Simulated Annealing, Vertex-Reinforced Random Walks and Learning in Games*

Michel Benaïm
Université de Neuchâtel, Suisse

Olivier Raimond
Université Paris Sud, France

October 8, 2007

Abstract

This paper studies a class of non Markovian and non homogeneous stochastic processes on a finite state space. It provides an unified approach to *simulated annealing* type processes, certain *vertex reinforced random walks* and certain models of learning in games including *Markovian fictitious play.*

Keywords: Stochastic approximation, Processes with reinforcement, Differential Inclusions, Learning in Games, Simulated Annealing

1 Introduction

Let E be a finite set called *the state space,* $\mathsf{M} = \mathsf{M}(E)$ the set of Markov matrices over E, and Σ a compact convex subset of an Euclidean space called *the observation space.* Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability space equipped with an increasing sequence of sub σ-fields $\{\mathcal{F}_n, n \in \mathbb{N}\} : \mathcal{F}_n \subset \mathcal{F}_{n+1} \subset \mathcal{F}$.

Our main object of interest is a discrete time random process $(X, M, V) = ((X_n, M_n, V_n))$ defined on $(\Omega, \mathcal{F}, \mathsf{P})$ taking values in $E \times \mathsf{M}(E) \times \Sigma$ such that:

*We acknowledge financial support from the Swiss National Science Foundation grant 200020-112316 .

(i) (X, M, V) is *adapted* (to $\{\mathcal{F}_n, n \in \mathbb{N}\}$), meaning that (X_n, M_n, V_n) is $\mathcal{F}_n$-measurable for each n.

(ii) For all $y \in E$

$$\mathsf{P}(X_{n+1} = y | \mathcal{F}_n) = M_n(X_n, y). \tag{1}$$

We refer to X_n (respectively V_n) as the *state* (respectively, the *observation*) variable at time n; and to the sequence (M_n) as the *strategy.* We let

$$v_n = \frac{1}{n} \sum_{i=1}^{n} V_i$$

denote the empirical average up to time n of the sequence of observations.

A well studied situation is when

$$M_n = K(v_n) \tag{2}$$

where K maps continuously probability vectors to irreducible Markov matrices. In such a case (X_n) is called a "Markov chain controlled" by (v_n) and the behavior of (v_n) can be analyzed through the ODE

$$\dot{v} = -v + \pi(v) \tag{3}$$

where $\pi(v)$ is the invariant probability of $K(v)$. This approach to controlled Markov chains goes back to the work of Métivier and Priouret (1987) (see also the books Benveniste, Métivier and Priouret (1990), Duflo (1996)) strongly influenced by the pioneered works of Ljung (1977), Kushner and Clark (1978) on the ODE's method. It has been used in Benaïm (1997) for analyzing certain vertex reinforced random walks on finite graphs.

The main purpose of this paper is to investigate the long term behavior of (v_n) under less stringent assumptions than (2). In particular we are interested in situations where:

(a) M_n may depend on other (non-observable or hidden) variables than v_n and;

(b) The closure of $\{M_n \ : n \geq 0\}$ may contain degenerate (i.e non irreducible) Markov matrices.

Situation (a) typically occurs in game theory where players may have only partial information on the actions played by their opponents, and (b) is motivated by stochastic optimization algorithms.

Relying on a recent paper by Benaïm, Hofbauer and Sorin (2005) it will be shown that under certain assumptions (involving estimates on the log-Sobolev and spectral gap constants of (M_n)) the aymptotic behavior can be described in term of a certain set-valued deterministic dynamical system that generalizes the ODE (3). Applications to non-homogeneous Markov chains, vertex reinforced random walks and learning processes in game theory will be given.

Outline of contents

The organization of the paper is as follows. Section 2 states the notation, hypotheses and the main result. Our main assumption (Hypothesis 2.1) is somewhat abstract and more tractable conditions (expressed in term of spectral gaps and log-Sobolev constants) are given in section 3. Section 4 is devoted to examples and applications. The proof of the main result is postponed to section 5.

2 Notation, hypotheses and main results

A *probability vector* (or measure) over E is a map $\mu : E \to \mathbb{R}^+$ such that $\sum_x \mu(x) = 1$, and a *Markov matrix* is a map $M : E \times E \to \mathbb{R}^+$, such that

$$\forall x \in E, \sum_y M(x,y) = 1.$$

We let $\Delta = \Delta(E)$ denote the space of probability vectors over E and $\mathsf{M} = \mathsf{M}(E)$ denote the set of Markov matrices on E.

Given a function $f : E \to \mathbb{R}$ and $\mu \in \Delta$ we use the notation

$$\mu f = \sum_x \mu(x) f(x).$$

A Markov matrix M on E acts on functions f and measures μ according to the formulas

$$Mf(x) = \sum_y M(x,y) f(y),$$

$$\mu M(y) = \sum_x \mu(x) M(x, y).$$

We let M^n denote the Markov matrix obtained by matrix multiplication. Equivalently $M^n f = M(M^{n-1} f)$ for $n \geq 1$, with the convention that $M^0 f = f$.

Points $x, y \in E$ are said to be *related* if there exist $i, j \geq 0$ (depending on x and y) such that $M^i(x, y) > 0$ and $M^j(y, x) > 0$. An equivalence class for this relation is called a *recurrent class.* The Markov matrix M on E is said *indecomposable* if it has a unique recurrent class (possibly periodic) and is said *irreducible* if this recurrent class is E.

By standard results, indecomposability of M implies that M possesses a unique *invariant probability measure* π characterized by the relation $\pi M = \pi$. Moreover, the *generator* $L = -I + M$ has kernel $\mathbb{R}1$ and its restriction to $\{f : \pi f = 0\}$ is an isomorphism. It then follows that $-L$ admits a pseudo "inverse" Q characterized by

$$Q\mathbf{1} = 0,$$

and

$$Q(I - M) = (I - M)Q = I - \Pi;$$

where $\Pi \in \mathsf{M}$ denote the matrix defined by $\Pi(x, y) = \pi(y)$. To shorten notation we also call Q the pseudo inverse of M.

Given a vector f and a matrix N, we set $|f| = \max|f(x)|$ and $|N| = \max_{x,y} |N(x, y)|$.

Our main assumption is the following:

Hypothesis 2.1 *The matrices (M_n) are indecomposable and their pseudo inverses (Q_n) and invariant probabilities (π_n) satisfy*

(i)

$$\lim_{n \to \infty} \frac{|Q_n|^2 \log(n)}{n} = 0,$$

(ii)

$$\lim_{n \to \infty} |Q_{n+1} - Q_n| = 0,$$

(iii)

$$\lim_{n \to \infty} |\pi_{n+1} - \pi_n| = 0.$$

The verification of hypothesis 2.1 is the subject of section 3 where sufficient and more tractable conditions will be detailed.

Let $\hat{V}_n : E \to \Sigma$ be an $\mathcal{F}_n$-measurable map defined by

$$\hat{V}_n(x) = \frac{\mathsf{E}(V_{n+1}\mathbf{1}_{X_{n+1}=x}|\mathcal{F}_n)}{M_n(X_n, x)}$$

for $M_n(X_n, x) \neq 0$. In addition to hypothesis 2.1 we assume that

Hypothesis 2.2

$$\lim_{n\to\infty} M_{n+1}Q_{n+1}(\hat{V}_{n+1} - \hat{V}_n) = 0.$$

Remark 2.3 Here are some sufficient conditions ensuring hypothesis 2.2.

(i) Assume that $x \mapsto \hat{V}_{n+1}(x) - \hat{V}_n(x)$ is a constant map. Then hypothesis 2.2 holds since $Q_n 1 = 0$. This will be used in section 4.

(ii) More generally, let $T\Sigma$ be the affine hull of Σ (the smallest affine space containing Σ). Assume that for all $n \in \mathbb{N}$ there exists a vector $A_n \in T\Sigma$ and a map $B_n : E \to T\Sigma$ such that

- **(a)** For all $x \in E$, $\hat{V}_{n+1}(x) - \hat{V}_n(x) = A_n + B_n(x)$
- **(b)** $\limsup_{n\to\infty} |B_n|\sqrt{\frac{n}{\log(n)}} < \infty$.

Then $|M_{n+1}Q_{n+1}((\hat{V}_{n+1} - \hat{V}_n))| = |M_{n+1}Q_{n+1}B_n| \leq |Q_{n+1}||B_n| \to 0$ by hypothesis 2.1.

(iii) Assume that $M_n(x, y) = \pi_n(y)$. Then $M_{n+1}Q_{n+1} = 0$ so that hypothesis 2.2 holds.

2.1 Adapted set-valued dynamical systems

The purpose of this section is to introduce certain differential inclusions on Σ that will prove to be useful for analyzing the long term behavior of (v_n). Recall that we let π_n denote the invariant probability of M_n. Let

$$\theta_n = \pi_n \hat{V}_n = \sum_x \pi_n(x)\hat{V}_n(x). \tag{4}$$

We let $C_n \subset \Sigma \times \Sigma$ denote the *topological support* of the law of (v_n, θ_n). That is the smallest closed set $F \subset \Sigma \times \Sigma$ such that

$$\mathsf{P}((v_n, \theta_n) \in F)) = 1.$$

Let $\mathsf{clos}\{C_n\}$ denote the set of all possible limit points $z = \lim z_{n_k}$ with $z_{n_k} \in C_{n_k}$ and $n_k \to \infty$. It is easily seen that $\mathsf{clos}\{C_n\}$ is a nonempty compact subset of $\Sigma \times \Sigma$.

A nonempty set $G \subset \Sigma \times \Sigma$ is called a *graph* (or a bundle) over Σ, if the projection

$$p : G \to \Sigma,$$
$$(u, v) \mapsto u$$

is onto. A graph G over Σ defines a *set-valued function* mapping each point $u \in \Sigma$ to a set $G(u) = \{v \in \Sigma\ : (u, v) \in G\}$.

Definition 2.4 *A set $C \subset \Sigma \times \Sigma$ is said to be* adapted *to* $\{(v_n, \theta_n)\}$ *(or simply adapted) if*

(i) *C is a closed graph over Σ.*

(ii) *For all $u \in \Sigma$, $C(u)$ is a nonempty convex set.*

(iii) $\mathsf{clos}\{C_n\} \subset C$.

To an adapted set C we associate the differential inclusion

$$\dot{v} \in -v + C(v). \tag{5}$$

A solution to (5) is an absolutely continuous mapping $v : \mathbb{R} \to \Sigma$ verifying $\dot{v}(t) + v(t) \in C(v(t))$ for almost every t. A set $A \subset \Sigma$ is said to be *invariant* if for all $x \in A$ there exists a solution $\mathbf{x}$ to (5) with $\mathbf{x}(0) = x$ and such that $\mathbf{x}(\mathbb{R}) \subset A$.

Given a set $A \subset \Sigma$ and $(x, y) \in A^2$ we write $x \hookrightarrow_A y$ if for every $\varepsilon > 0$ and $T > 0$ there exists an integer $n \in \mathbb{N}$, solutions $\mathbf{x}_1, \dots \mathbf{x}_n$ to (5) and real numbers $t_1, t_2, \dots, t_n$ greater than T such that

(a) $\mathbf{x}_i([0, t_i]) \subset A$,

(b) $\|\mathbf{x}_i(t_i) - \mathbf{x}_{i+1}(0)\| \leq \varepsilon$ for all $i = 1, \dots, n-1$,

(c) $\|\mathbf{x}_1(0) - x\| \leq \varepsilon$ and $\|\mathbf{x}_n(t_n) - y\| \leq \varepsilon$.

Definition 2.5 *A set $A \subset \Sigma$ is said to be* internally chain transitive *provided A is compact and $x \hookrightarrow_A y$ for all $x, y \in A$.*

It is not hard to verify (see e.g Benaïm, Hofbauer and Sorin (2005) Lemma 3.5) that an *internally chain transitive* set is invariant.

The *limit set* of (v_n) is the set $L = L((v_n))$ consisting of all points $p = \lim v_{n_k}$ for some sequence $n_k \to \infty$. The next theorem 2.6 is the main result of the paper. Its proof heavily relies on Benaïm, Hofbauer and Sorin (2005) and is given in section 5.

Theorem 2.6 *Assume that hypotheses 2.1 and 2.2 hold. Let C be an adapted graph. Then the limit set of (v_n) is an internally chain transitive set for the differential inclusion*

$$\dot{v} \in -v + C(v).$$

2.2 Background : How to use Theorem 2.6

The notion of "internally chain transitive set" was introduced by Benaïm and Hirsch (1996) in order to analyze the long term behavior of certain perturbations of flows and has been recently extended to multivalued dynamical systems by Benaïm Hofbauer and Sorin (2005). We refer the reader to this paper for more details, examples and properties. For convenience this section briefly reviews a few useful properties of internally chain transitive sets.

The differential inclusion (5) induces a set-valued dynamical system $\{\Phi_t\}_{t \in \mathbb{R}}$ defined by

$$\Phi_t(x) = \{\mathbf{x}(t) : \mathbf{x} \text{ is a solution to (5) with } \mathbf{x}(0) = x \in \Sigma\}.$$

A non empty compact set A is an *attracting* set if there exists a neighborhood U of A and a function $\mathbf{t}$ from $(0, \varepsilon_0)$ to $\mathbb{R}^+$ with $\varepsilon_0 > 0$ such that

$$\Phi_t(U) \subset A^{\varepsilon}$$

for all $\varepsilon < \varepsilon_0$ and $t \geq \mathbf{t}(\varepsilon)$, where A^{ε} stands for the $\varepsilon-$neighborhood of A. If additionally A is invariant, then A is an *attractor*.

Given an attracting set (resp. attractor) A, its *basin of attraction* is the set

$$B(A) = \{x \in \Sigma : \exists t \geq 0, \Phi_t(x) \in U\}.$$

When $B(A) = \Sigma$, A is a *globally* attracting set (resp. a global attractor).

Given a closed invariant set S, the induced dynamical system Φ^S on S is defined by

$$\Phi_t^S(x) = \{\mathbf{x}(t) : \mathbf{x} \text{ is a solution to (5) with } \mathbf{x}(0) = x \text{ and } \mathbf{x}(\mathbb{R}) \subset S\}.$$

An invariant set S is *attractor free* if there exists no proper subset A of S which is an attractor for Φ^S.

Throughout the remainder of this section we let L denote an internally chain transitive set (for instance the limit set $L = L(v_n)$). Properties of L will then be obtained through the next result (Benaïm, Hofbauer and Sorin (2005), Lemma 3.5, Proposition 3.20 and Theorem 3.23):

Proposition 2.7 **(i)** *The set $L = L((v_n))$ is non–empty, compact, invariant and attractor free.*

(ii) *If A is an attracting set with $B(A) \cap L \neq \emptyset$, then $L \subset A$.*

Some useful properties of attracting sets or attractors are the two following (Benaïm, Hofbauer and Sorin (2005), Propositions 3.25 and 3.27).

Proposition 2.8 *Let $\Lambda \subset \Sigma$ be compact with a bounded open neighborhood U and $V : \overline{U} \to [0, \infty[$. Assume the following conditions:*

(i) *$\Phi_t(U) \subset U$ for all $t \geq 0$,*

(ii) *$V^{-1}(0) = \Lambda$,*

(iii) *V is continuous and for all $x \in U \backslash \Lambda, y \in \Phi_t(x)$ and $t > 0$, $V(y) < V(x)$.*

Then Λ contains an attractor whose basin contains U.

The map V introduced in this proposition is called a *strong Lyapounov function* associated to Λ.

Let now Λ be a subset of Σ and $U \subset \Sigma$ an open neighborhood of Λ. A continuous function $V : U \to \mathbb{R}$ is called a *Lyapunov function* for $\Lambda \subset \Sigma$ if $V(y) < V(x)$ for all $x \in U \setminus \Lambda$, $y \in \Phi_t(x)$, $t > 0$; and $V(y) \leq V(x)$ for all $x \in \Lambda$, $y \in \Phi_t(x)$ and $t \geq 0$.

Proposition 2.9 (Lyapounov) *Suppose $V : U \to \mathbb{R}$ is a Lyapunov function for Λ and $L = L((v_n)) \subset U$. Assume that $V(\Lambda)$ has an empty interior. Then $L \subset \Lambda$ and the restriction of V to L is constant.*

3 Verification of hypothesis 2.1

This section is devoted to the verification of Hypothesis 2.1. The results given here will be used in section 4 to analyze specific situations.

3.1 Estimates based on compactness

Let $\mathsf{M}_{ind}(E)$ denote the open set of indecomposable Markov matrices.

Proposition 3.1 *Suppose that the sequence* (M_n) *lies in a compact subset of* $\mathsf{M}_{ind}(E)$ *and verifies* $\lim_{n\to\infty}(M_{n+1}-M_n)=0$. *Then hypothesis 2.1 holds.*

This proposition is a direct consequence of the next lemma.

Lemma 3.2 *Let* $T\mathsf{M}(E)$ *be the space of matrices* $K=K(x,y)$ *such that* $\sum_y K(x,y)=0$. *The map* $\mathsf{Q}:\mathsf{M}_{ind}(E)\to T\mathsf{M}(E)$ *which associates to* M *its pseudo inverse and the map* $\Pi:\mathsf{M}_{ind}(E)\to\Delta$ *which associates to* M *its invariant measure are smooth maps.*

Proof : Set $M\in\mathsf{M}_{ind}(E)$. The invariant probability of M, $\Pi(M)$, is solution to $\phi(M,\pi)=0$ where $\phi:\mathsf{M}_{ind}(E)\times\Delta\to T\Delta$, is the smooth map defined by

$$\phi(M,\mu)=\mu(I-M),$$

with $T\Delta=\{\mu:E\to\mathbb{R}\ :\sum_x\mu(x)=0\}$. For all $\nu\in T\Delta$,

$$\frac{\partial\phi}{\partial\mu}(M,\mu).\nu=\nu(I-M).$$

Hence, by uniqueness of the invariant probability measure, $\frac{\partial\phi}{\partial\mu}(M,\mu)$ has kernel $\{0\}$ and the fact that Π is smooth follows from the implicit function theorem.

We denote by $\hat{\Pi}(M)\in\mathsf{M}(E)$ the matrix defined by $\hat{\Pi}(M)(x,y)=\Pi(M)(y)$. The pseudo inverse of M is solution to $\psi(M,Q)=0$ where $\psi:\mathsf{M}_{ind}(E)\times T\mathsf{M}(E)\to T\mathsf{M}(E)$, is the smooth map defined by

$$\psi(M,Q)=Q(I-M)-(I-\hat{\Pi}(M)).$$

For all $A\in T\mathsf{M}(E)$

$$\frac{\partial\psi}{\partial Q}(M,Q).A=A(I-M).$$

Hence, by uniqueness of the invariant probability measure, $\frac{\partial \psi}{\partial Q}(M,Q)$ has kernel $\{0\}$ and the fact that Q depends smoothly on M follows from the implicit function theorem. **QED**

Let K be a continuous mapping from Γ a compact set into $\mathsf{M}(E)$ such that $K(w)$ is indecomposable for all $w \in \Gamma$. Assume (w_n) is a sequence of Γ-valued random variables such that $M_n = K(w_n)$. If in addition $\lim_{n\to\infty}(M_{n+1} - M_n) = 0$, then proposition 3.1 applies.

3.2 Estimates based on log-Sobolev and spectral gap constants

Propositions 3.3 and 3.4 below can be used to verify hypothesis 2.1 when the sequence (M_n) is not bounded away from $\mathsf{M}_{ind}(E)$. The strategy is then to verify assertions (ii) and (iii) of proposition 3.3 and to use the estimates given by proposition 3.4 to verify assertion (i).

Proposition 3.3 *Suppose that the matrices (M_n) are indecomposable and that their pseudo inverse (Q_n) and invariant probabilities (π_n) satisfy*

(i)

$$\lim_{n\to\infty} \frac{|Q_n|^2 \log(n)}{n} = 0,$$

(ii)

$$\limsup_{n\to\infty} |M_{n+1} - M_n| \frac{n}{\log(n)} < \infty$$

(iii)

$$\limsup_{n\to\infty} |\pi_{n+1} - \pi_n| \sqrt{\frac{n}{\log(n)}} < \infty$$

Then hypothesis 2.1 holds

Proof : The proof amounts to show that hypothesis 2.1 (ii) holds. Set $L_n = M_n - I$ and $\Pi_n = \hat{\Pi}(M_n)$. Using the characterization of Q_n one has

$$Q_{n+1}(L_{n+1} - L_n) + (Q_{n+1} - Q_n)L_n = \Pi_{n+1} - \Pi_n.$$

Hence,

$$Q_{n+1}(L_{n+1} - L_n)Q_n + (Q_{n+1} - Q_n)L_nQ_n = (\Pi_{n+1} - \Pi_n)Q_n.$$

That is (using $Q_n \Pi_n = Q_n \Pi_{n+1} = 0$ and $L_n Q_n = \Pi_n - I$)

$$Q_{n+1}(M_{n+1} - M_n)Q_n + (Q_n - Q_{n+1}) = (\Pi_{n+1} - \Pi_n)Q_n.$$

Therefore

$$|Q_n - Q_{n+1}| \leq c(|Q_{n+1}||Q_n||M_{n+1} - M_n| + |\pi_{n+1} - \pi_n||Q_n|),$$

for some constant $c > 0$ and conditions $(i), (ii), (iii)$ imply hypothesis 2.1 (ii). **QED**

Let $\mathsf{M}_{irr}(E)$ denote the open set of irreducible Markov matrices. Let $M \in \mathsf{M}_{irr}(E)$ with invariant probability π and let $f : E \to \mathbb{R}$. The *variance*, *entropy* and *energy* of f are respectively defined as

$$var(f) = \pi(f^2) - (\pi f)^2$$

$$\mathcal{L}(f) = \sum_x f(x)^2 \log\left(\frac{f(x)^2}{\pi f^2}\right)\pi(x)$$

$$\mathcal{E}(f) = \frac{1}{2}\sum_{x,y}(f(y) - f(x))^2 M(x,y)\pi(x).$$

The *spectral gap* and *log-Sobolev* constants of M are then defined to be

$$\lambda = \min\left\{\frac{\mathcal{E}(f)}{var(f)} \ : var(f) \neq 0\right\}$$

$$\alpha = \min\left\{\frac{\mathcal{E}(f)}{\mathcal{L}(f)} \ : \mathcal{L}(f) \neq 0\right\}.$$

The following estimates follows from the quantitative results for finite Markov chains as given in Saloff-Coste (1997) theorems.

Proposition 3.4 *Let $M \in \mathsf{M}_{irr}(E)$ with invariant probability π log-Sobolev constant α and spectral gap λ. For all $(x,y) \in E$ the following estimates hold:*

(i)

$$|Q(x,y)| \leq \sqrt{\frac{\pi(y)}{\pi(x)}}\frac{1}{\lambda}$$

(ii)

$$|Q(x,y)| \le \frac{1}{\alpha}\log_+\left(\log\left(\frac{1}{\pi(x)}\right)\right) + \frac{e}{\lambda}$$

where $\log_+(t) = \max(0, \log(t))$.

In particular

$$|Q| \le \frac{1}{\alpha}\left[\log_+\left(\log\left(\frac{1}{\pi_*}\right)\right) + \frac{e}{2}\right]$$

and

$$|Q| \le \frac{1}{\lambda}\left[\log_+\left(\log\left(\frac{1}{\pi_*}\right)\right)\frac{\log((1-\pi_*)/\pi_*)}{1-2\pi_*}) + e\right].$$

Proof : Let $L = -I + M$ and let $\{P_t\}$ be the continuous time semi-group $P_t = e^{tL}$. Then Q can be written as

$$Q(x,y) = \int_0^\infty (P_t(x,y) - \pi(y))dt.$$

The first assertion then easily follows from the estimate

$$|P_t(x,y) - \pi(y)| \le \sqrt{\frac{\pi(y)}{\pi(x)}}e^{-\lambda t}$$

whose proof can be found in Saloff-Coste (1997, Corollary 2.1.5).

We now pass to the second assertion. If $\pi(x) \ge e^2$ the inequality to be proved follows from inequality (i). Hence we assume that $\pi(x) < e^2$, and we follow the line of the proof of Theorem 2.2.5 in Saloff-Coste (1997). For $q \ge 1$, we let $||.||_q$ denotes the norm in $l^q(\pi)$. We let P_t^* denote the adjoint of P_t in $l^2(\pi)$, and $p_t(x,y) = p_t^*(y,x) = P_t(x,y)/\pi(y)$. Let g_x denote the function given by $g_x(y) = 0$ for $x \ne y$ and $g_x(x) = 1/\pi(x)$. Then

$$|P_t(x,y) - \pi(y)| \le ||p_t(x,.) - 1||_2 = ||(P_t^* - \pi)g_x||_2$$

Therefore

$$\begin{aligned} |P_{t+s}(x,y) - \pi(y)| &\le ||p_{t+s}(x,.) - 1||_2 \le ||P_t^* - \pi||_{2\to 2}||P_s^*\delta_x||_2 \\ &\le e^{-\lambda t}||P_s^*||_{k\to 2}||g_x||_k \end{aligned}$$

for any $k \ge 1$. where we have used the fact that $||P_t^* - \pi||_{2\to 2} \le e^{-\lambda t}$. Let q be the Hölder conjugate of k. Then $||P_s^*||_{k\to 2} = ||P_s||_{2\to q}$. Now choose

$q(s) = 1 - e^{2\alpha s}$. By hypercontractivity (see Theorem 2.2.4 in Saloff-Coste (1997)), $||P_s||_{2\to q(s)} \leq 1$ so that

$$|P_{t+s}(x,y) - \pi(y)| \leq e^{-\lambda t}\pi(x)^{-1/q(s)}.$$

Hence

$$|Q(x,y)| \leq 2s + \frac{1}{\lambda}\pi(x)^{-1/q(s)}.$$

For $s = \frac{1}{2\alpha}\log_+(\log(\frac{1}{\pi(x)}))$ this gives the desired result.

The uniform bounds on $|Q|$ follow from the rough estimates

$$\frac{1-2\pi_*}{\log((1-\pi_*)/\pi_*)}\lambda \leq \alpha \leq \lambda/2$$

given in Saloff-Coste (1997, Lemma 2.2.2 and Corollary 2.2.10) **QED**

4 Some applications

In sections 4.1 and 4.2, we are interested in the long term behavior of the *empirical occupation measure of the process*. We then let $\Sigma = \Delta, V_n = \delta_{X_n}$ and

$$v_n = \frac{1}{n}\sum_{i=1}^{n}\delta_{X_i}.$$

Hence, $\hat{V}_n(x) = \delta_x$ and $\theta_n = \pi_n$.

4.1 Markov chains

Let (M_n) be a deterministic (or $\mathcal{F}_0$ measurable) sequence of Markov matrices over E. A *non homogeneous* Markov chain with transition matrices (M_n) is an adapted process (X_n) on E verifying (1).

Proposition 4.1 *Let $L((\pi_n)) \subset \Delta$ denote the limit set of (π_n) and let $\mathsf{conv}[L((\pi_n))]$ denote its convex hull. Suppose that hypothesis 2.1 holds. Then $L((v_n)) \subset \mathsf{conv}[L((\pi_n))]$ with probability one.*

Proof : The set $C = \Delta \times \mathsf{conv}[L((\pi_n))]$ is adapted to (v_n, π_n). The induced differential equation $\dot{v} \in -v + \mathsf{conv}[L((\pi_n))]$ has a unique global attractor $\mathsf{conv}[L((\pi_n))]$. Hence, by Theorem 2.6 and Proposition 2.7, (ii), $L((v_n)) \subset \mathsf{conv}[L((\pi_n))]$. **QED**

Corollary 4.2 *Suppose that the sequence* (M_n) *lies in a compact subset of* $\mathsf{M}_{ind}(E)$ *and verifies* $M_{n+1} - M_n \to 0$*. Then conclusion of proposition 4.1 holds.*

Proof : Follows from proposition 4.1 and corollary 3.1. **QED**

Corollary 4.3 *Assume that* $M_n \to M \in \mathsf{M}_{ind}(E)$*. Then* $v_n \to \pi$ *the invariant probability of* M*.*

Markov chains with rare transitions

Among the well studied chains that motivate our analysis are the *chains with rare transitions.*

Let M_0 be an irreducible Markov matrix over E, reversible with respect to a reference probability π_0. That is

$$\pi_0(x)M_0(x,y) = \pi_0(y)M_0(y,x).$$

We sometimes call such an M_0, an *exploration matrix* since it provides a way to explore the state space.

Let $W : E \times E \to \mathbb{R}$, be a map and (β_n) a sequence of positive numbers. Set

$$M_n(x,y) = M(\beta_n, x, y) \tag{6}$$

where

$$M(\beta, x, y) = \begin{cases} M_0(x,y)\psi[\exp(-\beta W(x,y))] & \text{if } x \neq y, \\ 1 - \sum_{y \neq x} M(\beta, x, y) & \text{if } x = y, \end{cases}$$

and

$$\psi(u) = \min(1, u) \tag{7}$$

or

$$\psi(u) = \frac{u}{1+u}.$$

In particular, let $U : E \to \mathbb{R}$ be a map, and let

$$W(x,y) = U(y) - U(x), \tag{8}$$

then (M_n) are the transition matrices of the so-called *Metropolis-Hasting* $(\beta_n = \beta)$ or *simulated annealing* $(\beta_n \to \infty)$ algorithm (Hajek (1982), Holley and Stroock (1988), Miclo (1992)).

Consider the Markov chain with rare transitions (6) where W is given by (8). For $x, y \in E$ a path γ from x to y is a sequence of points $x_0 = x, x_1, \ldots x_n = y$ such that $M_0(x_i, x_{i+1}) > 0$. We let $\Gamma_{x,y}$ denote the set of all paths from x to y. The *elevation* from x to y is defined as

$$\mathsf{Elev}(x,y) = \min\{\max\{U(z)\ : z \in \gamma\}\ : \gamma \in \Gamma_{x,y}\}$$

and the *energy barrier* as

$$U^{\#} = \max\{\mathsf{Elev}(x,y) - U(x) - U(y) + \mathsf{min}U\ : x \in E, y \in E\} \tag{9}$$

Proposition 4.4 *Consider the Markov chain with rare transitions (6) with W given by (8). Assume that $\beta_n = \beta(n)$ where $\beta : \mathbb{R}^+ \to \mathbb{R}^+$ is differentiable and verify*

$$0 \leq \dot{\beta}(t) \leq \frac{A}{t}$$

for some $A < 1/2U^{\#}$. Then $v_n \to \pi$ where

$$\pi(x) \propto \pi_0(x)\mathbf{1}_{\mathsf{Argmin}U}(x).$$

Proof : Our first goal is to verify hypothesis 2.1. Let $\lambda(\beta)$ denote the spectral gap of $M(\beta,\cdot,\cdot)$. It follows from Theorem 2.1 in Holley and Stroock (1988) that

$$\lim_{\beta\to\infty} \frac{\log(\lambda(\beta))}{\beta} = -U^{\#}. \tag{10}$$

The invariant probability measure of $M(\beta,\cdot,\cdot)$ is the *Gibbs measure*

$$\pi_\beta(x) \propto \exp(-\beta U(x))\pi_0(x). \tag{11}$$

Since $\beta_n \leq \beta_1 + A\log(n)$, by application of the last inequality of Proposition 3.4, one gets that hypothesis 2.1 (i) holds.

For $x \neq y$

$$\frac{\partial M(\beta,x,y)}{\partial\beta} = -M_0(x,y)W(x,y)\psi'(\exp(-\beta W(x,y))\exp(-\beta W(x,y)).$$

Using the fact that $|\psi'(t)t| \leq 1$, one gets that

$$|\frac{\partial M(\beta,x,y)}{\partial\beta}| \leq c$$

for some $c > 0$. Hence by the mean value theorem

$$|M_{n+1} - M_n| \leq c|\beta_{n+1} - \beta_n| \leq (Ac)/n.$$

This proves assertion (ii) of proposition 3.3. The proof of assertion (iii) is similar since

$$|\frac{\partial \pi_\beta(x)}{\partial \beta}| = |\pi_\beta(x)(U(x) - \sum_y \pi_\beta(y)U(y))| \leq 2||U||.$$

This concludes the verification of hypothesis 2.1.

Here $\pi_n(x) \propto \exp(-\beta_n U(x))\pi_0(x)$ so that $\pi_n \to \pi$. The result follows from Proposition 4.1. **QED**

Remark 4.5 For general W, it is always possible to define a *quasipotential* U (defined in term of W and M_0) and an energy barrier $U^\#$ (in general not given by (9)) such that both equations (10) and (11) hold. We refer the reader to Miclo (1992) for more details and proofs. With this quasi-potential and barrier Proposition 4.4 holds.

4.2 Vertex reinforced random walks

Vertex-reinforced random walks (VRRW) were first introduced by Pemantle (1988, 1992).

Suppose $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$. A general VRRW on E is defined by

$$M_n(x, y) = K_n(x, y, v_n)$$

where for each integer n and $v \in \Delta$, $K_n(\cdot, \cdot, v)$ is a deterministic Markov matrix over E, which specifies the rule of the reinforcement.

The following result was proved in Benaïm (1997).

Proposition 4.6 *Assume that $K_n(x, y, v) = K(x, y, v)$ is indecomposable for each $v \in \Delta$ and that the map $v \mapsto K(x, y, v)$ is continuous on Δ. Let $\pi(v)$ denote its invariant measure. Then the limit set of (v_n) is almost surely an internally chain transitive set of the differential equation*

$$\dot{v} = -v + \pi(v). \tag{12}$$

Proof : This follows from Proposition 3.1 and Theorem 2.6. **QED**

Linear reinforcement

The original VRRW as defined by Pemantle (1998, 1992) corresponds to a *linear reinforcement:*

$$M_n(x, y) \propto U(x, y) \left[1 + \sum_{i=1}^{n} \mathbf{1}_{X_i = y} \right],$$

where U is a matrix with nonnegative entries. Equivalently, with the notation of the previous paragraph,

$$K_n(x, y, v) \propto U(x, y) \left[\frac{1}{n} + v(y) \right]. \tag{13}$$

On a finite graph, this process was first analyzed by Pemantle (1992) for symmetric positive matrices ($U(x, y) = U(y, x) > 0$) and later by Benaïm (1997) for general positive matrices using proposition 4.6. As an example of what can be proved is the following result first due to Pemantle (1992)

Proposition 4.7 *Suppose* $U(x, y) = U(y, x) > 0$. *Then the limit set of* (v_n) *is a compact connected subset of the critical set of the map*

$$v \mapsto U(v, v) = \sum_{x,y} U(x, y) v(x) v(y).$$

Proof : This follows from the fact that $v \mapsto U(v, v)$ is a strict lyapounov function of (12) whose critical points are the zeroes of (12). **QED**

When the matrix U has zero entries, $K(x, y, v)$ may no longer be indecomposable for some $v \in \partial\Delta$ and proposition 4.6 cannot be applied. This makes the analysis of VRRW with linear reinforcement much more difficult. Beautiful results on $\mathbb{Z}$ and $\mathbb{Z}^d$ have been obtained by Pemantle and Volkov (1999), Volkov (2001) and Tarres (2004). We refer the reader to Pemantle (2007) for a survey and further references.

Non homogeneous linear reinforcement

Let (a_n) be a positive sequence and denote $r_n = \sum_{i=1}^{n} a_i$. Consider the VRRW corresponding to:

$$M_n(x, y) \propto U(x, y) \left[1 + \sum_{i=1}^{n} a_i \mathbf{1}_{X_i = y} \right],$$

where U is a matrix with nonnegative entries. Equivalently, $M_n(x,y) = K(x,y,\epsilon_n,w_n)$ with

$$K(x,y,\epsilon,w) \propto U(x,y)\,[\epsilon + w(y)]\,, \tag{14}$$

$\epsilon_n = 1/r_n$ and $w_n = \frac{1}{r_n}\sum_{i=1}^n a_i\delta_{X_i}$. Using proposition 3.1, it is not hard to check that hypothesis 2.1 and hypothesis 2.2 (with $V_i = \delta_{X_i}$) are satisfied, so that theorem 2.6 applies.

Since $\delta_{X_i} = v_i + (i-1)(v_i - v_{i-1})$, using the convention $r_0 = v_0 = 0$,

$$\begin{aligned} w_n &= \frac{1}{r_n}\sum_{i=1}^{n}(r_i - r_{i-1})v_i + \frac{1}{r_n}\sum_{i=1}^{n}(i-1)(v_i - v_{i-1})a_i \\ &= v_n + \frac{1}{r_n}\sum_{i=1}^{n-1} r_i(v_i - v_{i+1}) + \frac{1}{r_n}\sum_{i=1}^{n-1} i a_{i+1}(v_{i+1} - v_i) \\ &= v_n - \frac{1}{r_n}\sum_{i=1}^{n}(r_i - i a_{i+1})(v_{i+1} - v_i). \end{aligned}$$

Since $|v_{i+1} - v_i| \le 2/i$,

$$|w_n - v_n| \le \frac{2}{r_n}\sum_{i=1}^{n}\left|\frac{r_i}{i} - a_{i+1}\right|.$$

Consider now the two following classes of sequences (a_i):

(i) $a_i = a(i)$ where a is an increasing continuous function such that for all positive $s \in]0,1]$, $\lim_{t\to\infty}\frac{a(ts)}{a(t)} = 1$.

(ii) $a_i = a(i)$ where a is a decreasing continuous function such that for all positive $s \in]0,1]$, $\lim_{t\to\infty}\frac{a(ts)}{a(t)} = 1$, and there exists $b : [0,1] \to \mathbb{R}^+$ measurable such that $\int_0^1 b(s)ds < \infty$ and for all $(s,t) \in]0,1] \times \mathbb{R}^+$,

$$0 \le \frac{a(ts)}{a(t)} - 1 \le b(s).$$

For example $a_i = (\log(i))^\alpha$ satisfies (i) for $\alpha \ge 0$ and (ii) for $\alpha < 0$.

Lemma 4.8 *Assume (i) or (ii) holds, then* $\lim_{n\to\infty}|w_n - v_n| = 0$.

Proof : Note that it suffices to prove that $\frac{r_i}{i} - a_{i+1} = o(a_i)$. Assume first(i) holds. Then

$$\begin{aligned} 0 &\leq a_{i+1} - \frac{r_i}{i} \\ &\leq a_i \left(\frac{a_{i+1}}{a_i} - 1 + \int_0^1 \left(1 - \frac{a(is)}{a(i)} \right) ds \right) \\ &= o(a_i). \end{aligned}$$

Assume now (ii) holds. Then

$$\begin{aligned} 0 &\leq \frac{r_i}{i} - a_{i+1} \\ &\leq a_i \left(1 - \frac{a_{i+1}}{a_i} + \int_0^1 \left(\frac{a(is)}{a(i)} - 1 \right) ds \right) \\ &= o(a_i) \end{aligned}$$

by dominated convergence theorem. **QED**

Let $\pi(v)$ denote the invariant probability of $K(x, y, 0, v)$. Then the previous lemma implies that when (i) or (ii) holds, $\lim_{n\to\infty} |\pi_n - \pi(v_n)| = 0$. This last property with theorem 2.6 implies the

Theorem 4.9 *Assume that (i) or (ii) holds, then the limit set of (v_n) is almost surely an internally chain transitive set of the differential equation*

$$\dot{v} = -v + \pi(v). \tag{15}$$

Note that proposition 4.7 also holds for sequences (a_i) satisfying (i) or (ii).

Exponential reinforcement

Let $U : E \times E \to \mathbb{R}$ be a map. For $x \in E$ and $v \in \Delta$, set

$$U(x, v) = \sum_{y \in E} U(x, y) v(y),$$

$$W(x, y, v) = U(y, v) - U(x, v),$$

$$K(\beta, x, y, v) = \begin{cases} M_0(x, y)\psi[\exp(-\beta W(x, y, v))] & \text{if } x \neq y, \\ 1 - \sum_{y \neq x} K(\beta, x, y, v) & \text{if } x = y, \end{cases}$$

and

$$K_n(x,y,v) = K(\beta_n, x, y, v). \tag{16}$$

Here M_0 is an exploration matrix and ψ is given by (7). When $\beta_n = \beta$, such a VRRW can be seen as a discrete time version of the self-interacting diffusions on compact manifolds that have been thoroughly analyzed by Benaïm, Ledoux and Raimond (2002), Benaïm and Raimond (2003, 2005, 2006). When $\beta_n = A\log(n)$, the VRRW can be seen as a discrete time version of the self-interacting diffusions on compact manifolds studied by Raimond (2006).

Let $U^{\#}(\cdot, y)$ be the energy barrier as defined by equation (9) of the map $x \mapsto U(x,y)$

Theorem 4.10 *Consider the VRRW with exponential reinforcement defined by (16). Assume that $\beta_n = \beta(n)$ where $\beta : \mathbb{R}^+ \to \mathbb{R}^+$ is differentiable and verify*

$$0 \leq \dot{\beta}(t) \leq \frac{A}{t}$$

for some $A1/2\max\{U^{\#}(\cdot, y)\ : y \in E\}$. Let

$$C(v) = \Delta(\mathsf{Argmin} U(\cdot, v))$$

denote the set of probabilities supported by $\mathsf{Argmin} U(\cdot, v)$. Then the limit set of (v_n) is an internally chain transitive set of

$$\dot{v} \in -v + C(v).$$

Proof : This is an application of Theorem 2.6. The verification of hypothesis 2.1 is similar to the one given in proposition 4.4. Details are left to reader.

It is easily seen that C is a closed-valued set with convex values. For $v \in \Delta$, let

$$\pi_n[v](x) \propto \pi_0(x)\exp(-\beta_n U(x,v))$$

and

$$\pi[v](x) \propto \pi_0(x)\mathbf{1}_{\mathsf{Argmin} U(\cdot,v)}(x).$$

The invariant probability of K_n is $\pi_n[v_n]$ and

$$\lim_{n\to\infty} \pi_n[v](x) = \pi[v](x).$$

This proves that C is adapted to $(v_n, \pi_n[v_n])$ and the result follows from Theorem 2.6. **QED**

Corollary 4.11 (symmetric interaction) *Assume that hypotheses of Theorem 4.10 hold and assume furthermore that U is* symmetric *(i.e $U(x,y) = U(y,x)$). Then (v_n) converges almost surely to a connected component of the set*

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

Proof : Let

$$H(v) = \frac{1}{2}U(v,v) = \frac{1}{2}\sum_{x,y} U(x,y)v(x)v(y).$$

We claim that H is a lyapouvov function of the differential inclusion (5). Let $t \mapsto v(t)$ be a solution to (5) then, for almost all $t \geq 0$

$$\begin{aligned} \frac{d}{dt}H(v(t)) &= \frac{1}{2}[U(\dot{v}(t),v(t)) + U(v(t),\dot{v}(t)] = U(\dot{v}(t),v(t)) \\ &= U(\dot{v}(t)+v(t),v(t)) - U(v(t),v(t)) \\ &= \min_x U(x,v(t)) - U(v(t),v(t)), \end{aligned}$$

where we have used the symmetry of U, the fact that $\dot{v}+v \in C(v)$ and the definition of $C(v)$. Since $t \mapsto H(v(t))$ is locally Lipchitz, it is nondecreasing. If now $t \mapsto H(v(t))$ is constant over a time interval, then $v(t) \in C(v(t))$ over this time interval. This proves that H is a Lyapounov function for $\Lambda = \{v \in \Delta \ : v \in C(v)\}$. The result now follows from Proposition 2.9 (compare to Benaïm, Hofbauer and Sorin (2005), Theorem 5.5) provided we show that $H(\Lambda)$ has empty interior.

Let $v \in \Lambda \cap \mathsf{int}(\Delta)$. Since the mapping $x \mapsto U(x,v)$ is constant, for all for all $w \in \Delta$, $U(w,v) = U(v,v)$. Therefore $H(v) = U(w,v)$ for all $w \in \Delta$. It follows that H restricted to $\Lambda \cap \mathsf{int}(\Delta)$ is a constant map. The same reasonning applies to prove that H restricted to each face of Δ is a constant map. We thus have proved that $H(\Lambda)$ takes finitely many values. **QED**

Remark 4.12 Corollary 4.11 still holds true under the weaker assumption that the map $v \mapsto U(x,v)$ is smooth and convex in v.

Corollary 4.13 *Assume that U is symmetric and nonnegative and that*

$$Ker(U) \cap T\Delta = \{0\}.$$

Then $\{v \in \Delta : v \in C(v)\}$ reduces to a singleton v^* and (v_n) converges almost surely to v^*.

Proof : Let $v \in C(v), w \in \Delta$ and $h = w - v$. Since $v \in C(v), U(v,h) \geq 0$. Thus $U(w,w) - U(v,v) = 2U(v,h) + U(h,h) \geq 0$, proving that v is a global minimum of $v \mapsto U(v,v)$. Since $U(h,h) > 0$ for $h = w - v \neq 0$, such a global minimum is unique. **QED**

4.3 Games

Consider a two-players game. We let E_1 (respectively E_2) denote the finite *set of actions* available to player 1 (respectively player 2) and

$$U = (U^1, U^2) : E_1 \times E_2 \to \mathbb{R} \times \mathbb{R}$$

denote the payoff function of the game. If player 1 and player 2 choose respectively the actions $x \in E_1$ and $y \in E_2$, then player 1 gets $U^1(x,y)$ and player 2 gets $U^2(x,y)$.

Let $((X_n, Y_n))$ denote the sequence of plays. In noncooperative game theory we assume that $((X_n, Y_n))$ is adapted to some filtration $(\mathcal{F}_n)$ and that at the beginning of round $n+1$, players have no information on the action to be played by their opponents: for all $(x,y) \in E_1 \times E_2$ and $n \in \mathbb{N}$

$$\mathsf{P}(X_{n+1} = x, Y_{n+1} = y | \mathcal{F}_n) = \mathsf{P}(X_{n+1} = x | \mathcal{F}_n) \mathsf{P}(Y_{n+1} = y | \mathcal{F}_n).$$

4.3.1 Markovian fictitious play

For $x \in E_1$ and $v^2 \in \Delta(E_2)$ set

$$U^1(x, v^2) = \sum_{z \in F} U^1(x,z) v^2(z).$$

Let

$$v_n^2 = \frac{1}{n} \sum_{i=1}^n \delta_{Y_i}.$$

A well studied strategy known as "fictitious play" consists for player 1 to play at time $n+1$ an action maximizing $U^1(\cdot, v_n^2)$, that is

$$X_{n+1} \in \mathsf{Argmax} U^1(\cdot, v_n^2). \tag{17}$$

This strategy relies on the idea that in absence of information on the next move of his opponent, player 1 assumes that he (the opponent) will play accordingly to the past empirical distribution of his moves. While fictitious play was originally proposed in 1951 by Brown as an algorithm to compute Nash equilibria it has been recently rediscover as a "learning model" (Fudenberg and Kreps (1993); Fudenberg and Levine (1998)) and has been extensively studied (Monderer and Shapley (1996); Benaïm and Hirsch (1999a, b); Hofbauer and Sandholm (2002); Benaïm, Hofbauer and Sorin (2005, 2006), see also Pemantle (2007) for an overview and further references).

Fictitious plays requires to solve the maximization problem (17) at each stage of the game. If the cardinal of E_1 is too large (or if players have computational limits) such a computation may be problematic. An alternative strategy proposed first in Benaïm, Hofbauer and Sorin (2006), based on pairwise comparison of payoffs, is as follows: The strategy of player 1 is such that $\mathsf{P}(X_{n+1} = y|\mathcal{F}_n) = M_n(X_n, y)$ with M_n the Markov matrix defined by

$$M_n(x,y) = \begin{cases} M_0(x,y)\psi[\exp(-\beta_n W_n(x,y))] & \text{if } x \neq y, \\ 1 - \sum_{y\neq x} M_n(x,y) & \text{if } x = y, \end{cases} \tag{18}$$

where

$$W_n(x,y) = U^1(x, v_n^2) - U^1(y, v_n^2),$$

M_0 is an exploration matrix, ψ is given by (7) and β_n is an increasing positive sequence. Such a strategy will be called a *Markovian fictitious play strategy*.

Adopting the view point of player 1, we choose, as an observation space,

$$\Sigma = \Delta(E_1) \times \Delta(E_2)$$

and as an observation variable

$$V_n = (\delta_{X_n}, \delta_{Y_n}).$$

Hence (v_n) is the *empirical frequency of the actions* played up to time n, and

$$\hat{V}_n(x) = (\delta_x, \nu_n),$$

where $\nu_n = \mathsf{E}(\delta_{Y_{n+1}}|\mathcal{F}_n)$.

We let $U^{1,\#}(y)$ denote the energy barrier, as defined by (9), of the map $x \mapsto U^1(x,y)$.

Theorem 4.14 *Assume that player 1 plays a Markovian fictitious play strategy as given by (18). Assume that $\beta_n = \beta(n)$ where β is differentiable and verify*

$$0 \leq \dot{\beta}(t) \leq \frac{A}{t}$$

for some $A < 1/2 \max\{U^{1,\#}(y) \, : y \in E\}$.

For $v = (v^1, v^2) \in \Delta(E) \times \Delta(F)$ let

$$C_1(v^2) = \Delta(\mathsf{Argmax} U^1(\cdot, v^2))$$

and

$$C(v) = C_1(v^2) \times \Delta(F)$$

Then the limit set of (v_n) is an internally chain transitive set of

$$\dot{v} \in -v + C(v).$$

Proof : This still an application of Theorem 2.6. The verification of hypothesis 2.1 is similar to the one given in proposition 4.4. Let

$$\pi_n[v^2](x) \propto \pi_0(x) \exp(\beta_n U^1(x, v^2))$$

and

$$\pi[v^2](x) \propto \pi_0(x) \mathbf{1}_{\mathsf{Argmax}(U^1(\cdot, v^2))}(x).$$

Then, the invariant probability of M_n is $\pi_n = \pi_n[v_n^2]$ and $\theta_n = \pi_n \hat{V}_n = (\pi_n, \nu_n)$ with $\nu_n = \mathsf{E}(\delta_{Y_{n+1}} | \mathcal{F}_n)$. Since $\pi_n[v^2] \to \pi[v^2] \in C^1(v^2)$ it follows that C is an adapted graph. **QED**

Much more can be said under the assumption that **both** players adopt a Markovian fictitious play strategy: $\mathsf{P}(X_{n+1} = y | \mathcal{F}_n) = M_n^1(X_n, y)$ and $\mathsf{P}(Y_{n+1} = y | \mathcal{F}_n) = M_n^2(Y_n, y)$, with M_n^1 and M_n^2 the Markov matrices defined by (with $i \in \{1, 2\}$)

$$M_n^i(x, y) = \begin{cases} M_0(x, y) \psi[\exp(-\beta_n^i W_n^i(x, y))] & \text{if } x \neq y, \\ 1 - \sum_{y \neq x} M_n^i(x, y) & \text{if } x = y, \end{cases} \tag{19}$$

where

$$\begin{aligned} W_n^1(x, y) &= U^1(x, v_n^2) - U^1(y, v_n^2), \\ W_n^2(x, y) &= U^2(v_n^1, x) - U^2(v_n^1, y), \end{aligned}$$

M_0^i is an exploration matrix, ψ is given by (7) and β_n^i is an increasing positive sequence.

Let $\mathsf{Conv}(U)$ denote the convex hull in $\mathbb{R}^2$ of the set $\{U(x,y) : x \in E_1, y \in E_2\}$ of all possible payoffs. We now choose

$$\Sigma = \Delta(E_1) \times \Delta(E_2) \times \mathsf{Conv}(U)$$

as an observation space, and

$$V_n = (\delta_{X_n}, \delta_{Y_n}, U(X_n, Y_n))$$

as the observation variable. Hence

$$\hat{V}_n(x,y) = (\delta_x, \delta_y, U(x,y))$$

where $\nu_n = \mathsf{E}(\delta_{Y_{n+1}}|\mathcal{F}_n)$.

Theorem 4.15 *Assume that both players adopt a Markovian fictitious play strategy. Assume that for $i \in \{1,2\}$, $\beta_n^i = \beta^i(n)$ where β^i is differentiable and verify*

$$0 \leq \dot{\beta}^i(t) \leq \frac{A^i}{t}$$

for some $A^i < 1/2 \max\{U^{i,\#}(y) \ : y \in E_i\}$.

For $v = (v^1, v^2, u) \in \Delta(E_1) \times \Delta(E_2) \times \mathsf{Conv}(U)$, let

$$C(v) = \{\{(\alpha,\beta,\gamma) \in \Sigma : \alpha \in C_1(v^2), \beta \in C_2(v^1), \gamma = U(\alpha,\beta)\}$$

where $C_1(v^2)$ is like in Theorem 4.14 and $C_2(v^1)$ is analogously defined for player 2. Then the limit set of (v_n) is an internally chain transitive set of

$$\dot{v} \in -v + C(v).$$

Proof : Let (M_n^i) denote the strategy of Player i. Let π_n^i, λ_n^i be the invariant measure and spectral gap of M_n^i. On the state space $E_1 \times E_2$ the strategy of the pair of players is $M_n = M_n^1 \otimes M_n^2$ which invariant measure is $\pi_n = \pi_n^1 \otimes \pi_n^2$ and spectral gap $\lambda_n = \min(\lambda_n^1, \lambda_n^2)$. Thus hypothesis 2.1 holds for (M_n). The rest of the proof is similar to the proof of Theorem 4.14 and is left to the reader. **QED**

Corollary 4.16 (zero sum games) *Suppose that $U^2 = -U^1$. Then under the assumption of Theorem 4.15, (v_n^1, v_n^2) converges almost surely to the set of* Nash equilibria

$$\{(v_1, v_2) \ : v_1 \in C_1(v^2), v_2 \in C_2(v^1)\},$$

and $(U^1(X_n, Y_n))$ converges almost surely to the value of the game

$$u^* = \max_{v^1 \in E} \min_{v^2 \in F} U^1(v^1, v^2) = \min_{v^1 \in E} \max_{v^2 \in F} U^1(v^1, v^2).$$

Proof : This follows from theorem 2.6, proposition 2.7 (ii) and the fact that the set $\{(v_1, v_2, u) \ : v_1 \in C_1(v^2), v_2 \in C_2(v^1), u = u^*\}$ is a global attractor of the differential inclusion, as proved in full generality by Benaïm, Hofbauer and Sorin (2005). **QED**

Corollary 4.17 (Potential games) *Suppose that $U^2 = U^1$. Then under the assumption of Theorem 4.15, (v_n^1, v_n^2) converges almost surely to a connected subset of the set of Nash equilibria*

$$\{(v_1, v_2) \ : v_1 \in C_1(v^2), v_2 \in C_2(v^1)\}$$

on which U^1 is constant, and $(U^1(X_n, Y_n))$ converges almost surely towards this constant.

Proof : Follows from theorem 2.6, proposition 2.9, and the fact that $U^1 = U^2$ is a Lyapounov function of the differential inclusion. The proof of this later point is given in (Benaïm Hofbauer and Sorin, 2005, Theorem 5.5). It is similar to the proof Corollary 4.11. **QED**

4.3.2 A remark on hypothesis 2.2

We give here a simple example showing the necessity of hypothesis 2.2.

Consider the zero sum game where $E_1 = E_2 = \{0, 1\}$, $U^1 = -U^2$ and

$$U^1 = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.$$

Let $V_n = U^1(X_n, Y_n)$ be the payoff to player 1 at time n. One has

$$\hat{V}_n(x) = U^1(x, 1)\nu_n + U^1(x, 0)(1 - \nu^n)$$

with $\nu_n = \mathsf{E}(Y_{n+1}|\mathcal{F}_n)$.

Suppose player 1 adopts the strategy given by

$$M_n = \left[\begin{array}{cc} 1/2 & 1/2 \\ 1/2 & 1/2 \end{array}\right].$$

Then $\pi_n(0) = \pi_n(1) = 1/2$ and $\theta_n = \pi_n \hat{V}_n = -1/2$ (the value of the game). Hypotheses 2.1 and 2.2 are easily seen to be satisfied (see Lemma 3.2 and remark 2.3 (*iii*)) and by theorem 2.6

$$v_n \to -\frac{1}{2}$$

regardless of the strategy played by 2.

Suppose now that player 1 adopts the strategy given as

$$M_n = \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right]$$

and that player 2 plays $Y_{n+1} = X_n$ for all $n \geq 1$.

Again $\pi_n(0) = \pi_n(1) = 1/2$ and $\theta_n = -1/2$. However, hypothesis 2.2 is not verified and the prediction given by (a wrong application of) theorem 2.6 fails since

$$v_n = \frac{1}{n}(U^1(X_1, Y_1) + (n-1)) \to 1.$$

5 Proof of Theorem 2.6

Let F denote a set–valued function mapping each point $x \in \mathbb{R}^m$ to a set $F(x) \subset \mathbb{R}^m$. We call F a *standard set valued-map* provided it verifies the three following conditions:

(i) F is a closed set-valued map. That is

$$Graph(F) = \{(x, y) \ : y \in F(x)\}$$

is a closed subset of $\mathbb{R}^m \times \mathbb{R}^m$.

(ii) F has nonempty compact convex values, meaning that $F(x)$ is a nonempty compact convex subset of $\mathbb{R}^m$ for all $x \in \mathbb{R}^m$.

(iii) There exists $c > 0$ such that for all $x \in \mathbb{R}^m$

$$\sup_{z \in F(x)} \|z\| \leq c(1 + \|x\|)$$

where $\|\cdot\|$ denotes any norm on $\mathbb{R}^m$.

Given a standard set-valued map F, set

$$F^\delta(u) = \{w \in \mathbb{R}^m \ : \exists v \in \mathbb{R}^m \ : \ d(u, v) \leq \delta, \ d(w, F(v)) \leq \delta\}.$$

The following proposition follows from the results of Benaïm, Hofbauer and Sorin (2005).

Proposition 5.1 *Let (x_n) and (U_n) be discrete time processes living in $\mathbb{R}^m$ and (γ_n) a sequence of nonnegative numbers. Let (F_n) be a sequence of set-valued maps and let F be a standard set valued-map. Assume that*

(i)

$$x_{n+1} - x_n - \gamma_{n+1} U_{n+1} \in \gamma_{n+1} F_n(x_n)$$

(ii)

$$\sum_n \gamma_n = \infty, \ \lim_{n \to \infty} \gamma_n = 0.$$

(iii) *For all $T > 0$*

$$\lim_{n \to \infty} \ \sup \left\{ \left\| \sum_{i=n}^{k-1} \gamma_{i+1} U_{i+1} \right\| \ : k = n+1, \ldots, m(\tau_n + T) \right\} = 0$$

where

$$\tau_n = \sum_{i=1}^{n} \gamma_i$$

and

$$m(t) = \sup\{k \geq 0 \ : t \geq \tau_k\} \tag{20}$$

(iv) $\sup_n \|x_n\| = M < \infty$,

(v) *For all* $\delta > 0$ *there exists* n_0 *such that*

$$F_n(x_n) \subset F^\delta(x_n)$$

for all $n \geq n_0$.

Then the limit set of (x_n) *is an attractor free set of the dynamics induced by* F.

Remark 5.2 If condition (v) is strengthen to $F_n = F$, Proposition 5.1 follows from Proposition 1.3 and Theorem 4.3 of Benaïm, Hofbauer and Sorin (2005). Under the weaker hypothesis (v), it suffices to verify that the arguments given in the proof of Proposition 1.3 adapt verbatim.

With the notation of the preceding sections, write

$$v_{n+1} - v_n = \frac{1}{n+1}[-v_n + V_{n+1}] = \frac{1}{n+1}[-v_n + \theta_n + U_{n+1}]$$

where

$$U_{n+1} = V_{n+1} - \theta_n. \tag{21}$$

Hence, conditions $(i), (ii)$ and (iv) of the previous proposition are satisfied with $F_n(u) = -u + C_n(u)$ and $\gamma_n = \frac{1}{n}$. Condition (v) follows from the next lemma.

Lemma 5.3 *Let* C *be adapted. For* $u \in \Sigma$ *and* $\delta > 0$ *set*

$$C^\delta(u) = \{w \in \Sigma \, : \exists v \in \Sigma \, : \, d(u, v) \leq \delta, d(w, C(v)) \leq \delta\}.$$

Then for all $\delta > 0$ *there exists* n_0 *such that*

$$C_n(u) \subset C^\delta(u)$$

for all $n \geq n_0$ *and* $u \in p(C_n)$.

Proof : Let $\Gamma_n = p(C_n)$. Assume to the contrary that there exist sequences $u_{n_k} \in \Gamma_{n_k}$ and $v_{n_k} \in C_{n_k}(u_{n_k})$ such that $n_k \to \infty$ and $v_{n_k} \notin C^\delta(u_{n_k})$. By compactness we may assume that $u_{n_k} \to u, v_{n_k} \to v \in C(u)$. Hence for n_k large enough $d(u_{n_k}, u) < \delta$ and $d(v_{n_k}, v) < \delta$ proving that $v_{n_k} \in C^\delta(u_{n_k})$. **QED**

To conclude the proof of theorem 2.6 it remains to verify condition (iii) of proposition 5.1.

Lemma 5.4 *Under hypothesis 2.1, the sequence (U_n) defined by (21) verifies hypothesis (iii) of proposition 5.1.*

Proof : Set $\frac{1}{n+1}U_{n+1} = \epsilon^0_{n+1} + \epsilon_{n+1}$ with

$$\epsilon^0_{n+1} = \frac{1}{n+1}(V_{n+1} - \hat{V}_n(X_{n+1})),$$

and

$$\begin{aligned} \epsilon_{n+1} &= \frac{1}{n+1}(\hat{V}_n(X_{n+1}) - \pi_n \hat{V}_n) \\ &= \frac{1}{n+1}(Q_n - M_n Q_n)\hat{V}_n(X_{n+1}) \end{aligned}$$

where the last equality follows from the definition of Q_n. Now, write $\epsilon_{n+1} = \sum_{i=1}^4 \epsilon^i_{n+1}$, where

$$\epsilon^1_{n+1} = \frac{1}{n+1}[Q_n\hat{V}_n(X_{n+1}) - M_nQ_n\hat{V}_n(X_n)],$$

$$\epsilon^2_{n+1} = \frac{1}{n+1}M_nQ_n\hat{V}_n(X_n) - \frac{1}{n}M_nQ_n\hat{V}_n(X_n),$$

$$\epsilon^3_{n+1} = \frac{1}{n}M_nQ_n\hat{V}_n(X_n) - \frac{1}{n+1}M_{n+1}Q_{n+1}\hat{V}_{n+1}(X_{n+1}),$$

$$\epsilon^4_{n+1} = \frac{1}{n+1}M_{n+1}Q_{n+1}(\hat{V}_{n+1} - \hat{V}_n)(X_{n+1}),$$

$$\epsilon^5_{n+1} = \frac{1}{n+1}[M_{n+1}Q_{n+1}\hat{V}_n(X_{n+1}) - M_nQ_n\hat{V}_n(X_{n+1})].$$

For $i = 0, \dots, 5$, let

$$\epsilon^i_n(T) = \sup\left\{\left\|\sum_{j=n}^{k-1}\epsilon^i_j\right\| : k = n+1, \dots, m(\tau_n + T)\right\}.$$

Since Σ is compact there exists a finite constant R such that $||V_n|| + \sum_x ||\hat{V}_n(x)|| \leq R$. Sequence (ϵ^0_n) is a martingale difference with $\mathsf{E}(||\epsilon^0_{n+1}||^2|\mathcal{F}_n) \leq R^2/(n+1)^2$. Therefore, by Doob's convergence theorem for L^2 martingales, $\lim_{n\to\infty}\epsilon^0_n(T) = 0$.

Sequence (ϵ_n^1) is a martingale difference with $||\epsilon_{n+1}^1|| \leq R|Q_n|/(n+1)$. Thus by a classical application of exponential martingale inequality (inequality (18) in Benaïm (1999)) we have for all positive α,

$$\mathsf{P}(\epsilon_n^1(T) \geq \alpha) \leq c \exp\left(\frac{-\alpha^2}{c\sum_{i=n}^{m(\tau_n+T)}(R^2|Q_i|^2/i^2)}\right)$$

for some positive constant c. By hypothesis 2.1, for any $\epsilon > 0$ and n large enough

$$\sum_{i=n}^{m(\tau_n+T)} (R^2|Q_i|^2/i^2) \leq \sum_{i=n}^{m(\tau_n+T)} \frac{1}{i}\frac{\epsilon}{\log(i)} \leq T\frac{\epsilon}{\log(n)}.$$

Thus

$$\sum_n \mathsf{P}(\epsilon_n^1(T) \geq \alpha)) < \infty$$

and $\lim_{n\to\infty} \epsilon_{n+1}^1 = 0$ by Borel-Cantelli Lemma.

For $n+1 \leq k \leq m(\tau_n+T)$,

$$\sum_{j=n}^{k-1} \epsilon_j^2 = \sum_{j=n}^{k-1} \frac{M_j Q_j \hat{V}_j(X_j)}{(j+1)j}.$$

Thus

$$\epsilon_n^2(T) \leq R \sum_{j=n}^{m(\tau_n+T)} \frac{|Q_j|}{j(j+1)} \leq RT \sup_{j\geq n} \frac{|Q_j|}{j}.$$

By hypothesis 2.1, this goes to zero when $n \to \infty$.

For $n+1 \leq k \leq m(\tau_n+T)$,

$$\sum_{j=n}^{k-1} \epsilon_j^3 = \frac{1}{n} M_n Q_n \hat{V}_n(X_n) - \frac{1}{k} M_k Q_k \hat{V}_k(X_k),$$

so that

$$\epsilon_n^3(T) \leq 2 \sup_{i\geq n} R\frac{1}{i}|Q_i|$$

and $\epsilon_n^3(T) \to 0$ as $n \to \infty$ by hypothesis 2.1.

The term $\epsilon_n^4(T)$ is dominated by

$$T \sup_{i\geq n} \sup_x |M_{i+1} Q_i (\hat{V}_{i+1} - \hat{V}_i)(x)|$$

which converges towards 0 as $n \to \infty$ by hypothesis 2.2.

Finally, since $M_n Q_n = Q_n + \Pi_n + I$

$$\epsilon_n^5 = \frac{1}{n+1}\left[(Q_{n+1} - Q_n)\hat{V}_n(X_{n+1}) - (\Pi_{n+1} - \Pi_n)\hat{V}_n\right].$$

Hence

$$\epsilon_n^5(T) \leq RT \sup_{i \geq n}(|Q_{i+1} - Q_i| + |\pi_{i+1} - \pi_i|) \to 0$$

by hypothesis 2.1. This completes the proof of (iii). **QED**

References

[1] M. Benaïm, (1999), *Dynamics of stochastic approximation algorithms*, Séminaire de Probabilités XXXIII, Lecture Notes in Math. **1709**, 1–68, Springer.

[2] M. Benaïm and M.W. Hirsch, (1996), *Asymptotic pseudotrajectories and chain recurrent flows, with applications*, J. Dynam. Differential Equations, 8,141–176.

[3] M. Benaïm and M.W. Hirsch, (1999), *Mixed equilibria and dynamical systems arising from fictitious play in perturbed games,* Games and Economic Behavior, **29**, 36-72.

[4] M. Benaïm, J. Hofbauer and S. Sorin, (2005), *Stochastic approximations and differential inclusions*, SIAM Journal on Control and Optimization, **44**, 328-348.

[5] M. Benaïm, J. Hofbauer and S. Sorin, (2006), *Stochastic approximations and differential inclusions. Part II: Applications*, Mathematics of Operations Research, 31, 673-695.

[6] M. Benaïm, M. Ledoux and O. Raimond, (2002), *Self-interacting diffusions*, Probab. Theor. Relat. Fields **122**, 1-41.

[7] M. Benaïm and O. Raimond, (2003), *Self-interacting diffusions II: Convergence in Law.*, Annales de l'institut Henri-Poincaré **6**, 1043-1055.

[8] M. Benaïm and O. Raimond, (2005), *Self-interacting diffusions III: Symmetric interactions.*, Annals of Probability **33**, no. 5, 1717–1759.

[9] A. Benveniste and M. Métivier and P. Priouret, (1990), "Stochastic Approximation and Adaptive Algorithms", Springer-Verlag, Berlin and New York.

[10] D. Blackwell, (1956), *An analog of the minmax theorem for vector payoffs*, Pacific Journal of Mathematics, **6**, 1-8.

[11] G. Brown, (1951), *Iterative solution of games by fictitious play*, in Koopmans T.C. (ed.) Activity Analysis of Production and Allocation, Wiley, 374-376.

[12] M. Duflo, (1996), "Algorithmes Stochastiques", Mathématiques et Applications, Springer-Verlag, vol 23.

[13] D. Fudenberg and D. K. Levine, (1998) "The Theory of Learning in Games", MIT Press.

[14] B. Hajek, (1982), *Cooling schedules for optimal annealing*, Math. Oper. Res., 13, 2, pp 311-329.

[15] J. Hofbauer and W. Sandholm., (2002) *On the global convergence of stochastic fictious play*, Econometrica, 70:2265-2294.

[16] R. Holley and D. Stroock, (1988), *Simulated Annealing via Sobolev Inequalities*, Commun. Math. Phys. 115, 553-568.

[17] H.J. Kushner and C.C. Clarck, (1978), "Stochastic Approximation for Constrained and Unconstrained Systems", Springer-Verlag , Berlin and New York.

[18] L. Ljung, (1977), *Analysis of recursive stochastic algorithms*, IEEE Trans. Automat. Control, AC-22", 551-575.

[19] M. Métivier and P. Priouret, (1987), *Théorèmes de convergence presque sure pour une classe d'algorithmes stochastiques à pas décroissant*, Probability Theory and Related Fields, 74 ,403-428.

[20] L. Miclo, (1992) *Recuit simul sans potentiel sur un ensemble fini,* Sminaire de probabilits (Strasbourg), tome 26, pp. 47-60.

[21] D. Monderer and L.S. Shapley, (1996) *Fictitious play property for games with identical interests*, J. Economic Theory, **68**, 258–265.

[22] L. Saloff-Coste,(1997), "Lectures on finite Markov Chains", Lectures on Probability Theory and Statistics 1996, Lecture Notes in Mathematics, Vol 1665.

[23] R. Pemantle, (1988), "Random processes with reinforcement" Doctoral Dissertation, M.I.T.

[24] R. Pemantle, (1992), *Vertex Reinforced Random Walk*, Probab. Theor. Relat. Fields **92**, 117-136.

[25] R. Pemantle, (2007), *A survey of random processes with reinforcement,* Probability survey, Vol 4, 1-79.

[26] R. Pemantle and S. Volkov, (1999), *Vertex-reinforced random walk on* $\mathbb{Z}$ *has finite range*, Annals of Probability, **27**, 1368-1388.

[27] O. Raimond, (2006), *Self-interacting diffusions: A simulated annealing version*, preprint.

[28] P. Tarrès, (2004), *VRRW on* $\mathbb{Z}$ *enventually get stuck at a set of five points*, Annals of probability 32, **3**, 1455-1478.

[29] S. Volkov, (2001), *Vertex-reinforced random walks on arbitrary graphs.* Annals of probability 29, 66-91.