random walks: Let $G \subset E \times E$ be an oriented graph (i.e., a set of oriented edges) with E as vertices set. Let

$$\Delta(G) = \left\{ W \in M_d(\mathbf{R}): W_{i,j} \geq 0; (i,j) \notin G \Rightarrow W_{i,j} = 0; \sum_{i,j} W_{i,j} = 1 \right\}$$

and let $\hat{M}: \Delta(G) \to M$ be a map. Consider the process $Y = \{Y_n\}$ on E defined by

$$P(Y_{n+1} = j | F_n) = \hat{M}_{Y_n, j}(W(n)),$$

where $W(n) \in \Delta(G)$ is such that $W_{i,j}(n)$ represents the proportion of time Y has moved from site i to site j between times 1 and n. Here, the probability transitions are influenced by the moves along each edge of the graph.

We call such a process a *generalized edge-reinforced random walk*. The original idea is due to Coppersmith and Diaconis (1986) who studied thoroughly the case of nonoriented graphs where each edge is given initial weight 1 and each time Y travels over an edge, 1 is added to the edge-weight.

Let $Z_n = (Y_{n-1}, Y_n) \in E \times E$. It is readily seen that Z is a generalized vertex-reinforced random walk on $E \times E$. Thus, the results given in this paper can be used to analyze the asymptotic behavior of the sequence $W(n)$. The empirical occupation measure of Y_n is related to $W(n)$ by the relation $v_i(n) = \sum_j W_{j,i}(n)$. A detailed example will be considered in Section 4.

3. Main properties. As usual, a finite homogeneous Markov chain is called *indecomposable* if it has only one recurrent class (either periodic or aperiodic). From now on, we suppose that the following condition holds.

HYPOTHESIS 3.1. *For each $v \in \Delta$, the matrix $M(v)$ is indecomposable.*

By a standard result of Markov chain theory this implies the existence of a unique invariant measure $\pi(v)$ for $M(v)$. Observe that the recurrence class of $M(v)$ may be periodic and may depend on v in a nontrivial way.

Following Pemantle, we define the critical set of M as

$$C = \{ v \in \Delta: \pi(v) = v \}.$$

To the map M we associate the vector field

$$F: \Delta \to T\Delta,$$
$$v \to -v + \pi(v).$$

The main result of this section (Corollary 3.5) shows that the asymptotic behavior of the sequence $\{v_n\}$ can be precisely related to the asymptotic behavior of the flow generated by F regardless of the dynamics of F. We begin by a lemma which records a few elementary properties of F.

LEMMA 3.2. (i) *The vector field F is C^k.*

(ii) *Δ is positively invariant under the dynamics of F.*

(iii) *If for all $j \in E$, $v_j = 0$ implies $M_{i,j}(v) = 0$, then Δ and its faces are globally invariant.*

PROOF. (i) Let $L: \Delta \times \mathrm{aff}(\Delta) \to T\Delta$ be a function defined by

$$L(v, \pi) = M^T(v)\pi - \pi,$$

where $M^T(v)$ denotes the transpose of $M(v)$.

The invariant measure $\pi = \pi(v)$ is the solution to the implicit equation $L(v, \pi) = 0$. Identifying the tangent space of $\mathrm{aff}(\Delta)$ at any point with $T\Delta$, we see that

$$\frac{\partial L}{\partial \pi}(v, \pi(v)) = \left(M^T(v) - \mathrm{Id}\right)|T\Delta.$$

By uniqueness of the invariant measure, the kernel of $M(v)^T - \mathrm{Id}$ is $\mathbf{R} \cdot \pi(v)$ and since $\pi(v) \in \mathrm{aff}(\Delta)$, $\mathbf{R} \cdot \pi(v) \cap T\Delta = \{0\}$ showing that $(\partial L/\partial \pi)(v, \pi(v))$ is invertible. Now, the implicit function theorem shows that $v \to \pi(v)$ is C^k. Thus F is C^k.

(ii) $F(v)$ points into Δ whenever v belongs to the boundary of Δ. Thus any forward trajectory based in Δ remains in Δ.

(iii) If $M_{i,j}(v) = 0$ when $v_j = 0$, the invariant measure $\pi(v)$ satisfies the same property [i.e., $\pi_j(v) = 0$ when $v_j = 0$]. Thus, $F_i(v) = 0$ when $v_i = 0$ and the result follows. □

Being smooth, the vector field F generates a smooth semiflow on Δ,

$$\Phi: \mathbf{R}_+ \times \Delta \to \Delta,$$
$$(t, v) \to \Phi_t(v),$$

such that the solution to the initial value problem $(du/dt) = F(u)$ with initial condition $u(0) = v$ is the curve $t \to \Phi_t(v)$. When Δ is globally invariant, Φ extends to a *flow* on Δ, meaning that the solutions are defined for all $t \in \mathbf{R}$.

Let $X: \mathbf{R}_+ \to \Delta$ be a continuous function. Using the terminology introduced in Benaïm and Hirsch (1996), we say that X is an *asymptotic pseudotrajectory* of Φ if

$$\lim_{t \to \infty} d\big(X(t+T), \Phi_T(X(t))\big) = 0$$

locally uniformly in $T \in \mathbf{R}_+$. Set $\tau_0 = 0$, $\tau_n = \sum_{i=1}^{n}(1/(i + n_0))$, $n \geq 1$ and define the interpolated process $V: \mathbf{R}_+ \to \Delta$ as shown in the following.

1. $V(\tau_n) = v(n)$.
2. V is affine on $[\tau_n, \tau_{n+1}]$.

PROPOSITION 3.3. *V is almost surely an asymptotic pseudotrajectory of the semiflow Φ induced by F.*

For convenience, the proof of this proposition is postponed to Section 5.

The asymptotic behavior of asymptotic pseudotrajectories is related to the dynamics of Φ by the following theorem proved in Benaïm (1996) for stochastic approximation processes and Benaïm and Hirsch (1996) for the more general class of asymptotic pseudotrajectories.

THEOREM 3.4. *Let X be a (precompact) asymptotic pseudotrajectory of Φ. Let*

$$L(X) = \bigcap_{t \geq 0} \overline{X[t, \infty)}$$

denote the limit of X. Then $L(X)$ is compact Φ-invariant connected and internally chain recurrent.

Let $L(\{v_n\}) = \bigcap_{n \geq 0} \overline{\{v_k : k \geq n\}}$ denote the limit set of the sequence $\{v(n)\}$. Clearly, $L(\{v_n\}) = L(V)$. The corollary follows.

COROLLARY 3.5. *$L(\{v_n\})$ is almost surely compact Φ-invariant connected and internally chain recurrent.*

This result, which is the main tool to analyze the behavior of $\{v(n)\}$, requires some explanations: a set $L \subset \Delta$ is invariant if for all $x \in L$ the solution $t \to \Phi_t(x)$ of $(du/dt) = F(u)$, $u(0) = x$ is defined for all $t \in \mathbf{R}$ and remains in L. In this case the semiflow Φ induces a flow on L which we denote $\Phi|L$. Let $L \subset \Delta$ be a compact invariant set and let $\Psi = \Phi|L$. Let $\delta > 0$, $T > 0$. A (δ, T) pseudo-orbit for Ψ from $a \in L$ to $b \in L$ is a finite sequence of partial trajectories

$$\Psi_t(x_i): 0 \leq t \leq t_i; \qquad i = 0, \dots, k-1; \, t_i \geq T,$$

such that

$$x_j \in L, \qquad j = 0, \dots, k-1;$$
$$d(x_0, a) < \delta,$$
$$d\big(\Psi_{t_j}(x_j), x_{j+1}\big) < \delta, \qquad j = 0, \dots, k-1;$$
$$x_k = b.$$

We write $a \rightsquigarrow b$ if for every $\delta > 0$, $T > 0$ there exists a (δ, T)-pseudo-orbit from a to b. If $a \rightsquigarrow a$ then a is a *chain recurrent* point for the flow $\Psi = \Phi|L$. If every point of L is chain recurrent for Ψ, then L is said to be *internally chain recurrent*.

If $a \rightsquigarrow b$ for all $a, b \in L$, we say that L is *internally chain transitive*. When L is connected, this is equivalent to "internally chain recurrent." This is also equivalent to the condition that there are no proper attractors for Ψ (see Section 6 for a definition of attractors). For example, any compact alpha or omega limit set for a flow Φ is internally chain transitive. The omega limit set of x is the set of points p such that $p = \lim_{k \to \infty} \Phi_{t_k}(x)$ for some sequence $t_k \to \infty$. The alpha limit set of x is defined as the omega limit set of x for the reversed flow $\{\Phi_{-t}\}$. The equilibria set of Φ is internally chain recurrent. The nonwandering set of Φ is an example of a set which is not (always) internally chain recurrent but consists of points which are chain recurrent for Φ. x is wandering if there is a neighborhood U of x and a positive time T such that $U \cap \Phi_t(U)$ is empty for all $t > T$. Otherwise x is nonwandering.

For more details and the basic theory of chain recurrence, we refer the reader to Conley (1978) or Bowen (1975).

REMARK. Given a sequence of positive weights $\{\alpha_k\}_{k \geq 0}$, define the weighted sum

$$S_i(n) = S_i(0) + \sum_{k=1}^{n} \alpha_k \delta_{Y_k, i}$$

and the weighted occupation measure

$$v(n) = \frac{S(n)}{n_0 + \sum_{k=1}^{n} \alpha_j}.$$

Then Corollary 3.5 remains valid provided that the average weight $\gamma_k = \alpha_k / (n_0 + \sum_{j=1}^{n} \alpha_j)$ satisfies the conditions: $\sum_k \gamma_k = \infty$ and $\sum_k \gamma_k^{1+\delta} < \infty$ for some $\delta \geq 1$.

4. Some applications.

Pemantle's convergence theorem. To illustrate Corollary 3.5 we will give a short proof of Pemantle's main convergence result in the case where the matrix R has positive entries. Set

$$N_i(v) = \sum_{k=1}^{d} R_{i,k} v_k \tag{1}$$

and suppose the map M is defined by

$$M_{i,j}(v) = \frac{R_{i,j} v_j}{N_i(v)} \tag{2}$$

The matrix R is called *reversible* if there exist positive numbers α_i, $i = 1, \dots, d$ such that for all $(i, j) \in E \times E$ with $i \neq j$:

$$\alpha_i R_{i,j} = \alpha_j R_{j,i}.$$

Suppose R is reversible. Set $h_i(v) = \alpha_i N_i(v)$. Then for $i \neq j$,

$$v_i h_i(v) M_{i,j}(v) = v_j h_j(v) M_{j,i}(v).$$

It follows that the invariant measure $\pi(v)$ of $M(v)$ is given as

$$\pi_i(v) = \frac{v_i h_i(v)}{H(v)}, \qquad i = 1, \dots, d,$$

where $H(v)$ is the normalization number $H(v) = \sum_{i=1}^{d} v_i h_i(v)$. Since $H(v) > 0$ for all $v \in \Delta$, we can multiply the vector field $F = -\mathrm{Id} + \pi$ by H without changing the phase portrait (this only changes the length of the vector field). This leads us to the following differential equation on Δ:

$$\frac{dv_i}{dt} = v_i\big(h_i(v) - H(v)\big). \tag{3}$$

Notice that

$$\frac{\partial H}{\partial v_i} = h_i(v) + \sum_j v_j \frac{\partial h_j(v)}{\partial v_i} = h_i(v) + \sum_j v_j \alpha_j R_{j,i} v_i = 2 h_i(v),$$

where the last equality follows from the reversibility of R and definition of $h_i(v)$. Thus, taking the derivative of H along trajectories of (3) gives

$$\begin{aligned}\frac{dH}{dt} &= \sum_i \frac{\partial H}{\partial v_i} v_i\big(h_i(v) - H\big) = 2\sum_i h_i(v) v_i\big(h_i(v) - H(v)\big)\\ &= 2\sum_i v_i\big(h_i(v) - H(v)\big)^2.\end{aligned}$$

Therefore, we are in the situation where the following holds.

1. H is a continuous function which is strictly increasing along nonstationary orbits (i.e., nonequilibrium points) of (3).

2. The restriction of H to the equilibria set $Eq(F) = \{v \in \Delta: F(v) = 0\} = C$ takes a finite number of values.

In this situation it is not very difficult to prove that any compact internally chain recurrent set consists of equilibria. This follows, for example, from Proposition 3.2 of Benaïm (1996). Then from Corollary 3.5 we deduce the following result which has been proved by Pemantle (1992) for symmetric R.

THEOREM 4.1. *Suppose R is reversible. Then $L(\{v_n\})$ is almost surely a compact connected subset of C.*

A class of edge-reinforced random walks. We describe here a class of edge-reinforced random walks which is directly inspired by the work of Coppersmith and Diaconis (1986) and Diaconis (1988).

Let $G \subset E \times E$ be an oriented graph (i.e., a set of oriented edges) with E as a set of vertices. We suppose that G is symmetric $[(i, j) \in G \Rightarrow (j, i) \in G]$, connected and aperiodic. Initially each edge $(i, j) \in G$ is given a positive weight $A_{i,j} = A_{j,i}$.

Let $0 < \varepsilon < \inf_{i,j} A_{i,j}$. A random walk Y starts at a particular vertex. Each time Y travels from i to j or from j to i, the weight of the edge (i, j) is positively reinforced by $A_{i,j}$ and each time Y travels over another edge, the weight of (i, j) is reinforced by $\varepsilon A_{i,j}$. At the next step, an edge leading out from the current vertex is chosen with a probability proportional to its weight.

The case $\varepsilon = 0$ and $A_{i,j} = 1$ was introduced by Coppersmith and Diaconis (1986) [see also Diaconis (1988)] as a simple model of exploring a new city. Imagine a person visiting this city. At first she explores at random the area where she lives, but as time goes on, routes that have been used more in the past become more familiar and are more likely to be traveled.

The choice of different numbers $A_{i,j}$ means that in this city (as in most cities) some paths are more attractive than others. The parameter $\varepsilon > 0$ models the fact that this person (like me) tends to forget the areas she has not visited for a long time. Indeed, let $A_{i,j}(n)$ denote the weight of (i, j) at time n [with the convention that $A_{i,j}(n) = 0$ if $(i, j) \notin G$] and suppose that the process Y does not visit vertex i between times n and $n + p$. Then, at time $n + p$ the probability of transition from i to j is given by

$$\frac{A_{i,j}(n) + p\varepsilon A_{i,j}}{\sum_k (A_{i,k}(n) + p\varepsilon A_{i,k})}$$

and we see that this last quantity tends to the initial transition probability $A_{i,j}/\sum_k A_{i,k}$ when p tends to infinity.

THEOREM 4.2. *Let $W(n) = \{W_{i,j}(n)\}_{i,j=1,\dots,d}$ denote the empirical occupation measure of oriented edges (i.e., $W_{i,j}(n)$ is the proportion of time Y has moved from vertex i to vertex j between times 1 and n). Then we have the following.*

(i) *The sequence* $\{W(n)\}_{n\geq 0}$ *converges almost surely toward a point* $W^{\varepsilon} \in \Delta(G)$.

(ii) *Let* $A_{\max} = \sup_{i,j} A_{i,j}$ *and* $G_{\max} = \{(i,j) \in G\colon\ A_{i,j} = A_{\max}\}$. *Point* W^{ε} *satisfies*

$$\lim_{\varepsilon\to 0} W^{\varepsilon}_{i,j} = \frac{\chi_{G_{\max}}(i,j)}{\mathrm{card}(G_{\max})},$$

where $\chi_{G_{\max}}(i,j)$ *equals* 1 *for* $(i,j) \in G_{\max}$ *and* 0 *otherwise.*

COROLLARY 4.3. *Let* $v(n)$ *denote the empirical occupation measure of the process at time* n. *Then*

(i) *The sequence* $\{v(n)\}_{n\geq 0}$ *converges almost surely toward a point* $v^{\varepsilon} \in \Delta$.

(ii) $$\lim_{\varepsilon\to 0} v^{\varepsilon}_i = \frac{\mathrm{card}\{j \in E\colon (i,j) \in G_{\max}\}}{\mathrm{card}(G_{\max})}.$$

Note that there is no presumption here that $G_{\max}$ is connected. Thus Corollary 4.3 has the interpretation that—for small ε—our visitor tends to visit *all* the most attractive spots of the town.

PROOF OF THEOREM 4.2. Define functions $f_{i,j}\colon \Delta(G) \to \mathbf{R}_+$ by

$$f_{i,j}(W) = (A_{i,j} - \varepsilon)(W_{i,j} + W_{j,i}) + \varepsilon A_{i,j}$$

if (i,j) is an edge and

$$f_{i,j}(W) = 0$$

otherwise. Set $S_{i,j}(0) = 1$, $S_{i,j}(n) = 1 + \sum_{k=1}^{n} \delta_{Y_{k-1},i}\delta_{Y_k,j}$ and $W_{i,j}(n) = S_{i,j}(n)/n + \mathrm{card}(G)$. With these notations we have

$$P(Y_{n+1} = j \mid F_n) = \hat{M}_{Y_n,j}(W(n)),$$

where

$$\hat{M}(W)_{i,j} = \frac{f_{i,j}(W)}{\sum_k f_{i,k}(W)}.$$

Let $\mu(W)$ be the measure given by

$$\mu_i(W) = \frac{\sum_k f_{i,k}(W)}{\sum_{l,k} f_{l,k}(W)}, \qquad i = 1,\dots,d.$$

A straightforward computation shows that $(\hat{M}(W), \mu(W))$ is reversible. Thus $\hat{M}(W)$ admits $\mu(W)$ as invariant measure. Now introduce the expended process $Z_n = \{(Y_{n-1}, Y_n)\colon\ n \geq 1\}$. This new process Z is a VRRW, which satisfies

$$P(Z_{n+1} = (i,j) \mid F_n) = N_{Z_n,(i,j)}(W(n))$$

with

$$N_{(k,l),(i,j)}(W) = \delta_{l,i}\hat{M}_{i,j}(W).$$

The map N satisfies Hypothesis 3.1 and for all $W \in \Delta(G)$ the homogeneous Markov chain $N(W)$ admits a unique invariant measure ν defined by

$$\nu_{i,j}(W) = \mu_i(W)\hat{M}_{i,j}(W).$$

Thus, the vector field $F: \Delta(G) \to T\Delta(G)$ associated to Z is given by

$$F_{i,j}(W) = -W_{i,j} + \nu_{i,j}(W) = -W_{i,j} + \frac{f_{i,j}(W)}{\sum_{l,k} f_{l,k}(W)}.$$

Multiplying F by the positive function $\sum_{l,k} f_{l,k}(W)$ leads to the differential system

$$\frac{dW_{i,j}}{dt} = -W_{i,j}\left(\sum_{l,k} f_{l,k}(W)\right) + f_{i,j}(W). \tag{4}$$

LEMMA 4.4. *Let* $\Theta = \{W \in \Delta(G): W_{i,j} > 0; W_{i,j} = W_{j,i}\}$. *Then any compact connected internally chain recurrent set for* (4) *is contained in* Θ.

PROOF. If $W_{i,j} = 0$ then $dW_{i,j}/dt = f_{i,j}(W) > \varepsilon A_{i,j}$. Thus, the vector field F points inward to $\Delta(G)$. This implies that $L \subset \mathrm{Int}(\Delta(G)) = \{W \in \Delta(G): W_{i,j} > 0\}$.

The property $f_{i,j}(W) = f_{j,i}(W)$ implies that

$$\frac{d(W_{i,j} - W_{j,i})^2}{dt} = -2(W_{i,j} - W_{j,i})^2\left(\sum_{l,k} f_{l,k}(W)\right) \le -C(W_{i,j} - W_{j,i})^2$$

for some positive constant C. Thus, $(W_{i,j} - W_{j,i})$ converges exponentially toward zero along the trajectories of (4). This implies that the set $\{W \in \Delta(G): W_{i,j} = W_{j,i}\}$ is a global attractor for (4). The fact that a chain recurrent set contains no proper attractor concludes the proof of the lemma. □

Define functions $g_{i,j}: \mathbf{R}_+ \to \mathbf{R}_+$ by $g_{i,j}(x) = 2(A_{i,j} - \varepsilon)x + \varepsilon A_{i,j}$ if (i, j) is an edge and $g_{i,j}(x) = 0$ otherwise. The dynamics induced by (4) on the positively invariant set Θ are given by

$$\frac{dW_{i,j}}{dt} = -W_{i,j}\left(\sum_{l,k} g_{l,k}(W_{k,l})\right) + g_{i,j}(W_{i,j}) \tag{5}$$

Let $(i_0, j_0) \in G_{\max}$. A straightforward computation shows that (5) admits one unique equilibrium W^ε in Θ which is given by the equations

$$\frac{A_{i,j}}{W^\varepsilon_{i,j}} - \frac{A_{\max}}{W^\varepsilon_{i_0,j_0}} = 2\frac{A_{\max} - A_{i,j}}{\varepsilon}, \qquad (i,j) \in G,$$

where $W^\varepsilon_{i_0,j_0}$ is given by the normalization condition

$$\sum_{(i,j)\in G} \frac{A_{i,j}}{A_{\max}/W^\varepsilon_{i_0,j_0} + 2(A_{\max} - A_{i,j})/\varepsilon} = 1.$$

From these relations it is easy to deduce that W^{ε} satisfies assertion (ii) of Theorem 4.2.

Now, let $L \subset \Delta(G)$ be a compact connected internally chain recurrent set. According to the preceding lemma and by compactness of L, there exists an open subset U of Θ, such that $L \subset U \subset \operatorname{clos}(U) \subset \Theta$. Let $\eta = \inf\{W_{i,j}\colon W \in \operatorname{clos}(U)\} > 0$. Consider the function $H\colon U \to \mathbf{R}$ given by

$$H(W) = \frac{1}{2}\Big(\sum_{k,l} g_{k,l}(W_{k,l})\Big)^2 - \sum_{k,l} \int_{\eta}^{W_{k,l}} \frac{g_{k,l}(x)\, g'_{k,l}(x)}{x}.$$

Taking the derivative of H along trajectories of (5) gives

$$\begin{aligned}\frac{dH}{dt} &= \sum_{k,l} \frac{\partial H}{\partial W_{k,l}} \frac{dW_{k,l}}{dt}\\ &= -\sum_{k,l} \frac{g'_{k,l}(W_{k,l})}{W_{k,l}} \bigg(-W_{k,l}\Big[\sum_{i,j} g_{i,j}(W_{i,j})\Big] + g_{k,l}(W_{k,l})\bigg)^2\\ &= -\sum_{k,l} \frac{g'_{k,l}(W_{k,l})}{W_{k,l}} \bigg(\frac{dW_{k,l}}{dt}\bigg)^2.\end{aligned}$$

Since $g'_{k,l}(W_{k,l})/W_{k,l} > 0$ for $W \in U$, this implies that H is strictly decreasing along nonstationary orbits of the dynamics induced by (5) on L. We conclude exactly as in the proof of Theorem 4.1 that L must consist of equilibria. It follows that $L = \{W^{\varepsilon}\}$. This concludes the proof of theorem 4.2. □

5. Proof of Proposition 3.3. The proof of Proposition 3.3 uses the ideas introduced in Métivier and Priouret (1987) with some simplifications due to the particular form of the process considered here. The main idea is to introduce a solution to the Poisson equation for the Markov chain $M(v)$.

Let $M^{\infty}(v)$ denote the matrix defined by

$$M^{\infty}_{i,j}(v) = \pi_j(v).$$

Let $H_t(v)$ denote the matrix solution to the linear differential equation

$$\frac{dH_t(v)}{dt} = (-\mathrm{Id} + M(v))\, H_t(v)$$

with initial condition $H_0(v) = \mathrm{Id}$. In other words $t \to H_t(v) = \exp(t(-\mathrm{Id} + M(v))$ is the transition matrix function of the continuous time Markov process associated with $M(v)$. Indecomposability of $M(v)$ implies that $H_t(v)$ converges toward $M^{\infty}(v)$ at an exponential rate. Thus, the map

$$Q\colon \Delta \to TM$$

$$v \to Q(v) = \int_0^{\infty} (H_t(v) - M^{\infty}(v))\, dt$$

is well defined.

LEMMA 5.1. (i) $(\mathrm{Id} - M(v))Q(v) = \mathrm{Id} - M^\infty(v) = Q(v)(\mathrm{Id} - M(v))$.
(ii) *Q is C^k.*

PROOF. (i) It is easily seen that $M(v)M^\infty(v) = M^\infty(v) = M^\infty(v)M(v)$. From these relations we deduce (i).

(ii) Consider the map

$$L\colon \Delta \times TM \to TM$$
$$L(v, Q) = Q(\mathrm{Id} - M(v)) - (\mathrm{Id} - M^\infty(v)).$$

By assertion (i), $Q = Q(v)$ is solution to the implicit equation $L(v, Q) = 0$. We have

$$\frac{\partial L}{\partial Q}(v, Q(v))\colon TM \to TM$$
$$A \to A \cdot (\mathrm{Id} - M(v)).$$

Uniqueness of the invariant measure implies that the kernel of $(\partial L/\partial Q)(v, Q(v))$ is zero. Indeed if $A = AM(v)$, each column of A must be proportional to $\pi(v)$ but since A is in TM, A must be the zero matrix. We conclude that Q is C^k by the implicit function theorem. □

For each $v \in \Delta$, $y \in E$ let $U(v, y) \in T\Delta$ be the vector defined by

$$U(v, y)_j = \delta_{y, j} - \pi_j(v), \qquad j = 1, \dots, d.$$

Since

$$S_i(n+1) - S_i(n) = \delta_{Y_{n+1}, i},$$

we have

$$v(n+1) - v(n) = \frac{1}{n+1+n_0} F(v(n)) + \varepsilon_{n+1},$$

where

$$\varepsilon_{n+1} = \frac{U(v(n), Y_{n+1})}{n+1+n_0}.$$

Or, equivalently, using the interpolated process

$$V(\tau_{n+1}) - V(\tau_n) = (\tau_{n+1} - \tau_n) F(V(\tau_n)) + \varepsilon_{n+1} \tag{6}$$

Set $m(t) = \sup\{p \in \mathbf{N}\colon \tau_p \le t\}$ and

$$\varepsilon(n, T) = \sup_{\{k\colon 0 \le \tau_k - \tau_n \le T+1\}} \left\| \sum_{i=n}^{k-1} \varepsilon_{i+1} \right\|$$

Now, if one compares (6) with the solution to the deterministic equation

$$\frac{dW}{ds} = F(W(s)), \qquad t \le s \le T$$

with initial condition $W(t) = V(t)$, it is easily seen by a standard Gronwall's inequality [see, e.g., Lemma 4.4 of Benaïm (1996)] that

$$\sup_{0 \le h \le T} d\big(\Phi_h(V(t)), V(t+h)\big) \le C(T)\varepsilon(m(t), T) \tag{7}$$

for some constant $C(T) > 0$. Thus the proof of Proposition 3.3 reduces to show that $\lim_{n\to\infty} \varepsilon(n, T) = 0$.

Notation. If A is a matrix and $y \in E$, we let $A[y]$ denote the vector whose jth component is $A_{y,j}$. We let Cste denote an arbitrary positive constant depending only on the map M.

By Lemma 5.1 we have $U(v, y) = (\mathrm{Id} - M^\infty(v))[y] = Q(v)[y] - (M(v)Q(v))[y]$. Thus

$$\varepsilon_{i+1} = \varepsilon_{1,i+1} + \varepsilon_{2,i+1} + \varepsilon_{3,i+1} + \varepsilon_{4,i+1},$$

where

$$\varepsilon_{1,i+1} = \frac{Q(v(i))[Y_{i+1}] - (M(v(i))Q(v(i)))[Y_i]}{i+1+n_0},$$

$$\varepsilon_{2,i+1} = \frac{(M(v(i))Q(v(i)))[Y_i]}{i+1+n_0} - \frac{M(v(i))Q(v(i))[Y_i]}{i+n_0},$$

$$\varepsilon_{3,i+1} = \frac{(M(v(i))Q(v(i)))[Y_i]}{i+n_0} - \frac{(M(v(i+1))Q(v(i+1)))[Y_{i+1}]}{i+1+n_0},$$

and

$$\varepsilon_{4,i+1} = \frac{(M(v(i+1))Q(v(i+1)))[Y_{i+1}]}{i+1+n_0} - \frac{(M(v(i))Q(v(i)))[Y_{i+1}]}{i+1+n_0}.$$

Continuity of Q, M and compactness of Δ ensure that

$$\|\varepsilon_{2,i+1}\| \le \frac{\mathrm{Cste}}{(i+n_0)^2} \quad \text{and} \quad \left\| \sum_{i=n}^{k} \varepsilon_{3,i+1} \right\| \le \frac{\mathrm{Cste}}{n+1+n_0}.$$

By smoothness of M and Q we get

$$\|\varepsilon_{4,i+1}\| \le \mathrm{Cste}\frac{\|v(i+1) - v(i)\|}{i+1} \le \mathrm{Cste}\frac{1}{(i+1)^2}.$$

Thus

$$\varepsilon(n, T) \le \sup_{\{k:\, 0 \le \tau_k - \tau_n \le T+1\}} \left\| \sum_{i=n}^{k-1} \varepsilon_{1,i+1} \right\| + \frac{C_1}{n+1+n_0} \tag{8}$$

for some positive constant C_1 (depending only on M). To conclude, notice that $E(\varepsilon_{1,i+1} | F_i) = 0$ and $\|\varepsilon_{1,i+1}\|^2 \le C_2/(i+n_0)^2$ for some positive constant C_2 (depending only on M). Thus $Z_n = \sum_{i=1}^n \varepsilon_{1,i+1}$ is a martingale which converges almost surely and, therefore, $\lim_{n\to\infty} \varepsilon(n, T) = 0$. This concludes the proof of Proposition 3.3.

REMARK. Notice that the variance of the martingale $\sum_{i=n}^{k-1} \varepsilon_{1,\,i+1}$ is bounded by

$$C_2 \sum_{i=n}^{k-1} \frac{1}{(n_0+i+1)^2} \le C_2 \sum_{i=n}^{k-1} \frac{\tau_{i+1}-\tau_i}{(n_0+i+1)} \le C_2 \frac{\tau_k-\tau_n}{n_0+n}.$$

Thus from estimate (8) and a Doob's inequality, we get the estimate: for any $\alpha > 0$ and $n + n_0 > 2C_1/\alpha$,

$$\begin{aligned} P(\varepsilon(n,T) > \alpha) &\le P\left(\sup_{\{k:\, 0 \le \tau_k - \tau_n \le T+1\}} \left\| \sum_{i=n}^{k-1} \varepsilon_{1,\,i+1} \right\| > \alpha/2 \right) \\ &\le \frac{C_1(T)}{\alpha^2(n_0+n)} \end{aligned} \tag{9}$$

for some constant $C_1(T)$ depending on T and M.

6. Reachable sets and attractors. The aim of this section is to show that the behavior of general VRRW can be as complicated as the behavior of any vector field on Δ and to give the practical condition ensuring that a given attractor of F has a positive probability to be the limit set of $\{v_n\}$. We begin modestly in Lemma 6.1 by showing that (up to a time reparametrization) any vector field on Δ is the vector field associated to a VRRW.

LEMMA 6.1. *Let G: $\Delta \to T\Delta$ be a C^k, $k \ge 1$ vector field leaving Δ positively invariant. Then there exists a C^k map M: $\Delta \to M$ and $\varepsilon > 0$ such that we have the following.*

(i) *M satisfies Hypothesis* 3.1.
(ii) *The vector field F associated to M is $F = \varepsilon G$.*

PROOF. Set $\pi(\varepsilon, v) = v + \varepsilon G(v)$. We claim that for $\varepsilon > 0$ small enough, $\pi(\varepsilon, v) \in \Delta$. Clearly $\pi(\varepsilon, v) \in \mathrm{aff}(\Delta)$ for all ε.

Let $p^i(v) = (p^i(v)_1, \dots, p^i(v)_d) \in \Delta$ denote the point defined by $p^i(v)_j = v_j$ for $j \ne i$ and $p^i(v)_i = 0$. By smoothness of G there exists a continuous function $\tilde{G}$ such that $G_i(v) - G_i(p^i(v)) = v_i \tilde{G}_i(v)$. Thus $\pi(\varepsilon, v) = v_i(1 + \varepsilon \tilde{G}_i(v)) + \varepsilon G_i(p^i(v))$. Choose ε small enough so that $1 + \varepsilon \tilde{G}_i(v) > 0$. Since G leaves Δ positively invariant, $G_i(p^i(v)) \ge 0$. Therefore $\pi_i(\varepsilon, v) \ge 0$. To conclude, it suffices to choose $M_{i,\,j}(v) = \pi_j(\varepsilon, v)$. $\square$

Let $G_M(v) \subset E \times E$ denote the graph of the Markov transition matrix $M(v)$. That is:

$$(i,j) \in G_M(v) \quad \Leftrightarrow \quad M_{i,\,j}(v) > 0.$$

A sequence $x_0, \dots, x_k \in \Delta$ is called *admissible* if there exist vertices $i_0, \dots, i_k \in E$ such that statements (1) and (2) hold.

1. $(i_l, i_{l+1}) \in G_M(x_l)$; $l = 0, \dots, k-1$;
2. $x_{l+1} - x_l = 1/(l+1+n_0)(-x_l + e_{i_l})$, where $e_1, \dots, e_d$ denote the vectors of the canonical basis of $\mathbf{R}^d$.

The points x_0 and x_k are, respectively, the *initial* and *terminal* points of the sequence. The length of the sequence is the integer $k+1$. The vertices $i_0, \dots, i_k$ are the states of the sequence.

A point $w \in \Delta$ which is the terminal point of an admissible sequence with initial point v and initial state y is called *reachable* from (v, y). The set of all points which are reachable from (v, y) is called the *reachable set* of (v, y) and is denoted $R^+(v, y)$. The set of all points which are reachable from (v, y) by admissible sequences of length greater than $k+1$ is denoted $R^+(k, v, y)$. Thus $R^+(v, y) = \bigcup_{k \geq 0} R^+(k, v, y)$.

Define the G_M-limit set of v with initial state y, as

$$S^+(v, y) = \bigcap_{k \geq 0} \overline{R^+(k, v, y)}.$$

Equivalently, a point $w \in \Delta$ belongs to $S^+(v, y)$ if and only if for every neighborhood U of w and every integer k:

$$P\big(v(n) \in U \text{ for some } n \geq k | v(0) = v, Y_0 = y\big) > 0.$$

LEMMA 6.2. *Let $v \in \Delta$ and $y \in E$. Then we have the following.*

(i) *$S^+(v, y)$ is a nonempty compact set positively invariant under the flow Φ [i.e., $\Phi_t(S^+(v, y)) \subset S^+(v, y)$ for all $t \geq 0$].*
(ii) $P(L(\{v_n\}) \subset S^+(v, y) | v(0) = v, Y_0 = y) = 1$.

PROOF. (i) By compactness of Δ and definition of $S^+(v, y)$, it is clear that $S^+(v, y)$ is a nonempty compact subset of Δ. Let $w \in S^+(v, y)$ and $T > 0$. Our next goal is to show that $\Phi_T(w) \in S^+(v, y)$. Fix $\varepsilon > 0$. By uniform continuity of Φ_T there exists $\alpha > 0$ such that

$$d(x, x') \leq \alpha \quad \Rightarrow \quad d\big(\Phi_T(x), \Phi_T(x')\big) < \varepsilon.$$

Since $w \in S^+(v, y)$, for every $k \in \mathbf{N}$, there exists an admissible sequence $v = x_0, \dots, x_m$ with $m \geq k$ and $d(x_m, w) \leq (\alpha/2)$. Consider now the process $\{Y_n, S(n)\}_{n \geq 0}$ with initial value $v(0) = x_m$, initial mass $n'(0) = m + n_0$ and where the initial state Y_0 is chosen to be the terminal state i_m of the sequence. According to estimates (7), (8) and (9), we see that

$$P\Big(d\big(V(T), \Phi_T(v(0)) > \varepsilon/2\big)\Big) \leq C_2(T) \frac{1}{(n'(0) + 1)\varepsilon^2}$$

for some positive constant $C_2(T)$, where V denotes the interpolated process associated to $\{Y_n, S(n)\}$. Since $n'(0) = n_0 + m$ we see that for k large enough this probability can be made arbitrarily small. Therefore there exists an admissible sequence $z_0 = x_m, \dots, z_l$ from x_m to a point z_l which is in the $\varepsilon/2$ neighborhood of $\Phi_T(x_m)$. Since $d(x_m, w) \leq \alpha$ we get $d(\Phi_T(w), z_l) \leq \varepsilon$.

Putting together the sequences $v = x_0, \dots, x_m$ and $z_0, \dots, z_l$ gives a new admissible sequence of length greater than k from v to $z_l \in B_\varepsilon(\Phi_T(w))$. Since k and ε are arbitrary, this shows that $\Phi_T(w) \in S^+(v, y)$.

(ii) follows from the definitions of $S^+(v, y)$ and $L\{(v_n)\}$. $\square$

A subset $A \subset \Delta$ is an attractor for Φ if the two following conditions hold:

1. A is nonempty, compact and invariant ($\Phi_t A = A$).
2. A has a neighborhood $W \subset \Delta$ such that $\operatorname{dist}(\Phi_t x, A) \to 0$ as $t \to \infty$ uniformly in $x \in W$.

The basin of A denoted $B(A)$ is the positively invariant open set comprising all points x such that $\operatorname{dist}(\Phi_t x, A) \to 0$ as $t \to \infty$.

A global attractor is an attractor whose basin is all the space Δ.

An attractor A is said to be *minimal* if it contains no proper invariant subset or, equivalently, every orbit in A is dense in A. Simple examples of minimal attractors are asymptotically stable equilibria and periodic orbits.

The next two theorems are proved at the end of this section.

THEOREM 6.3. *Let $A \subset \Delta$ be an attractor with basin of attraction $B(A)$. Let $v \in \operatorname{Int}(\Delta)$ and $y \in E$. Suppose that $S^+(v, y) \cap B(A) \neq \varnothing$, then*

$$P\big(L(\{v_n\}) \subset A | v(0) = v, Y_0 = y\big) > 0.$$

In order to state the next theorem we need to define a few terms. Let $G \subset E \times E$ be an oriented graph whose vertices are the elements of E. We use the notation $i \to j$ to represent the arrow $(i, j) \in G$. We suppose that G is indecomposable. By *indecomposable*, we mean that there exists a nonempty set $R_G \subset E$, called the *recurrence class* of G, such that for every pair $(i, j) \in E \times R_G$ it is possible to find a sequence of arrows $i \to i_1 \to \cdots \to i_n \to j$ but no such sequence exists if $(i, j) \in R_G \times (E - R_G)$.

A set $C = \{i_0, i_1, \dots, i_{n-1}\} \subset E$ is called a *cycle* if there exists a permutation $\sigma\colon \{0, \dots, n-1\} \to \{0, \dots, n-1\}$ such that $i_{\sigma(0)} \to i_{\sigma(1)} \to \cdots \to i_{\sigma(n-1)} \to i_{\sigma(0)}$. A proper subcycle of C is a proper subset of C which is a cycle. We define the barycenter of C as the vector

$$B(C) = \frac{1}{\operatorname{card}(C)} \sum_{i \in C} e_i,$$

where $\operatorname{card}(C)$ is the cardinal of C. We call C a *basic cycle* of G if it satisfies the two following conditions.

1. $C \subset R_G$.
2. $B(C)$ cannot be written as a convex combination of the barycenters of proper subcycles of C.

Let $C_1, \dots, C_k$ denote the proper subcycles of C. The condition (2) is equivalent to the condition that the linear system

$$\sum_{j=1}^{k} 1_{\{i \in C_j\}} x_j = 1, \qquad i = 1, \dots, \operatorname{card}(C)$$

admits no solution $\{x_j\}$ with nonnegative entries.

Finally we will say that two graphs G and G' are equivalent (written $G \sim G'$) if they have the same set of basic cycles. Figure 1 gives an example of two equivalent graphs for which the basic cycles are $\{1\}, \{2\}, \{1, 2, 3\}$.

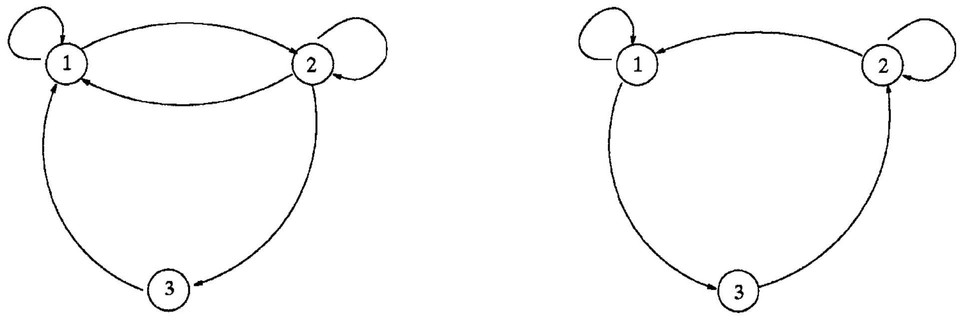

FIG. 1. *Two equivalent graphs. The basic cycles are* $\{1\}, \{2\}$ *and* $\{1, 2, 3\}$.

THEOREM 6.4. *Suppose that one of the two following conditions* (a) *or* (b) *holds.*

(a) *There exists an indecomposable graph* $G \subset E \times E$ *such that*:

$$\forall\ v \in \mathrm{Int}(\Delta), \qquad G \subset G_M(v) \text{ and } G \sim G_M(v).$$

(b) $\forall (v, v') \in \Delta \times \Delta$, $G_M(v) \sim G_M(v')$.

Then for all $v \in \mathrm{Int}(\Delta)$, $y \in E$, $S^+(v, y)$ *contains a global attractor of* Φ.

COROLLARY 6.5. *Suppose that condition* (a) *or* (b) *of Theorem* 6.4 *holds. Then for every attractor* $A \subset \Delta$, $v \in \mathrm{Int}(\Delta)$, $y \in E$,

$$P\big(L(\{v_n\}) \subset A | v(0) = v, Y_0 = y\big) > 0.$$

If furthermore, A is a minimal attractor then

$$P\big(L(\{v_n\}) = A | v(0) = v, Y_0 = y\big) > 0.$$

The first part of the corollary follows immediately from Theorems 6.3 and 6.4. The second part follows from Corollary 3.5

REMARKS. (i) When the map M is given by (2) and R has positive entries, condition (a) of Theorem 6.4 is satisfied. Thus Corollary 6.5 implies that any minimal attractor has a positive probability to be the limit set of $\{v(n)\}$. This has been proved by Pemantle (1992) for asymptotically stable equilibria when R is symmetric.

(ii) According to Lemma 6.1 and Corollary 6.5, we see that for an arbitrary VRRW the limit set of $\{v_n\}$ can be as complicated as any minimal attractor. For example, it is easy to construct a generalized VRRW on $E = \{1, 2, 3\}$ such that $L(\{v_n\})$ is homeomorphic to a circle or a VRRW on $E = \{1, 2, 3, 4\}$ for which $L(\{v_n\})$ is homeomorphic to a two-torus.

7. Shadowing properties. This section states some shadowing properties for generalized VRRW which are used to give a partial answer to Conjecture 1.2.

Let $K \subset \Delta$ be a compact set positively invariant by the flow Φ.

The expansion constant of Φ_t at K is

$$EC(\Phi_t, K) = \inf\{m(D\Phi_t(x)) \colon x \in K\},$$

where

$$m(D\Phi_t(x)) = \inf\{\|D\Phi_t(x)v\| \colon \|v\| = 1\} = \|D\Phi_t(x)^{-1}\|^{-1}.$$

The logarithmic expansion rate of Φ at K is defined as

$$l_{\exp}(\Phi, K) = \lim_{t\to\infty} \frac{1}{t} \log(EC(\Phi_t, K)),$$

where the limit exists by standard subadditivity arguments. The logarithmic expansion rate has been introduced and used by Hirsch (1994) to study shadowing properties of dynamical systems. The techniques of Hirsch (1994) convert nicely to prove shadowing theorems for stochastic approximation processes (Benaïm, 1996). Recently Schreiber (1995) proved that $l_{\exp}(\Phi, K)$ equals the smallest Liapounov exponent for ergodic measures of Φ with support in K. This has the nice consequence that

$$l_{\exp}(\Phi, K) = l_{\exp}(\Phi, M(\Phi) \cap K),$$

where $M(\Phi)$ is the smallest closed set such that $\mu(M) = 1$ for every Φ invariant probability measure. By the Poincaré recurrence theorem $M(\Phi) \cap K$ is contained in the Birkoff center of Φ restricted to K:

$$BC(\Phi, K) = \left\{x \in K \colon \liminf_{t\to\infty} d(x, \Phi_t(x)) = 0\right\}.$$

This last property is particularly useful to estimate $l_{\exp}(\Phi, K)$. The next result follows easily from Theorem 5.2 of Benaïm (1996)

THEOREM 7.1. *Let $A \subset \Delta$ be an attractor for F. Suppose that $l_{\exp}(\Phi, A) > -1/2$. Then with probability 1 on the set $\{L(v_n) \subset A\}$, there exists a random vector $W \in B(A)$ such that*

$$\lim_{n\to\infty} \|\Phi_{\tau_n}(W) - v_n\| = 0.$$

COROLLARY 7.2. *Let $A \subset C = \{v \in \Delta \colon v = \pi(v)\}$ be an attractor. Suppose that for each $v \in A$ all the eigenvalues of the linear map $D\pi(v) \colon T\Delta \to T\Delta$ have their real parts greater than $1/2$. Then with probability 1 on the set $\{L(v_n) \subset A\}$, there exists a random vector $W \in B(A)$ such that*

$$\lim_{n\to\infty} \|\Phi_{\tau_n}(W) - v_n\| = 0.$$

PROOF. If $v \in C$, $D\Phi_t(v) = \exp(tDF(v))$. Since $\Phi|A$ is the identity flow, ergodic invariant measures of $\Phi|A$ are Dirac measure. The smallest Liapounov exponent of $\Phi|A$ for the Dirac measure δ_v equals $\lim_{t\to\infty}(1/t)\log m(\exp(tDF(v)))$. A straightforward application of Jordan's decomposition shows that this last number equals the real part of the eigenvalue of $DF(v)$ having the smallest real part. Since $DF(v) = -\mathrm{Id} + D\pi(v)$, the result follows. $\square$

This corollary will now be used to give a partial answer to Pemantle's Conjecture 1.2.

COROLLARY 7.3. *Suppose* $M(v)$ *is given by* (2) *with* $R_{i,j} = R_{j,i} > 0$. *Suppose that*

$$\frac{\inf_{i,j} R_{i,j}}{\sup_{i,j} R_{i,j}} > \frac{1}{2}.$$

Then Conjecture 1.2 *is true.*

PROOF. Let $v \in C$. That is $v = \pi(v)$. First suppose that $v \in \mathrm{Int}(\Delta)$. A simple computation shows that

$$\frac{\partial \pi_i}{\partial v_j}(v) = v_i\left(\frac{R_{i,j}}{H(v)} - 2\right) + \delta_{i,j}.$$

Since $v \in C \cap \mathrm{Int}(\Delta)$, we have $h_i(v) = H(v)$. Thus

$$\frac{\partial \pi_i}{\partial v_j}(v) = v_i\left(\frac{R_{j,i}}{h_i(v)} - 2\right) + \delta_{i,j} = M_{j,i}(v) - 2v_i + \delta_{i,j}.$$

It follows that $D\pi(v)$ is given by:

$$\begin{aligned} &D\pi(v)\colon T\Delta \to T\Delta,\\ &w \to M^T(v)w + w \end{aligned}$$

and the computation of the eigenvalues of $D\pi(v)$ reduces to the computation of eigenvalues of $M(v)$. Recall that the Markov chain $M(v)$ is reversible with invariant measure $\pi(v) = v$. Therefore, eigenvalues of $M(v)$ are reals. Also, if we let $\beta(v)$ denote the smallest eigenvalue of $M(v)$, $\beta(v)$ satisfies the variational formula

$$1 + \beta(v) = \inf\left\{\frac{\langle (\mathrm{Id} + M(v))f, f\rangle_v}{\langle f, f\rangle_v} : f \neq 0\right\},$$

where $\langle f, g \rangle_v = \sum_i f_i g_i v_i$ denote the scalar product in $l^2(v)$. Now we use the following convenient fact:

$$\begin{aligned}\langle (\mathrm{Id} + M(v)) f, f \rangle_v &= \frac{1}{2} \sum_{i,j} (f_i + f_j)^2 M_{i,j}(v) \pi_i(v) \\ &= \frac{1}{2} \sum_{i,j} (f_i + f_j)^2 \frac{R_{i,j} v_j}{h_i(v)} v_i \\ &\geq \frac{\inf_{i,j} R_{i,j}}{\sup_{i,j} R_{i,j}} \frac{1}{2} \sum_{i,j} (f_i + f_j)^2 v_j v_i \\ &= \frac{\inf_{i,j} R_{i,j}}{\sup_{i,j} R_{i,j}} \big(\langle f, f \rangle_v + 2 \langle f, 1 \rangle_v^2 \big).\end{aligned}$$

It follows that

$$1 + \beta(v) \geq \frac{\inf_{i,j} R_{i,j}}{\sup_{i,j} R_{i,j}} > \frac{1}{2}.$$

In the situation where v is not in $\mathrm{Int}(\Delta)$ we can always suppose that $v = (v_1, \dots, v_k, 0, \dots, 0)$ with $v_i > 0$ for $i = 1, \dots, k$. Then it is easily seen that the matrix $((\partial \pi_i / \partial v_j)(v))_{i,j}$ takes the form

$$\left(\frac{\partial \pi_i}{\partial v_j}(v) \right)_{i,j} = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix},$$

where D is a $d - k \times d - k$ diagonal matrix with entries $D_{i,i} = h_i(v)/H(v)$, $i = 1, \dots, k$ and A is a $k \times k$ matrix whose smallest eigenvalue can be estimated exactly in the case $v \in \mathrm{Int}(\Delta)$.

Now, we shall use the fact that for every $x \in \Delta$ the deterministic trajectory $\{\Phi_t(x)\}$ converges, as t approaches ∞, to a unique limit point which is an equilibrium. A nice proof of this result has been given, in a very different context, by Losert and Akin (1983, Theorem 2). Losert and Akin's convergence theorem combined with Corollary 7.2 concludes the proof. □

8. Proofs of Theorems 6.3 and 6.4.

Proof of Theorem 6.3. Let $X: \mathbf{R}_+ \to \Delta$ be an asymptotic pseudotrajectory of Φ. For any $T > 0$, define

$$d_X(T) = \sup_{k \in \mathbf{N}} d\big(\Phi_T(X(kT)), X(kT + T)\big).$$

If a point $x \in \Delta$ belongs to the basin of attraction of an attractor $A \subset \Delta$, then $\Phi_t(x) \to A$ as $t \to \infty$. The next lemma shows that the same is true for an asymptotic pseudotrajectory X provided that $d_X(T)$ is small enough.

Lemma 8.1. *Let $A \subset \Delta$ be an attractor with basin $B(A)$ and let $K \subset B(A)$ be a nonempty compact set. There exist numbers $T > 0$, $\delta > 0$ depending only on K such that: if X is an asymptotic pseudotrajectory with $X(0) \in K$ and $d_X(T) < \delta$, then $L(X) \subset A$.*

PROOF. Choose an open set W such that $A \cup K \subset W \subset \text{clos}(W) \subset B(A)$ and choose $\delta > 0$ such that $N_{2\delta}(A)$ (the 2δ neighborhood of A) is contained in W. Since A is an attractor, there exists $T > 0$ such that $\Phi_T(W) \subset N_\delta(A)$. Now, if $X(0) \in K$ and $d_X(T) < \delta$, we have $\Phi_T(X(0)) \in N_\delta(A)$ and $d(X(T), \Phi_T(X(0))) < \delta$. Thus $X(T) \in N_{2\delta}(A) \subset W$. By induction it follows that $X(kT) \in W$ for all $k \in \mathbf{N}$. Thus, by compactness, $L(X) \cap \text{clos}(W) \neq \emptyset$.

By a theorem of Conley (1978), every chain recurrent point for Φ is contained in $A \cup (\Delta - B(A))$. Therefore, $L(X)$ (which is a connected set consisting of chain recurrent points by Theorem 3.4) is either contained in A or in $\Delta - B(A)$ but since $L(X) \cap \text{clos}(W) \neq \emptyset$, we must have $L(X) \subset A$. □

We will now use Lemma 8.1 to prove Theorem 6.3.

Let $K \subset B(A)$ be a compact neighborhood of A. First notice that $S^+(v, y) \cap K \neq \emptyset$. Indeed, by hypothesis there exists $w \in S^+(v, y) \cap B(A)$. Thus for $t > 0$ large enough, $\Phi_t(w) \in K$ and since by Lemma 6.2, $S^+(v, y)$ is positively invariant, $\Phi_t(w) \in S^+(v, y)$.

To the compact set K we associate the numbers $T > 0$, $\delta > 0$ given by Lemma 8.1. Let $n \in \mathbf{N}$ be such that $n > (2C_1/\delta)$ where C_1 is the constant which appears in formula (8). Since $S^+(v, y)$ intersects K, there exists $m \geq n$ such that $P(v(m) \in K | v(0) = v,\ Y_0 = y) > 0$. Now, consider the interpolated process X defined by $X(t) = V(t + \tau_m)$ for $t \geq 0$. According to Proposition 3.3, X is almost surely an asymptotic pseudotrajectory of Φ. On the other hand,

$$P\big(d_X(T) > \delta\big) \leq \sum_{k \in \mathbf{N}} P\big(d\big(\Phi_T(X(kT), X(kT + T)\big) > \delta\big)$$

Thus using estimates (7) and (9),

$$\begin{aligned} P\big(d_X(T) > \delta\big) &\leq \sum_{k \in \mathbf{N}} P\left(\varepsilon(m(\tau_m + k), T) > \frac{\delta}{C(T)}\right) \\ &\leq \frac{C_1(T) C(T)^2}{\delta^2} \sum_{k \in \mathbf{N}} \frac{1}{m(\tau_m + k)}. \end{aligned}$$

Since $\log(p) - \tau_p$ converges toward some constant (depending on n_0) as $p \to \infty$, it follows that

$$m(\tau_m + k) \geq C(n_0) m e^k - 1,$$

where $C(n_0)$ is a positive constant depending on n_0. Thus,

$$P\big(d_X(T) > \delta\big) \leq \frac{C(n_0, T)}{m\delta^2}$$

for some constant $C(n_0, T)$ depending on n_0 and T. It follows that for n large enough, the event $\{d_X(T) > \delta\}$ has a positive probability and the result follows from Lemma 8.1.

PROOF OF THEOREM 6.4. Let $G \subset E \times E$ be an indecomposable graph and R_G its class of recurrence. Define sets

$$\Sigma_G^n = \{\hat{y} = (y_0, \dots, y_{n-1}) \in R_G \times \cdots \times R_G \colon y_0 \to \cdots \to y_{n-1}\}, \qquad n \geq 1$$

and

$$\Sigma_G = \bigcup_{n \in \mathbf{N}^*} \Sigma_G^n.$$

The set Σ_G is the space of all finite sequences with symbols in R_G which are allowable for G. Given $\hat{y} = (y_0, \dots, y_{n-1}) \in \Sigma_G^n$, set

$$s(\hat{y}) = \sum_{i=0}^{n-1} e_{y_i}, \qquad |y| = n.$$

The occupation measure of G is the map $\nu \colon \Sigma_G \to \Delta$ defined by

$$\nu(\hat{y}) = \frac{s(\hat{y})}{|y|}.$$

A point $v \in \Delta$ will be called a *limit point* of ν if there exists a sequence $\hat{y}_n \in \Sigma_G$ with $\lim_{n\to\infty} |y_n| = \infty$ and $v = \lim_{n\to\infty} \nu(\hat{y}_n)$. We let $\nu^*(G)$ denote the set of all limit points for ν.

The following notation will be useful: let $\hat{y} = (y_0, \dots, y_{n-1}) \in \Sigma_G^n$ and $\hat{z} = (z_0, \dots, z_{m-1}) \in \Sigma_G^m$. If $y_{n-1} \to z_0$, we set $\hat{y}\hat{z} = (y_0, \dots, y_{n-1}, z_0, \dots, z_{m-1})$.

LEMMA 8.2. *The set $\nu^*(G)$ is the convex hull of the barycenters of the basic cycles of G.*

PROOF. We prove this result in the case $R_G = E$ (otherwise it suffices to consider the restriction of G to R_G).

Our first goal is to show that $\nu^*(G)$ is convex. Let $v, w \in \nu^*(G)$ and $0 < t < 1$. Fix $\varepsilon > 0$, $n \in \mathbf{N}$. Choose sequences $\hat{y}, \hat{z} \in \Sigma_G$ such that $|\hat{y}| > n$, $|\hat{z}| > n$, $\|v - \nu(\hat{y})\| < \varepsilon/4$ and $\|w - \nu(\hat{z})\| < \varepsilon/4$.

By adding at most d symbols at the beginning and the end of these sequences we can suppose that $y_0 = y_{|y|-1} = z_0 = z_{|z|-1}$. Choose integers p, q such that $|p/(p+q) - t| < \varepsilon/4$. Then we have

$$\nu\left(\hat{y}^{p|\hat{z}|}\hat{z}^{q|\hat{y}|}\right) = \frac{p|\hat{z}|s(\hat{y}) + q|\hat{y}|s(\hat{z})}{(p+q)|\hat{y}||\hat{z}|} = \frac{p}{p+q}\nu(\hat{y}) + \frac{q}{p+q}\nu(\hat{z}),$$

where we use the notation $\hat{y}^{q+1} = \hat{y}^q\hat{y}$. Thus

$$\left\| \nu\left(\hat{y}^{p|\hat{z}|}\hat{z}^{q|\hat{y}|}\right) - (tv + (1-t)w) \right\| \le \varepsilon.$$

Therefore $(tv + (1-t)w) \in \nu^*(G)$. Hence $\nu^*(G)$ is convex.

Let H denote the convex hull of barycenters of the cycles of G. Let $C = \{i_0, \dots, i_{n-1}\}$ be a cycle of G. Suppose $i_0 \to i_1 \to \cdots \to i_{n-1} \to i_0$ and set $\hat{c} = (i_0, \dots, i_{n-1})$. Clearly $B(C) = \nu(\hat{c}^m)$ for every $m \in \mathbf{N}$, $m \ge 1$. Thus, $B(C) \in \nu^*(G)$. It follows that $H \subset \nu^*(G)$. Our next goal is to show that conversely, $\nu^*(G) \subset H$. This will conclude the proof of the lemma because, by definition of basic cycles, H is also the convex hull of barycenters of basic cycles.

Let $\Sigma_G^{n'}$ denote the subset (possibly empty) of Σ_G^n consisting of sequences $\hat{y} = (y_0, \dots, y_{n-1})$ for which $y_{n-1} \to y_0$. We claim that $\nu(\Sigma_G^{n'}) \subset H$. We prove this claim by induction on n. Suppose $n = 1$ and $\Sigma_G^{1'} \neq \emptyset$ (otherwise the result is trivial). Let $\hat{y} = (y_0) \in \Sigma_G^{1'}$. By definition $\{y_0\}$ is a cycle, thus $\nu(\hat{y}) = e_{y_0} = B(\{y_0\})$ and the claim is true for $n = 1$.

Now fix $n > 1$ and suppose that the claim is true for every $1 \le k < n$. Suppose $\Sigma_G^{n'} \neq \emptyset$ and let $\hat{y} = (y_0, \dots, y_{n-1}) \in \Sigma_\Phi^{n'}$. If $\{y_0, \dots, y_{n-1}\}$ is a cycle we are done. Otherwise, write $\hat{y} = (y_0, \dots, y_k, \dots, y_l \dots, y_{n-1})$, where $y_i \neq y_j$ for $i, j \le l$, $i \neq j$ and $y_{l+1} = y_k$. Define sequences $\hat{c} = (y_k, \dots, y_l)$ and $\hat{z} = (y_0, \dots, y_{k-1}, y_{l+1}, \dots, y_{n-1})$. We have

$$\nu(\hat{y}) = \frac{|\hat{c}| B(C) + |\hat{z}| \nu(\hat{z})}{|\hat{z}| + |\hat{c}|},$$

where $C = \{y_k, \dots, y_l\}$. Since by induction hypothesis, $\nu(\hat{z}) \in H$, it follows that $\nu(\hat{y}) \in H$ and the claim is proved.

Now, let $v \in \nu^*(G)$. As already noticed, v can be written as $v = \lim_{k\to\infty} \nu(\hat{y}^{n_k})$ with $n_k \to \infty$ and $\hat{y}^{n_k} \in \Sigma_G^{n'_k}$. It follows that $\nu^*(G) \subset H$. $\square$

We now pass to the proof of Theorem 6.4. By assumption (condition (a) or (b) of Theorem 6.4), the graphs $G_M(v)$ of the matrices $M(v)$, $v \in \mathrm{Int}(\Delta)$ have the same set of basic cycles. Let H denote the convex hull of their barycenters. The next two lemmas prove the theorem.

LEMMA 8.3. *H contains a global attractor of the flow Φ.*

PROOF. First remark that $\pi(v) \in H$. Indeed, let Y^v denote the Markov chain whose transition matrix is $M(v)$. With probability 1, there exists $m \in \mathbf{N}$ such that for all $l \ge m$, Y_l^v is in the recurrence class of $M(v)$ and by the ergodic theorem,

$$\lim_{n\to\infty} \frac{\sum_{l=0}^{n-1} e_{Y_{m+l}^v}}{n} = \pi(v).$$

Thus, by Lemma 8.2, $\pi(v) \in H$.

Define a function $G: \Delta \to \mathbf{R}_+$ by $G(x) = \inf\{\|x - w\|: w \in H\}$. By compactness and convexity of H, G is convex and continuous. Therefore, it admits a right derivative $DG(x)$ at any point $x \in \Delta$, which is convex and continuous.

Let $0 \le s \le 1$. By convexity,

$$\frac{G((1-s)x + s\pi(x)) - G(x)}{s} \le -G(x) + G(\pi(x)) = -G(x),$$

where the last equality stands because $\pi(x) \in H$. Thus, by taking the limit $s \to 0^+$, we get

$$DG(x)F(x) \le -G(x).$$

Hence, setting $x = \Phi_t(v)$, gives

$$\frac{d}{dt}\big(G(\Phi_t(v)\big) \le -G\big(\Phi_t(v)\big).$$

It follows that

$$d\big(\Phi_t(v), H\big) \le e^{-t} d(v, H)$$

for every $v \in \Delta$. This proves that H contains every attractor of F. In particular it contains the set $A = \bigcap_{t \ge 0} \Phi_t(\Delta)$ which is a global attractor. □

LEMMA 8.4. *Let $v \in \mathrm{Int}(\Delta)$, $y \in E$. Then $H \subset S^+(v, y)$.*

PROOF. First suppose that condition (a) holds. Let $w \in H$, $y \in E$. According to Lemma 8.2 there exists a sequence $\hat{z}_n = (z_{0,n}, \dots, z_{|\hat{z}|-1,n}) \in \Sigma_G$, with $\lim_{n\to\infty} \nu(\hat{z}_n) = w$ and $\lim_{n\to\infty} |\hat{z}_n| = \infty$. By indecomposability, there exists $\hat{a} \in \Sigma_G$ such that $a_0 = y$, $a_{|\hat{a}|-1} \to z_{0,n}$ and $|\hat{a}| \le d$. Set $\hat{y}_n = \hat{a}\hat{z}_n$.

Since $G \subset G_M(v)$ the sequence defined by $x_{0,n} = v$, and

$$x_{l+1,n} - x_{l,n} = \frac{1}{l+1+n_0}\left(-x_{l,n} + e_{y_{l,n}}\right), \qquad l = 0, \dots, |\hat{y}_n| - 1$$

is an admissible sequence for which $x_{|\hat{y}^n|,n} \to w$, $n \to \infty$. Thus, $w \in S^+(v, y)$.

Now, suppose that condition (b) holds. For all $v \in \Delta$ the graphs $G_M(v)$ have the same class of recurrence that we denote $R \subset E$.

CLAIM. *There exists $N \in \mathbf{N}$ such that for all $n_0 \ge N$, $x \in \Delta$ the two following statements are true.*

(i) *If $i \in E$ and $j \in R$ there exists an admissible sequence $x_0, \dots, x_k \in \Delta$ of length $k + 1 \le d$ with initial point $x_0 = x$, initial state $i_0 = i$ and terminal state $i_k = j$.*

(ii) *If $C = \{i_0, \dots, i_{n-1}\}$ is a basic cycle and $i \in C$, there exists a permutation σ: $\{0, \dots, n-1\} \to \{0, \dots, n-1\}$ with $i_{\sigma(0)} = i$ such that the sequence defined by*

$$x_0 = v, \qquad x_{l+1} - x_l = \frac{1}{l+1+n_0}\left(-x_l + e_{i_{\sigma(l)}}\right), \qquad l = 0, \dots, n-1$$

is admissible.

PROOF OF THE CLAIM. By compactness of Δ and continuity of M there exists $\delta > 0$ such that for every $v \in \Delta$ and every pair $(i, j) \in E \times R$, it is possible to find a sequence of states $(i_0, i_1, \dots, i_{k-1}, i_k)$, $i_0 = i$, $i_k = j$, $k \le d$, such that $\|w - v\| < \delta$ implies $\prod_{l=0}^{k-1} M_{i_l, i_{l+1}}(w) > 0$. Now, for $n_0 > (2d/\delta)$ the sequence defined by

$$x_0 = v, \qquad x_{l+1} - x_l = \frac{1}{l+1+n_0}\left(-x_l + e_{i_l}\right), \qquad l = 0, \dots, k-1$$

satisfies the required properties in part (i). Part (ii) is obtained by similar arguments.

Now, let $C_1, \dots, C_k$ denote the basic cycles of the graphs. Let $w \in H$, $w = \sum_i t_i B(C_i)$, $t_i \geq 0$, $\sum_i t_i = 1$. Fix $\varepsilon > 0$. Chooses integers q_i such that

$$\left| \frac{q_i \operatorname{card}(C_i)}{\sum_j q_j \operatorname{card}(C_j)} - t_i \right| < \varepsilon.$$

Given $v \in \Delta$, $y \in E$, we construct an admissible sequence $x_0, x_1, \dots,$ as follows: $x_0 = v$ and $x_1, \dots, x_N$ are arbitrarily chosen. Then we use part (i) to construct an admissible sequence with initial point x_N and whose terminal state is in C_1, then use p_1 times part (ii), then use again part (i) and so on. This gives an admissible sequence of finite length whose terminal point is in a neighborhood of v whose diameter is of the order of ε. This implies that $w \in S^+(v, y)$. $\square$

9. Pemantle's conjecture. In this section we use the results of Sections 3 and 6 to construct a counterexample to Pemantle's conjecture. We suppose that M is given by formulas (1) and (2) and R is a matrix with positive entries. Our first goal is to give an explicit expression for the vector field F.

The vector field. Set $h_i(v) = \alpha_i(v) N_i(v)$ where $N_i(v)$ is given by (1) and $\alpha(v)$ is a positive vector. We can always write the invariant measure of $M(v)$ as

$$\pi_i(v) = \frac{v_i h_i(v)}{H(v)},$$

where

$$H(v) = \sum_i v_i h_i(v).$$

From the relation

$$\sum_{k \neq i} M_{k,i}(v) \pi_k(v) = \sum_{k \neq i} M_{i,k}(v) \pi_i(v),$$

which characterizes $\pi(v)$, we deduce

$$\sum_{k \neq i} R_{k,i} v_i \big(v_k \alpha_k(v) \big) = \sum_{k \neq i} R_{i,k} v_k \big(v_i \alpha_i(v) \big).$$

To compute $\alpha(v)$ we will use a lemma due to Freidlin and Wentzell (1984). Given $i \in E$, a graph consisting of arrows $m \to n$ (with m, $n \in E$, $m \neq i$, $m \neq n$,) is called an *i-graph* if the graph has no cycle and for every $m \in E$ $(m \neq i)$ is the origin of exactly one arrow. The set of all i-graphs is denoted $G(i)$. Using Lemma 3.1, Chapter 6 of Freidlin and Wentzell (1984), $\alpha_i(v)$ can be expressed as

$$v_i \alpha_i(v) = \sum_{g \in G(i)} \alpha(g),$$

where

$$\alpha(g) = \prod_{m \to n \in g} R_{m,n} v_n.$$

Notice that for $g \in G(i)$, v_i always divides $\alpha(g)$. Thus the formula given for $\alpha_i(v)$ defines $\alpha_i(v)$ as an homogeneous polynomial of degree $d-2$ in variables v_i. Now if we multiply the vector field F by the positive function H we see that the dynamic of F is equivalent to the dynamic induced on Δ by the following differential system:

$$\frac{dv_i}{dt} = v_i\big(h_i(v) - H(v)\big), \qquad i = 1, \dots, d \tag{10}$$

with

$$h_i(v) = \alpha_i(v)\, N_i(v) = \Big(\sum_k R_{i,k} v_k\Big)\Big(\frac{1}{v_i} \sum_{g \in G(i)} \prod_{m \to n \in g} R_{m,n} v_n\Big). \tag{11}$$

A counterexample. Suppose $d = 3$ and let

$$R = \begin{bmatrix} 2 & 1 & \lambda \\ \lambda & 2 & 1 \\ 1 & \lambda & 2 \end{bmatrix}$$

where $2 < \lambda < 3$.

Using formula (11) we find

$$h_1(v) = (2v_1 + v_2 + \lambda v_3)\big(\lambda v_1 + \lambda^2 v_2 + v_3\big),$$
$$h_2(v) = (\lambda v_1 + 2v_2 + v_3)\big(v_1 + \lambda v_2 + \lambda^2 v_3\big),$$
$$h_3(v) = (v_1 + \lambda v_2 + 2v_3)\big(\lambda^2 v_1 + v_2 + \lambda v_3\big).$$

Consider the flow induced by (10) on the face $\partial_3\Delta$. Putting $v_3 = 0$ and $v_2 = 1 - v_1$, we can use $v_1 \in [0,1]$ as a single variable and the dynamic on $\partial_3\Delta$ is equivalent to the single equation

$$\begin{aligned} \frac{dv_1}{dt} &= v_1\big(h_1(v_1, 1-v_1, 0) - H(v_1, 1-v_1, 0)\big) \\ &= v_1(1-v_1)\big((1-\lambda(v_1+\lambda)(2v_1+\lambda-2)\big). \end{aligned} \tag{12}$$

The constraint $\lambda > 2$ shows that (12) admits only two equilibria given by $v_1 = 0$ and $v_1 = 1$. Hence $C \cap \partial_3\Delta = \{e_1, e_2\}$. By permuting cyclically we get

$$C \cap \partial\Delta = \{e_1, e_2, e_3\}.$$

However, we will prove the following result which disproves Pemantle's conjecture.

PROPOSITION 9.1. *For all $v \in \mathrm{Int}(\Delta)$ and for all $y \in E$,*

$$P\big(L(\{v_n\}) = \partial\Delta | v(0) = v, Y_0 = y\big) > 0.$$

From equation (12) and the constraint $\lambda > 2$, we see that if $x \in \partial_3\Delta - \{e_1, e_2\}$ then $\lim_{t\to\infty} \Phi_t(x) = e_1$ and $\lim_{t\to-\infty} \Phi_t(x) = e_2$. Similarly, if $x \in \partial_i\Delta - \{e_{i+1}, e_{i+2}\}$, then $\lim_{t\to\infty} \Phi_t(x) = e_{i+1}$ and $\lim_{t\to-\infty} \Phi_t(x) = e_{i+2}$ where indices are counted mod 3. From this we deduce the following property (see Figure 2).

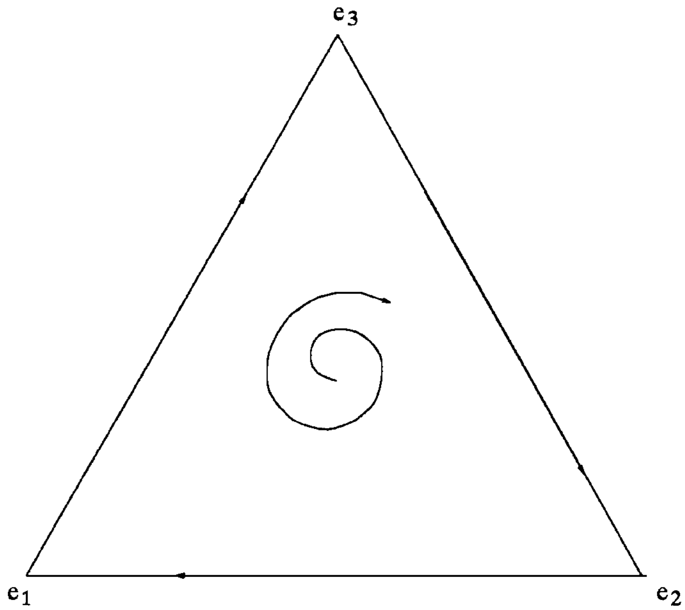

FIG. 2. $\lambda < \lambda^*$.

LEMMA 9.2. *The connected internally chain recurrent subsets of* $\partial\Delta$ *are the equilibria* $\{e_1\}, \{e_2\}, \{e_3\}$ *and the set* $\partial\Delta$.

For $x \in \Delta$ we let $DF(x)\colon T\Delta \to T\Delta$ denote the differential of F at x. If x is a zero of F (i.e., an equilibrium), x is said *hyperbolic* if eigenvalues of $DF(x)$ have nonzero real parts. Let $x \in \Delta$ be an hyperbolic equilibrium; x is called a *sink* if eigenvalues of $DF(x)$ have negative real parts, a *source* if they have positive real parts, and a *saddle* otherwise. From equation (12) we see that the eigenvalues of $DF(e_i)$ are

$$\rho_- = -\lambda, \qquad \rho_+ = \lambda(\lambda - 2).$$

Thus e_i is an hyperbolic saddle point. Also, the stable manifold of e_i is

$$W^s(e_i) = \left\{x \in \Delta\colon \lim_{t\to\infty} \Phi_t(x) = e_i\right\} = \partial_{i-1}\Delta$$

and the unstable manifold is

$$W^u(e_i) = \left\{x \in \Delta\colon \lim_{t\to-\infty} \Phi_t(x) = e_i\right\} = \partial_{i+1}\Delta.$$

LEMMA 9.3. *The set* $\partial\Delta$ *is an attractor.*

PROOF. Let $\varepsilon > 0$. Define points p_i^s, p_i^u as follows: $p_i^s \in W^s(e_i)$, $p_i^u \in W^u(e_i)$ and $d(p_i^s, e_i) = d(p_i^u, e_i) = \varepsilon$.

Let Σ_i^s (respectively, Σ_u^s) denote a compact interval (i.e., a one-dimensional convex set) parallel to $W^u(e_i)$ [respectively, $W^s(e_i)$] of length ε with p_i^s (respectively, p_i^u) as endpoint. It will be useful sometimes to identify Σ_i^s (respectively, Σ_u^s) with the interval $[0, \varepsilon]$ and p_i^s (respectively, p_i^u) with 0. When we use this identification, we write $\Sigma_i^s \simeq [0, \varepsilon]$. Σ_i^s and Σ_i^u are local sections through p_i^s and p_i^u transversal to the vector field at these points. Thus (by choosing ε small enough) we can suppose that the orbit through every point of $\Sigma_i^s - \{p_i^s\}$ intersects transversally Σ_i^u. This defines a Poincaré map $g_i\colon \Sigma_i^s - \{p_i^s\} \to \Sigma_i^u - \{p_i^u\}$, $p \to \Phi_\tau(p)$ where $\tau = \inf\{t > 0\colon \Phi_t(p) \in \Sigma_i^u\}$. Also, there exists a smaller compact interval $I_i^u \subset \Sigma_i^u$ and a smooth Poincaré map $h_i\colon I_i^u \to \Sigma_{i-1}^s$, $p \to \Phi_\tau(p)$ where $\tau = \inf\{t > 0\colon \Phi_t(p) \in \Sigma_{i-1}^s\}$.

CLAIM. *There exists $\eta > 0$ and $J_i^u \subset I_i^u$ ($J_i^u \simeq [0, \eta]$) such that (under the identifications $J_{i+1}^u \simeq [0, \eta]$ and $\Sigma_i^s \simeq [0, \varepsilon]$) $p \in J_{i+1}^u$ implies*

$$0 < g_i \circ h_{i+1}(p) \le \tfrac{1}{2} p.$$

PROOF OF THE CLAIM. A natural coordinate system around e_i is given by $x \in \Sigma_i^s \simeq [0, \varepsilon]$, $y \in \Sigma_i^u \simeq [0, \varepsilon]$. Expressed in this coordinate system, the dynamic of F can be rewritten as

$$\frac{dx}{dt} = x(\rho_+ + f_1(x, y))$$

$$\frac{dy}{dt} = y(\rho_- + f_2(x, y)),$$

where f_1 and f_2 are smooth functions such that $f_1(0, y) = f_2(x, 0) = 0$. Now, notice that the constraint $2 < \lambda < 3$ implies $0 < \rho_+ < -\rho_-$. Therefore for ε small enough it is possible to find $0 < \tilde{\rho}_+ < -\tilde{\rho}_-$ such that

$$\frac{dx}{dt} \le x\tilde{\rho}_+,$$

$$\frac{dy}{dt} \le y\tilde{\rho}_-.$$

Integrating this inequality from 0 to τ with the initial condition $x(0) = x$, $y(0) = \varepsilon$ and the constraint $x(\tau) = \varepsilon$, gives $x(\tau) = \varepsilon$ and $y(\tau) \le \varepsilon(x/\varepsilon)^\alpha$ where $\alpha = |\tilde{\rho}_-|/\tilde{\rho}_+$. Thus, under the identification $\Sigma_i^u \simeq [0, \varepsilon]$, we get $g_i(p) \le \varepsilon(p/\varepsilon)^\alpha$.

Let L denote a Lipschitz bound for the maps h_i, $i = 1, 2, 3$. For every $p \in I_{i+1}^u$, we have $g_i(h_{i+1}(p)) \le L^\alpha \varepsilon(p/\varepsilon)^\alpha$. Hence, setting

$$\eta = \varepsilon(1/2L^\alpha)^{1/(\alpha-1)}$$

proves the claim.

Let $P\colon J_3^u \to J_3^u$ be a map $P = (g_3 \circ h_1) \circ (g_1 \circ h_2) \circ (g_2 \circ h_3)$. Let q be the endpoint of J_3^u identified with η and let $q' = P(q)$. The point q' belongs to the forward trajectory of q, hence $q' = \Phi_t(q)$ for some $t > 0$. By Jordan curve theorem, the curve $\{\Phi_s(p)\colon 0 \le s \le t\} \cup [q, q']$ separates Δ. Let U be the connected component which contains $\partial\Delta$.

The claim shows that P is a contraction from which it follows that $\lim_{t\to\infty} d(\Phi_t(x), \partial\Delta) = 0$ uniformly in $x \in U$. This proves that $\partial\Delta$ is an attractor. □

The next lemma is Pemantle's Theorem 1.3 (1992) restated for nonsymmetric matrices. Since the proof given by Pemantle adapts easily to this case we do not reproduce it but refer the reader to Pemantle's paper.

LEMMA 9.4. *Let $p \in \Delta$ be an equilibrium of F [i.e., $\pi(p) = p$]. Suppose that $p \in \partial_i\Delta$ and $h_i(p) > H(p)$ for some $i \in E$. Then $P(\{L(v_n)\} = p) = 0$.*

PROOF OF PROPOSITION 9.1. Corollary 6.5 is applicable; thus, since $\partial\Delta$ is an attractor (Lemma 9.3) we have $P(L(\{v_n\}) \subset \partial\Delta) > 0$. According to Corollary 3.5 and Lemma 9.2 the event $\{L(\{v_n\}) \subset \partial\Delta)\}$ decomposes as

$$\{L(\{v_n\}) \subset \partial\Delta\} = \bigcup_{i=1}^{3} \{L(\{v_n\}) = e_i\} \cup \{L(\{v_n\}) = \partial\Delta\},$$

but $e_i \in \partial_{i-1}\Delta$ and $h_{i-1}(e_i) - H(e_i) = \lambda^2 - 2\lambda > 0$. Thus by Lemma 9.4 the events $\{L(\{v_n\}) = e_i\}$ have zero probability and the result follows. □

REMARKS.

(i) Notice that the point $v^* = (1/3, 1/3, 1/3)$ is always an equilibrium of the vector field F. A computation of the Jacobian matrix $DF(v^*)$ gives

$$\operatorname{Tr}(DF(v^*)) = -\frac{\lambda(-7 - \lambda + 2\lambda^2)}{9}$$

and

$$\operatorname{Det}(DF(v^*)) = \frac{(\lambda^2 - 3\lambda + 3)(\lambda^2 + \lambda + 1)^2}{81} > 0.$$

From this we see that there is a critical value

$$\lambda^* = \frac{(1 + \sqrt{57})}{4} \simeq 2.137$$

at which the stability of v^* changes.

(ii) When $2 < \lambda < \lambda^*$, v^* is a source. Thus, by a result of Brandiere and Duflo (1996) or by adapting the proof of Pemantle's Theorem 1.2, v^* has zero probability to be the limit set of $\{v_n\}$. When λ passes the critical value λ^*, it is easy to verify that a Hopf's bifurcation occurs: the equilibrium v^* becomes a sink and gives rise to a unstable hyperbolic periodic orbit (see Figures 2 and 3). By adapting the proof of Theorem 2.1 of Benaïm and Hirsch (1995), it can be proved that this periodic orbit has zero probability to be the limit set of $\{v_n\}$ while in virtue of Corollary 6.5 both $\partial\Delta$ and v^* have a positive probability to be this limit set. From these remarks and numerical simulations it seems reasonable to think (although we did not prove it) that for $2 < \lambda < \lambda^*$:

$$P(L\{v_n\} = \partial\Delta) = 1$$

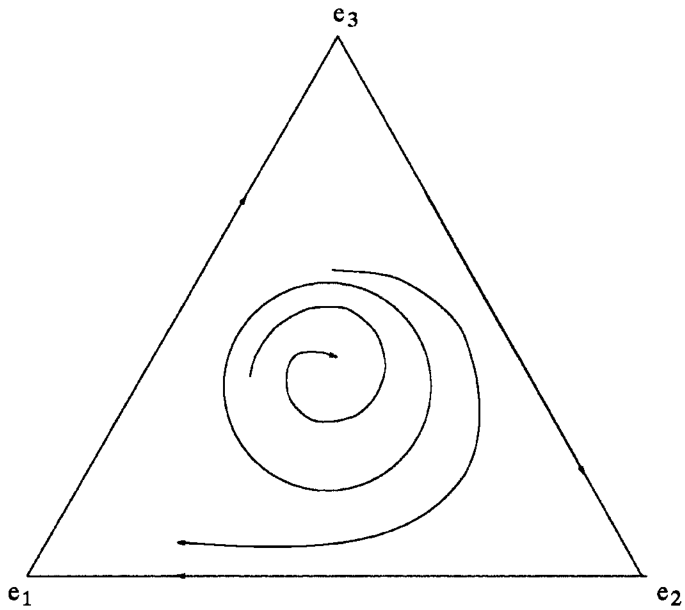

FIG. 3. $\lambda > \lambda^*$.

while

$$P\big(L\{v_n\} = \partial\Delta\big) + P\big(L\{v_n\} = v^*\big) = 1$$

when $3 > \lambda > \lambda^*$.

REFERENCES

ARTHUR, B., ERMOL'EV, Y. and KANIOVSKII, Y. (1983). A generalized urn problem and its applications. *Cybernetics* **19** 61–71.

BENAÏM, M. (1996). A dynamical system approach to stochastic approximations. *Siam J. Control Optim.* **34** 437–472.

BENAÏM, M. and HIRSCH, M. W. (1995). Dynamics of Morse–Smale urn processes. *Ergodic Theory Dynamical Systems* **15** 1005–1030.

BENAÏM, M. and HIRSCH, M. W. (1996). Asymptotic pseudo-trajectories and chain-recurrent flows with applications. *Journal of Dynamics and Differential Equations* **8** 141–176.

BOWEN, R. (1975). ω-Limit sets for axiom *A* diffeomorphisms. *J. Differential Equations* **18** 333–339.

BRANDIERE, O. and DUFLO, M. (1996). Les algorithmes stochastiques contournent ils les pieges? *Ann. Instit. H. Poincaré* **32** 395–427.

CONLEY, C. C. (1978). *Isolated Invariant Sets and the Morse Index*. Amer. Math. Soc., Providence, RI.

COPPERSMITH, D. and DIACONIS, P. (1986). Random walk with reinforcement. Unpublished manuscript.

DIACONIS, P. (1988). Recent progress on de Finetti's notions of exchangeability. *Bayesian Statistics* **3** 111–125.

FREIDLIN, M. I. and WENTZELL, A. D. (1984). *Random Perturbation of Dynamical Systems*. Springer, New York.

HILL, B. M., LANE, D. and SUDDERTH, W. (1980). A strong law for some generalized urn processes. *Ann. Probab.* **8** 214–226.

HIRSCH, M. W. (1994). Asymptotic phase, shadowing and reaction-diffusion systems. In *Control Theory, Dynamical Systems and Geometry of Dynamics* (K. D. Elworthy, W. N. Everitt and E. B. Lee, eds.). Dekker, New York.

LOSERT, V. and AKIN, E. (1983). Dynamics of games and genes: discrete versus continuous time. *J. Math. Bio.* **17** 241–251.

MÉTIVIER, M. and PRIOURET, P. (1987). Thèorèmes de convergence presque sure pour une classe d'algorithmes stochastiques à pas décroissants. *Probab. Theory Related Fields* **74** 403–438.

PEMANTLE, R. (1990). Nonconvergence to unstable points in urn models and stochastic approximations. *Ann. Probab.* **18** 698–712.

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

SCHREIBER, S. (1995). Expansion rate and Lyapounov exponents. Unpublished manuscript.

DEPARTEMENT DE MATHÉMATIQUES
LABORATOIRE DE STATISTIQUE ET PROBABILITÉS
UNIVERSITÉ PAUL SABATIER
118 ROUTE DE NARBONNE
31062 TOULOUSE CEDEX
FRANCE
E-MAIL: benaim@cict.fr