DOI: 10.1214/EJP.v13-489

Electronic Journal of Probability

Vol. 13 (2008), Paper no. 12, pages 341–373.

Journal URL
http://www.math.washington.edu/~ejpecp/

A penalized bandit algorithm *

Damien Lamberton †   Gilles Pagès ‡

Abstract

We study a two armed-bandit recursive algorithm with penalty. We show that the algorithm converges towards its "target" although it always has a noiseless "trap". Then, we elucidate the rate of convergence. For some choices of the parameters, we obtain a central limit theorem in which the limit distribution is characterized as the unique stationary distribution of a Markov process with jumps.

Key words: Two-armed bandit algorithm, penalization, stochastic approximation, convergence rate, learning automata.

AMS 2000 Subject Classification: Primary 62L20; Secondary: 93C40,68T05,91B32.

Submitted to EJP on June 20, 2006, final version accepted February 26, 2008.

*This work has benefitted from the stay of both authors at the Isaac Newton Institute (Cambridge University, UK) on the program *Developments in Quantitative Finance*.

†Université Paris-Est, Laboratoire d'analyse et de mathématiques appliquées, UMR CNRS 8050, 5 Bld Descartes, F-77454 Marne-la-Vallée Cedex 2, France. E-mail: damien.lamberton@univ-mlv.fr

‡Université Paris 6, Laboratoire de probabilités et modèles aléatoires, UMR CNRS 7599, case 188, 4, pl. Jussieu, F-75252 Paris Cedex 5. E-mail: gpa@ccr.jussieu.fr

Introduction

In a recent joint work with P. Tarrès ([15], see also [22]), we studied the convergence of the so-called two armed bandit algorithm. The purpose of the present paper is to investigate a modified version of this recursive algorithm, in which a penalization is introduced. In the terminology of learning theory (see [17; 18]), the algorithm studied in [15] appeared as a Linear Reward-Inaction (LRI) scheme (with possibly decreasing step), whereas the one we will introduce is related to the Linear Reward-Penalty (LRP) or ($LR\varepsilon P$) schemes.

In our previous paper, we introduced the algorithm in a financial context. However, historically, this recursive procedure was designed independently in the fields of mathematical psychology (see [19]) and of engineering (see [21]). Its name goes back to another interpretation as a model of slot machine with two "arms" providing two different rates of gain. It can also be interpreted as an adaptive procedure for clinical trials, based on its connections with generalized urn models (see [15]) that are often proposed in the literature for that purpose (see *e.g.* [1] and the references therein).

Another important motivation for investigating is that the two armed bandit algorithm is known in the field of stochastic approximation as the simplest example of a recursive stochastic algorithm having a noiseless trap in the following sense: one zero of its mean function is repulsive for the related ODE but the algorithm has no stochastic noise at this equilibrium. Therefore, the standard "first order" ODE method as well as the "second order" approach based on the repeling effect induced by the presence of noise (see *e.g.* the seminal paper [20] by Pemantle) do not seem to apply for proving the non-convergence of the algorithm toward this "noiseless trap".

Let us first present the (LRI) procedure (with possibly decreasing step) in a gambling framework: one considers a slot machine in a casino (a "bandit") with two arms, say A and B (by contrast with the famous "one-armed-bandit" machines). When playing arm A (resp. B), the average yield (for 1 Euro) is $p_A \in (0,1)$ (resp. p_B). These parameters are unknown to the gambler. For the sake of simplicity one may assume that the slot machine has a 0-1 gross profit: one wins 0 or 1. Then p_A and p_B are the respective theoretical frequencies of winning with the arms. More precisely, the events A_n (resp. B_n) "winning at time n using A (resp. B)" are iid with probability $\mathbb{P}(A_n) = p_A$ (resp. $\mathbb{P}(B_n) = p_B$) with $p_A, p_B \in (0,1)$. The (LRI) procedure is an adaptive natural method to detect the most performing arm: at every time n, the player selects an arm at random, namely A with probability X_n and B with probability $1 - X_n$. Once the selected arm has delivered its verdict, the probability X_n is updated as follows (in view of the arm selection at time $n+1$):

$$X_{n+1} = X_n + \gamma_{n+1}\left(\mathbf{1}_{\{U_{n+1}\leq X_n\}\cap A_{n+1}}(1-X_n) - \mathbf{1}_{\{U_{n+1}>X_n\}\cap B_{n+1}}X_n\right), \quad X_0 = x \in [0,1],$$

where $(U_n)_{n\geq 1}$ is an iid sequence of uniform random variables on the interval $[0,1]$, independent of $(A_n, B_n)_{n\geq 1}$. In words, if the player plays arm A (as a result of the biased tossing) and wins 1 Euro at time $n+1$, the probability to choose A (at time $n+2$) will be increased by $\gamma_{n+1}(1 - X_n)$ (*i.e.* proportionally to the probability of selecting B). If the gain is 0, the probability is left unchanged. One proceeds symmetrically when B is selected. The $(0,1)$-valued parameter sequence $(\gamma_n)_{n\geq 1}$ rules the intensity of the updating. When $\gamma_n = \gamma \in (0,1)$, the above alorithm reduces to the original (LRI) procedure (see [18]) and γ is known as the *reward rate parameter*. This sequence (γ_n) specifies how the recursive learning procedure keeps the memory of the past and how fast it forgets the starting value $X_0 = x \in (0,1)$ of the procedure. Also note that such a procedure is only based on rewarding: no arm is ever "penalized" when it provides no gain ("Reward" or "Inaction").

In a financial framework, A and B can be two traders who manage at time n, X_n % and $1-X_n$ % of a fund respectively. In the framework of clinical tests, A and B model two possible clinical protocols to be tested on patients. In the framework of engineering, one may think of two subcontractors which provide a car manufacturer with specific mechanical devices (windscreen-wipers, tires, gearbox, etc) with respective reliability p_A and p_B.

This procedure has been designed in order to be "infallible" (or "optimal" in the learning automata terminology) *i.e.* to always select asymptotically the most profitable arm. The underlying feature of the above (LRI) procedure that supports such an intuition is that it is the only recursive procedure of that type which is always a sub-(resp. super)-martingale as soon as $p_A > p_B$ (resp. $p_A < p_B$) as emphasized in [14]. To be more specific, by "infallibility" we mean that if $p_A > p_B$, then X_n converges to 1 with probability 1 provided $X_0 \in (0,1)$ (and if $p_A < p_B$, the limit is 0 with symmetric results).

Unfortunately it turns out that this intuition is misleading: the algorithm is often "fallible", depending on the choice of the step sequence γ. In fact "infallibility" needs some further stringent assumptions on this step sequence (see [15], and also [23] for an ergodic version of the algorithm). Furthermore, the rate of convergence of the procedure either to its "target" 1 or to its "trap" 0 is never ruled by a CLT with rate $\sqrt{\gamma_n}$ like standard stochastic approximation algorithms are (see [11]). It is shown in [14] that its rate structure is complex, highly non-standard and strongly depends on the (unknown) values p_A and p_B. As a result, this rate becomes quite poor as these probabilities get close to each other. This illustrates in a rather striking way the effects induced by a "noiselss trap" on the dynamics of a stochastic approximation procedure.

In order to improve the efficiency of the algorithm, *i.e.* to make it "unconditionally" infallible, a natural idea is to introduce a penalty when the selected arm delivers no gain. More precisely, if the selected arm at time n performs badly, its probability to be selected is decreased by a penalty factor $\rho_n\gamma_n$. This leads to introduce a (variant of the) *Linear Reward Penalty* – or "penalized two-armed bandit" – procedure:

$$\begin{aligned} X_{n+1} &= X_n + \gamma_{n+1}\left(\mathbf{1}_{\{U_{n+1}\le X_n\}\cap A_{n+1}}(1-X_n) - \mathbf{1}_{\{U_{n+1}>X_n\}\cap B_{n+1}}X_n\right) \\ &\quad -\gamma_{n+1}\rho_{n+1}\left(X_n\mathbf{1}_{\{U_{n+1}\le X_n\}\cap A^c_{n+1}} - (1-X_n)\mathbf{1}_{\{U_{n+1}>X_n\}\cap B^c_{n+1}}\right), \quad n\in\mathbb{N}, \end{aligned}$$

where the notation A^c is used for the complement of an event A. The precise assumptions on the *reward rate* γ_n and the *penalty rate* $\gamma_n\rho_n$ will be given in the following sections. When $\gamma_n = \gamma$ and $\rho_n = \rho$ the procedure reduces to the original (LRP) procedure.

From a stochastic approximation viewpoint as well as for practical applications, our main results are on the one hand that infallibility always holds and on the other hand that it is possible to specify the sequences (γ_n) and (ρ_n) *regardless of the values of* p_A *and* p_B so that the convergence rate satisfies a CLT theorem like standard stochastic approximation procedures. However with a quite important difference: the limiting distribution is never Gaussian: it can be characterized as the (absolutely continuous) invariant distribution of a homogeneous Markov process with jumps.

The paper is organized as follows. In Section 1, we discuss the convergence of the sequence $(X_n)_{n\ge0}$. First we show that, if ρ_n is a positive constant ρ, the sequence converges with probability one to a limit $x^*_\rho \in (0,1)$ satisfying $x^*_\rho > \frac{1}{2}$ if and only if $p_A > p_B$, so that, although the algorithm manages to distinguish which arm is the most performing, it does not prescribe to play exclusively with that arm. However, when ρ is small, one observes a kind of asymptotic infallibility, namely that $x^*_\rho \to 1$ as $\rho \to 0$ (see section 1.2 below). Note that a somewhat similar setting (but with $\gamma_n = \gamma$) has been investigated in learning automata theory as the "$LR\varepsilon P$" procedure (see [12] or also [13]). To obtain true infallibility, we consider a sequence $(\rho_n)_{n\geq 1}$ which goes to zero so that the penalty rate becomes negligible with respect to the reward rate ($\gamma_n\rho_n = o(\gamma_n)$). This framework ($\rho_n \to 0$) seems new in the learning theory literature. Then, we are able to establish the infallibility of the algorithm under very light conditions on the reward rate γ_n (and ρ_n) in which $p_{_A}$ and $p_{_B}$ are not involved. From a purely stochastic approximation viewpoint, this modification of the original procedure has the same mean function and time scale (hence the same target and trap, see (5)) as the LRI procedure with decreasing step but it always keeps the algorithm away from the trap, without adding noise at any equilibrium point. (In fact, this last condition was necessary in order to keep the algorithm inside its domain $[0, 1]$ since the equilibrium points are endpoints 0 and 1.)

The other two sections are devoted to the rate of convergence. In Section 2, we show that under some conditions (including $\lim\limits_{n\to\infty} \gamma_n/\rho_n = 0$) the sequence $Y_n = (1 - X_n)/\rho_n$ converges in probability to $(1 - p_{_A})/\pi$, where $\pi = p_{_A} - p_{_B} > 0$. With additional assumptions, we prove that this convergence occurs with probability 1. In Section 3, we show that if the ratio γ_n/ρ_n goes to a positive limit as n goes to infinity, then $(Y_n)_{n\geq 1}$ converges in a weak sense to a probability distribution ν. This distribution is identified as the unique stationary distribution of a discontinuous Markov process. This result is obtained by using weak functional methods applied to a re-scaling of the algorithm. This approach can be seen as an extension of the SDE method used to prove the CLT in a more standard framework of stochastic approximation (see [11]). Furthermore, we show that ν is absolutely continuous with continuous, possibly non-smooth, piecewise $\mathcal{C}^\infty$ density. An interesting consequence of these results for practical applications is that, by choosing ρ_n and γ_n proportional to $n^{-1/2}$, one can achieve *convergence at the rate* $1/\sqrt{n}$*, without any a priori knowledge about the values of* $p_{_A}$ *and* $p_{_B}$. This is in contrast with the case of the LRI procedure, where the rate of convergence depends heavily on these parameters (see [14]) and becomes quite poor when they get close to each other.

NOTATION. Let $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$ be two sequences of positive real numbers. The symbol $a_n \sim b_n$ means $a_n = b_n + o(b_n)$.

1 Convergence of the LRP algorithm with decreasing step

1.1 Some classical background on stochastic approximation

We will rely on the ODE lemma recalled below for a stochastic procedure (Z_n) taking its values in a given compact interval I.

Theorem 1. (a) KUSHNER & CLARK'S ODE LEMMA (SEE [10]): *Consider a function* $g : I \to \mathbb{R}$*, such that* $Id + g$ *leaves* I *stable*[1]*, and the stochastic approximation procedure defined on* I *by*

$$Z_{n+1} = Z_n + \gamma_{n+1}(g(Z_n) + \Delta R_{n+1}),\ n \geq 0, \quad Z_0 \in I,$$

[1] then, for every $\gamma \in [0, 1]$, $Id + \gamma g = \gamma(Id + g) + (1 - \gamma)Id$ still takes values in the convex set I.

where $(\gamma_n)_{n\geq 1}$ *is a sequence of* $[0,1]$*-valued real numbers satisfying* $\gamma_n \to 0$ *and* $\sum_{n\geq 1}\gamma_n = +\infty$. *Set* $N(t) := \min\{n : \gamma_1 + \cdots + \gamma_{n+1} > t\}$. *Let* z^* *be an attracting zero of* g *in* I *and* $G(z^*)$ *its attracting interval. If, for every* $T > 0$,

$$\max_{N(t)\leq n\leq N(t+T)} \left| \sum_{k=N(t)+1}^{n} \gamma_k \, \Delta R_k \right| \longrightarrow 0 \qquad \mathbb{P}\text{-}a.s. \;\; as \quad t \to +\infty, \tag{1}$$

then, $Z_n \xrightarrow{a.s.} z^*$ *on the event*

$$\{Z_n \textit{ visits infinitely often a compact subset of } G(z^*)\}.$$

(*b*) The Hoeffding condition (see [2]): *If* $(\Delta R_n)_{n\geq 0}$ *is a sequence of* L^∞*-bounded martingale increments, if* (γ_n) *is non-increasing and* $\sum_{n\geq 1} e^{-\frac{\vartheta}{\gamma_n}} < +\infty$ *for every* $\vartheta > 0$, *then Assumption (1) is satisfied.*

1.2 Basic properties of the algorithm

We first recall the definition of the algorithm. We are interested in the asymptotic behavior of the sequence $(X_n)_{n\in\mathbb{N}}$, where $X_0 = x$, with $x \in (0,1)$, and

$$\begin{aligned} X_{n+1} &= X_n + \gamma_{n+1}\left(\mathbf{1}_{\{U_{n+1}\leq X_n\}\cap A_{n+1}}(1-X_n) - \mathbf{1}_{\{U_{n+1}>X_n\}\cap B_{n+1}} X_n\right) \\ &\quad -\gamma_{n+1}\rho_{n+1}\left(X_n \mathbf{1}_{\{U_{n+1}\leq X_n\}\cap A^c_{n+1}} - (1-X_n)\mathbf{1}_{\{U_{n+1}>X_n\}\cap B^c_{n+1}}\right), \quad n \in \mathbb{N}. \end{aligned}$$

Throughout the paper, we assume that $(\gamma_n)_{n\geq 1}$ is a non-increasing sequence of positive numbers satisfying $\gamma_n < 1$, $\sum_{n=1}^{\infty}\gamma_n = +\infty$ and

$$\forall \vartheta > 0, \quad \sum_n e^{-\frac{\vartheta}{\gamma_n}} < \infty,$$

and that $(\rho_n)_{n\geq 1}$ is a sequence of positive numbers satisfying $\gamma_n\rho_n < 1$; $(U_n)_{n\geq 1}$ is a sequence of independent random variables which are uniformly distributed on the interval $[0,1]$, the events A_n, B_n satisfy

$$\mathbb{P}(A_n) = p_{_A}, \quad \mathbb{P}(B_n) = p_{_B}, \quad n \in \mathbb{N},$$

where $0 < p_{_B} \leq p_{_A} < 1$, and the sequences $(U_n)_{n\geq 1}$ and $(\mathbf{1}_{A_n}, \mathbf{1}_{B_n})_{n\geq 1}$ are independent. The natural filtration of the sequence $(U_n, \mathbf{1}_{A_n}, \mathbf{1}_{B_n})_{n\geq 1}$ is denoted by $(\mathcal{F}_n)_{n\geq 0}$ and we set

$$\pi = p_{_A} - p_{_B}.$$

With this notation, we have, for $n \geq 0$,

$$X_{n+1} = X_n + \gamma_{n+1}\left(\pi h(X_n) + \rho_{n+1}\kappa(X_n)\right) + \gamma_{n+1}\Delta M_{n+1}, \tag{2}$$

where the functions h and κ are defined by

$$h(x) = x(1-x), \quad \kappa(x) = -(1-p_{_A})x^2 + (1-p_{_B})(1-x)^2, \quad 0 \leq x \leq 1,$$

$\Delta M_{n+1} = M_{n+1} - M_n$, and the sequence $(M_n)_{n\geq 0}$ is the martingale defined by $M_0 = 0$ and

$$\begin{aligned}\Delta M_{n+1} &= \mathbf{1}_{\{U_{n+1}\leq X_n\}\cap A_{n+1}}(1-X_n) - \mathbf{1}_{\{U_{n+1}>X_n\}\cap B_{n+1}}X_n - \pi h(X_n) \\ &\quad -\rho_{n+1}\left(X_n\mathbf{1}_{\{U_{n+1}\leq X_n\}\cap A^c_{n+1}} - (1-X_n)\mathbf{1}_{\{U_{n+1}>X_n\}\cap B^c_{n+1}} + \kappa(X_n)\right). \qquad (3)\end{aligned}$$

Observe that the increments ΔM_{n+1} are bounded.

1.3 The case of a constant penalty rate (an $(LR\varepsilon P)$ setting)

In this subsection, we assume

$$\forall n \geq 1, \quad \rho_n = \rho,$$

with $0 < \rho \leq 1$. We then have

$$X_{n+1} = X_n + \gamma_{n+1}\left(h_\rho(X_n) + \Delta M_{n+1}\right),$$

where

$$h_\rho(x) = \pi h(x) + \rho\kappa(x), \quad 0 \leq x \leq 1.$$

Note that $h_\rho(0) = \rho(1-p_B) > 0$ and $h_\rho(1) = -\rho(1-p_A) < 0$, and that there exists a unique $x^*_\rho \in (0,1)$ such that $h_\rho(x^*_\rho) = 0$. By a straightforward computation, we have

$$\begin{aligned}x^*_\rho &= \frac{\pi - 2\rho(1-p_B) + \sqrt{\pi^2 + 4\rho^2(1-p_B)(1-p_A)}}{2\pi(1-\rho)} \quad \text{if } \pi \neq 0 \text{ and } \rho \neq 1 \\ *[.5em] &= \frac{(1-p_A)}{(1-p_A)+(1-p_B)} \quad \text{if } \pi = 0 \text{ or } \rho = 1.\end{aligned}$$

In particular, $x^*_\rho = 1/2$ if $\pi = 0$ regardless of the value of ρ. We also have $h_\rho(1/2) = \pi(1+\rho)/4 \geq 0$, so that

$$x^*_\rho > 1/2 \quad \text{if} \quad \pi > 0. \qquad (4)$$

Now, let x be a solution of the ODE $dx/dt = h_\rho(x)$. If $x(0) \in [0, x^*_\rho]$, x is non-decreasing and $\lim_{t\to\infty} x(t) = x^*_\rho$. If $x(0) \in [x^*_\rho, 1]$, x is non-increasing and $\lim_{t\to\infty} x(t) = x^*_\rho$. It follows that the interval $[0,1]$ is a domain of attraction for x^*_ρ. Consequently, using Kushner and Clark's *ODE* Lemma (see Theorem 1), one reaches the following conclusion.

Proposition 1. *Assume that $\rho_n = \rho \in (0,1]$, then*

$$X_n \xrightarrow{a.s.} x^*_\rho \qquad \textit{as } n \to \infty.$$

The natural interpretation, given the above inequalities on x^*_ρ, is that this algorithm never fails in pointing the best arm thanks to Inequality (4), but it will never select the best arm asymptotically as the original *LRI* procedure did. However, note that

$$x^*_\rho \to 1 \qquad \text{as} \qquad \rho \to 0$$

which makes the family of algorithms (indexed by ρ) "asymptotically" infallible as $\rho \to 0$. These results are in some way similar to those obtained in [12] for the so-called $(LR\varepsilon P)$ scheme (with constant reward and penalty rates γ and ρ). By considering a penalty rate ρ_n going to zero we will show that the resulting algorithm becomes "unconditionaly" infallible as n goes to infinity.

1.4 Convergence when the penalty rate goes to zero

Proposition 2. *Assume $\lim_{n\to\infty} \rho_n = 0$. The sequence $(X_n)_{n\in\mathbb{N}}$ is almost surely convergent and its limit X_∞ satisfies $X_\infty \in \{0,1\}$ with probability 1.*

PROOF: We first write the algorithm in its canonical form

$$X_{n+1} = X_n + \gamma_{n+1}(\pi\, h(X_n) + \Delta R_{n+1}) \qquad \text{with} \qquad \Delta R_n = \Delta M_n + \rho_n \kappa(X_{n-1}). \tag{5}$$

It is straightforward to check that the *ODE* $\dot{x} = h(x)$ has two equilibrium points, 0 and 1, 1 being attractive with $(0,1]$ as an attracting interval and 0 is unstable.

Since the martingale increments ΔM_n are bounded, it follows from the assumptions on the sequence $(\gamma_n)_{n\geq 1}$ and the Hoeffding condition (see Theorem 1(b)) that

$$\max_{N(t)\leq n\leq N(t+T)} \left| \sum_{k=N(t)+1}^{n} \gamma_k \Delta M_k \right| \stackrel{\mathbb{P}\text{-}a.s.}{\longrightarrow} 0 \quad \text{as } t\to+\infty,$$

for every $T>0$. On the other hand, the function κ being bounded on $[0,1]$ and ρ_n converging to 0, we have, for every $T>0$, with the notation $\|\kappa\|_\infty = \sup_{x\in[0,1]} |\kappa(x)|$,

$$\max_{N(t)\leq n\leq N(t+T)} \left| \sum_{k=N(t)+1}^{n} \gamma_k \rho_k \kappa(X_{k-1}) \right| \leq \|\kappa\|_\infty (T+\gamma_{N(t+T)}) \max_{k\geq N(t)+1} \rho_k \longrightarrow 0 \quad \text{as } t\to+\infty.$$

Finally, the sequence $(\Delta R_n)_{n\geq 1}$ satisfies Assumption (1). Consequently, either X_n visits infinitely often an interval $[\varepsilon,1]$ for some $\varepsilon>0$ and X_n converges toward 1, or X_n converges toward 0. $\diamond$

Remark 1. If $\pi=0$, *i.e.* $p_A = p_B$, the algorithm reduces to

$$X_{n+1} = X_n + \gamma_{n+1}\rho_{n+1}(1-p_A)(1-2X_n) + \gamma_{n+1}\Delta M_{n+1}.$$

The number 1/2 is the unique equilibrium of the ODE $\dot{x} = (1-p_A)(1-2x)$, and the interval $[0,1]$ is a domain of attraction. Assuming $\sum_{n=1}^{\infty} \rho_n\gamma_n = +\infty$, and that the sequence $(\gamma_n/\rho_n)_{n\geq 1}$ is non-increasing and satisfies

$$\forall \vartheta>0, \quad \sum_{n=1}^{\infty} \exp\left(-\vartheta \frac{\rho_n}{\gamma_n}\right) < +\infty,$$

it can be proved, using the Kushner-Clark *ODE* Lemma (Theorem 1), that $\lim_{n\to\infty} X_n = 1/2$ almost surely. As concerns the asymptotics of the algorithm when $\pi=0$ and $\gamma_n = g\,\rho_n$ (for which the above condition is not satisfied), we refer to the final remark of the paper.

From now on, we will assume that $p_A > p_B$. The next proposition shows that the penalized algorithm is infallible under very light assumptions on γ_n and ρ_n.

Proposition 3. *(Infallibility) Assume $\lim_{n\to\infty} \rho_n = 0$. If the sequence $(\gamma_n/\rho_n)_{n\geq 1}$ is bounded and $\sum_n \gamma_n\rho_n = \infty$, and if $\pi>0$, we have $\lim_{n\to\infty} X_n = 1$ almost surely.*

PROOF: We have from (2), since $h \geq 0$ on the interval $[0,1]$,

$$X_n \geq X_0 + \sum_{j=1}^{n} \gamma_j \rho_j \kappa(X_{j-1}) + \sum_{j=1}^{n} \gamma_j \Delta M_j, \quad n \geq 1.$$

Since the jumps ΔM_j are bounded, we have

$$\left\| \sum_{j=1}^{n} \gamma_j \Delta M_j \right\|_{L^2}^2 \leq C \sum_{j=1}^{n} \gamma_j^2 \leq C \sup_{j\in\mathbb{N}} (\gamma_j/\rho_j) \sum_{j=1}^{n} \gamma_j \rho_j,$$

for some positive constant C. Therefore, since $\sum_n \gamma_n \rho_n = \infty$,

$$L^2\text{-}\lim_{n\to\infty} \frac{\sum_{j=1}^{n} \gamma_j \Delta M_j}{\sum_{j=1}^{n} \gamma_j \rho_j} = 0 \quad \text{so that} \quad \limsup_n \frac{\sum_{j=1}^{n} \gamma_j \Delta M_j}{\sum_{j=1}^{n} \gamma_j \rho_j} \geq 0 \qquad a.s..$$

Here, we use the fact that a sequence which converges in L^2 has a subsequence which converges almost surely. Now, on the set $\{X_\infty = 0\}$, we have

$$\lim_{n\to\infty} \frac{\sum_{j=1}^{n} \gamma_j \rho_j \kappa(X_{j-1})}{\sum_{j=1}^{n} \gamma_j \rho_j} = \kappa(0) > 0.$$

Hence, it follows that, still on the set $\{X_\infty = 0\}$,

$$\limsup_{n\to\infty} \frac{X_n}{\sum_{j=1}^{n} \gamma_j \rho_j} > 0.$$

Therefore, we must have $\mathbb{P}(X_\infty = 0) = 0$. $\diamondsuit$

The following Proposition will give a control on the conditional variance process of the martingale $(M_n)_{n\in\mathbb{N}}$ which will be crucial to elucidate the rate of convergence of the algorithm.

Proposition 4. *We have, for $n \geq 0$,*

$$\mathbb{E}\left(\Delta M_{n+1}^2 \mid \mathcal{F}_n\right) \leq p_A(1 - X_n) + \rho_{n+1}^2(1 - p_B).$$

PROOF: We have

$$\Delta M_{n+1} = V_{n+1} - \mathbb{E}(V_{n+1} \mid \mathcal{F}_n) + W_{n+1} - \mathbb{E}(W_{n+1} \mid \mathcal{F}_n),$$

with

$$V_{n+1} = \mathbf{1}_{\{U_{n+1} \leq X_n\} \cap A_{n+1}}(1 - X_n) - \mathbf{1}_{\{U_{n+1} > X_n\} \cap B_{n+1}} X_n$$

and

$$W_{n+1} = -\rho_{n+1}\left(X_n \mathbf{1}_{\{U_{n+1} \leq X_n\} \cap A_{n+1}^c} - (1 - X_n)\mathbf{1}_{\{U_{n+1} > X_n\} \cap B_{n+1}^c}\right).$$

Note that $V_{n+1}W_{n+1} = 0$, so that

$$\begin{aligned}\mathbb{E}\left(\Delta M_{n+1}^2 \mid \mathcal{F}_n\right) &= \mathbb{E}(V_{n+1}^2 \mid \mathcal{F}_n) + \mathbb{E}(W_{n+1}^2 \mid \mathcal{F}_n) - (\mathbb{E}(V_{n+1} + W_{n+1} \mid \mathcal{F}_n))^2 \\ &\leq \mathbb{E}(V_{n+1}^2 \mid \mathcal{F}_n) + \mathbb{E}(W_{n+1}^2 \mid \mathcal{F}_n).\end{aligned}$$

Now, using $p_B \leq p_A$ and $X_n \leq 1$,

$$\begin{aligned}\mathbb{E}\left(V_{n+1}^2 \mid \mathcal{F}_n\right) &= p_A X_n(1-X_n)^2 + p_B(1-X_n)X_n^2 \\ &\leq p_A(1-X_n)\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}(W_{n+1}^2 \mid \mathcal{F}_n) &= \rho_{n+1}^2\left(X_n^3(1-p_A) + (1-X_n)^3(1-p_B)\right) \\ &\leq \rho_{n+1}^2(1-p_B).\end{aligned}$$

This proves the Proposition. ◇

2 The rate of convergence: pointwise convergence

2.1 Convergence in probability

Theorem 2. *Assume*

$$\lim_{n\to\infty} \rho_n = 0, \quad \lim_{n\to\infty} \frac{\gamma_n}{\rho_n} = 0, \quad \sum_n \rho_n\gamma_n = \infty, \quad \rho_n - \rho_{n-1} = o(\rho_n\gamma_n). \tag{6}$$

Then, the sequence $((1-X_n)/\rho_n)_{n\geq 1}$ converges to $(1-p_A)/\pi$ in probability.

Note that the assumptions of Theorem 2 are satisfied if $\gamma_n = C/n^a$ and $\rho_n = C'/n^r$, with $C, C' > 0$, $0 < r < a$ and $a + r < 1$. In fact, we will see that for this choice of parameters, convergence holds with probability one (see Theorem 3).

Before proving Theorem 2, we introduce the notation

$$Y_n = \frac{1-X_n}{\rho_n}.$$

We have, from (2),

$$\begin{aligned}1 - X_{n+1} &= 1 - X_n - \gamma_{n+1}\pi X_n(1-X_n) - \rho_{n+1}\gamma_{n+1}\kappa(X_n) - \gamma_{n+1}\Delta M_{n+1} \\ \frac{1-X_{n+1}}{\rho_{n+1}} &= \frac{1-X_n}{\rho_{n+1}} - \frac{\gamma_{n+1}}{\rho_{n+1}}\pi X_n(1-X_n) - \gamma_{n+1}\kappa(X_n) - \frac{\gamma_{n+1}}{\rho_{n+1}}\Delta M_{n+1}.\end{aligned}$$

Hence

$$\begin{aligned}Y_{n+1} &= Y_n + (1-X_n)\left(\frac{1}{\rho_{n+1}} - \frac{1}{\rho_n} - \frac{\gamma_{n+1}}{\rho_{n+1}}\pi X_n\right) - \gamma_{n+1}\kappa(X_n) - \frac{\gamma_{n+1}}{\rho_{n+1}}\Delta M_{n+1} \\ Y_{n+1} &= Y_n\left(1 + \gamma_{n+1}\varepsilon_n - \pi_n\gamma_{n+1}X_n\right) - \gamma_{n+1}\kappa(X_n) - \frac{\gamma_{n+1}}{\rho_{n+1}}\Delta M_{n+1},\end{aligned}$$

where

$$\varepsilon_n = \frac{\rho_n}{\gamma_{n+1}}\left(\frac{1}{\rho_{n+1}} - \frac{1}{\rho_n}\right) \text{ and } \pi_n = \frac{\rho_n}{\rho_{n+1}}\pi.$$

It follows from the assumption $\rho_n - \rho_{n-1} = o(\rho_n\gamma_n)$ that $\lim_{n\to\infty} \varepsilon_n = 0$ and $\lim_{n\to\infty} \pi_n = \pi$.

Lemma 1. *Assume (6) and consider two positive numbers* π^- *and* π^+ *with* $0 < \pi^- < \pi < \pi^+ < 1$. *Given* $l \in \mathbb{N}$, *let*

$$\nu^l = \inf\{n \geq l \mid \pi_n X_n - \varepsilon_n > \pi^+ \ \ or \ \ \pi_n X_n - \varepsilon_n < \pi^-\}.$$

We have

- $\lim_{l\to\infty} \mathbb{P}(\nu^l = \infty) = 1$,
- *for* $n \geq l$, *if* $\theta_n^+ = \prod_{k=l+1}^n (1 - \pi^+\gamma_k)$ *and* $\theta_n^- = \prod_{k=l+1}^n (1 - \pi^-\gamma_k)$ *(with the convention* $\theta_l^\pm = 1$*),*

$$\frac{Y_{n\wedge\nu^l}}{\theta^-_{n\wedge\nu^l}} \leq Y_l - \sum_{k=l+1}^{n\wedge\nu^l} \frac{\gamma_k}{\theta_k^-}\kappa(X_{k-1}) - \sum_{k=l+1}^{n\wedge\nu^l} \frac{\gamma_k}{\rho_k\theta_k^-}\Delta M_k \tag{7}$$

and

$$\frac{Y_{n\wedge\nu^l}}{\theta^+_{n\wedge\nu^l}} \geq Y_l - \sum_{k=l+1}^{n\wedge\nu^l} \frac{\gamma_k}{\theta_k^+}\kappa(X_{k-1}) - \sum_{k=l+1}^{n\wedge\nu^l} \frac{\gamma_k}{\rho_k\theta_k^+}\Delta M_k. \tag{8}$$

Moreover, with the notation $||\kappa||_\infty = \sup_{0<x<1}|\kappa(x)|$,

$$\sup_{n\geq l} \mathbb{E}\left(Y_n \mathbf{1}_{\{\nu^l \geq n\}}\right) \leq \mathbb{E}\, Y_l + \frac{||\kappa||_\infty}{\pi^-}.$$

Remark 2. Note that, as the proof will show, Lemma 1 remains valid if the condition $\lim_{n\to\infty} \gamma_n/\rho_n = 0$ in (6) is replaced by the boundedness of the sequence $(\gamma_n/\rho_n)_{n\geq 1}$. In particular, the last statement, which implies the tightness of the sequence $(Y_n)_{n\geq 1}$, will be used in Section 3.

Proof: Since $\lim_{n\to\infty}(\pi_n X_n - \varepsilon_n) = \pi$ a.s., we clearly have $\lim_{l\to\infty} \mathbb{P}(\nu^l = \infty) = 1$.

On the other hand, for $l \leq n < \nu^l$, we have

$$Y_{n+1} \leq Y_n(1 - \gamma_{n+1}\pi^-) - \gamma_{n+1}\kappa(X_n) - \frac{\gamma_{n+1}}{\rho_{n+1}}\Delta M_{n+1}$$

and

$$Y_{n+1} \geq Y_n(1 - \gamma_{n+1}\pi^+) - \gamma_{n+1}\kappa(X_n) - \frac{\gamma_{n+1}}{\rho_{n+1}}\Delta M_{n+1},$$

so that, with the notation $\theta_n^+ = \prod_{k=l+1}^n (1 - \pi^+\gamma_k)$ and $\theta_n^- = \prod_{k=l+1}^n (1 - \pi^-\gamma_k)$,

$$\frac{Y_{n+1}}{\theta^-_{n+1}} \leq \frac{Y_n}{\theta_n^-} - \frac{\gamma_{n+1}}{\theta^-_{n+1}}\kappa(X_n) - \frac{\gamma_{n+1}}{\rho_{n+1}\theta^-_{n+1}}\Delta M_{n+1}$$

and

$$\frac{Y_{n+1}}{\theta^+_{n+1}} \geq \frac{Y_n}{\theta_n^+} - \frac{\gamma_{n+1}}{\theta^+_{n+1}}\kappa(X_n) - \frac{\gamma_{n+1}}{\rho_{n+1}\theta^+_{n+1}}\Delta M_{n+1}.$$

By summing up these inequalities, we get (7) and (8).

By taking expectations in (7), we get

$$\begin{aligned}
\mathbb{E}\frac{Y_{n\wedge\nu^l}}{\theta^-_{n\wedge\nu^l}} &\leq \mathbb{E}\,Y_l + ||\kappa||_\infty \mathbb{E}\sum_{k=l+1}^{n\wedge\nu^l}\frac{\gamma_k}{\theta^-_k} \\
&= \mathbb{E}\,Y_l + \frac{||\kappa||_\infty}{\pi^-}\mathbb{E}\sum_{k=l+1}^{n\wedge\nu^l}\left(\frac{1}{\theta^-_k} - \frac{1}{\theta^-_{k-1}}\right) \\
&\leq \mathbb{E}\,Y_l + \frac{||\kappa||_\infty}{\pi^-}\frac{1}{\theta^-_n}.
\end{aligned}$$

We then have

$$\begin{aligned}
\mathbb{E}(Y_n\mathbf{1}_{\{\nu^l\geq n\}}) = \theta^-_n\mathbb{E}\left(\frac{Y_{n\wedge\nu^l}}{\theta^-_{n\wedge\nu^l}}\mathbf{1}_{\{\nu^l\geq n\}}\right) &\leq \theta^-_n\mathbb{E}\frac{Y_{n\wedge\nu^l}}{\theta^-_{n\wedge\nu^l}} \\
&\leq \theta^-_n\left(\mathbb{E}\,Y_l + \frac{||\kappa||_\infty}{\pi^-}\frac{1}{\theta^-_n}\right) \\
&\leq \mathbb{E}\,Y_l + \frac{||\kappa||_\infty}{\pi^-}. \qquad \diamond
\end{aligned}$$

Lemma 2. *Assume (6) and let $(\theta_n)_{n\in\mathbb{N}}$ be a sequence of positive numbers such that $\theta_n = \prod_{k=1}^n(1-p\gamma_k)$ for some $p\in(0,1)$. The sequence $\left(\theta_n\sum_{k=1}^n\frac{\gamma_k}{\theta_k\rho_k}\Delta M_k\right)_{n\in\mathbb{N}}$ converges to 0 in probability.*

Proof: It suffices to show convergence to 0 in probability for the associated conditional variances T_n, defined by

$$T_n = \theta_n^2\sum_{k=1}^n\frac{\gamma_k^2}{\theta_k^2\rho_k^2}\mathbb{E}\left(\Delta M_k^2 \mid \mathcal{F}_{k-1}\right).$$

We know from Proposition 4 that

$$\begin{aligned}
\mathbb{E}\left(\Delta M_k^2 \mid \mathcal{F}_{k-1}\right) &\leq p_A(1-X_{k-1}) + \rho_k^2(1-p_B) \\
&= p_A\rho_{k-1}Y_{k-1} + \rho_k^2(1-p_B).
\end{aligned}$$

Therefore, $T_n \leq p_A T_n^{(1)} + (1-p_B)T_n^{(2)}$, where

$$T_n^{(1)} = \theta_n^2\sum_{k=1}^n\frac{\gamma_k^2}{\theta_k^2\rho_k^2}\rho_{k-1}Y_{k-1}$$

and

$$T_n^{(2)} = \theta_n^2\sum_{k=1}^n\frac{\gamma_k^2}{\theta_k^2}.$$

We first prove that $\lim_{n\to\infty}T_n^{(2)} = 0$. Note that, since $p\gamma_k\leq 1$,

$$\frac{1}{\theta_k^2} - \frac{1}{\theta_{k-1}^2} = \frac{2p\gamma_k - p^2\gamma_k^2}{\theta_k^2} \geq p\frac{\gamma_k}{\theta_k^2}. \tag{9}$$

Therefore,

$$T_n^{(2)} \le \frac{\theta_n^2}{p}\sum_{k=1}^{n}\gamma_k\left(\frac{1}{\theta_k^2}-\frac{1}{\theta_{k-1}^2}\right),$$

and $\lim_{n\to\infty} T_n^{(2)} = 0$ follows from Cesaro's lemma.

We now deal with $T_n^{(1)}$. First note that the assumption $\rho_n - \rho_{n-1} = o(\rho_n\gamma_n)$ implies $\lim_{n\to\infty}\rho_n/\rho_{n-1} = 1$, so that, the sequence $(\gamma_n)_{n\ge 1}$ being non-increasing with limit 0, we only need to prove that $\lim_{n\to\infty}\bar{T}_n^{(1)} = 0$ in probability, where

$$\bar{T}_n^{(1)} = \theta_n^2\sum_{k=1}^{n}\frac{\gamma_k^2}{\theta_k^2\rho_k}Y_k.$$

Now, with the notation of Lemma 1, we have, for $n \ge l > 1$ and $\varepsilon > 0$,

$$\begin{aligned}\mathbb{P}\left(\bar{T}_n^{(1)} \ge \varepsilon\right) &\le \mathbb{P}(\nu^l < \infty) + \mathbb{P}\left(\theta_n^2\sum_{k=1}^{n}\frac{\gamma_k^2}{\theta_k^2\rho_k}Y_k\mathbf{1}_{\{\nu^l=\infty\}} \ge \varepsilon\right)\\ &\le \mathbb{P}(\nu^l < \infty) + \frac{1}{\varepsilon}\theta_n^2\sum_{k=1}^{n}\frac{\gamma_k^2}{\theta_k^2\rho_k}\mathbb{E}\left(Y_k\mathbf{1}_{\{\nu^l=\infty\}}\right).\end{aligned}$$

Using Lemma 1, $\lim_{n\to\infty}\gamma_n/\rho_n = 0$ and (9), we have

$$\lim_{n\to\infty}\theta_n^2\sum_{k=1}^{n}\frac{\gamma_k^2}{\theta_k^2\rho_k}\mathbb{E}\left(Y_k\mathbf{1}_{\{\nu^l=\infty\}}\right) = 0.$$

We also know that $\lim_{l\to\infty}\mathbb{P}(\nu^l < \infty) = 0$. Hence,

$$\lim_{n\to\infty}\mathbb{P}\left(\bar{T}_n^{(1)} \ge \varepsilon\right) = 0. \qquad \diamond$$

Proof of Theorem 2: First note that if $\theta_n = \prod_{k=1}^{n}(1-p\gamma_k)$, with $0 < p < 1$, we have

$$\sum_{k=1}^{n}\frac{\gamma_k}{\theta_k}\kappa(X_{k-1}) = \frac{1}{p}\sum_{k=1}^{n}\left(\frac{1}{\theta_k}-\frac{1}{\theta_{k-1}}\right)\kappa(X_{k-1}).$$

Hence, using $\lim_{n\to\infty}X_n = 1$ and $\kappa(1) = -(1-p_A)$,

$$\lim_{n\to\infty}\theta_n\sum_{k=1}^{n}\frac{\gamma_k}{\theta_k}\kappa(X_{k-1}) = -\frac{1-p_A}{p}.$$

Going back to (7) and (8) and using Lemma 2 with $p = \pi^+$ and π^-, and the fact that $\lim_{l\to\infty}\mathbb{P}(\nu^l = \infty) = 1$, we have, for all $\varepsilon > 0$, $\lim_{n\to\infty}\mathbb{P}(Y_n \ge \frac{1-p_A}{\pi^-}+\varepsilon) = \lim_{n\to\infty}\mathbb{P}(Y_n \le \frac{1-p_A}{\pi^+}-\varepsilon) = 0$, and since π^+ and π^- can be made arbitrarily close to π, the Theorem is proved. $\diamond$

2.2 Almost sure convergence

Theorem 3. *In addition to (6), we assume that for all* $\beta \in [0,1]$,

$$\gamma_n \rho_n^{\beta} - \gamma_{n-1}\rho_{n-1}^{\beta} = o(\gamma_n^2 \rho_n^{\beta}), \tag{10}$$

and that, for some $\eta > 0$, *we have*

$$\forall C > 0, \quad \sum_n \exp\left(-C\frac{\rho_n^{1+\eta}}{\gamma_n}\right) < \infty. \tag{11}$$

Then, with probability 1,

$$\lim_{n\to\infty} \frac{1-X_n}{\rho_n} = \frac{1-p_A}{\pi}.$$

Note that the assumptions of Theorem 3 are satisfied if $\gamma_n = Cn^{-a}$ and $\rho_n = C'n^{-r}$, with $C, C' > 0$, $0 < r < a$ and $a + r < 1$.

The proof of Theorem 3 is based on the following lemma, which will be proved later.

Lemma 3. *Under the assumptions of Theorem 3, let* $\alpha \in [0,1]$ *and let* $(\theta_n)_{n\in\geq 1}$ *be a sequence of positive numbers such that* $\theta_n = \prod_{k=1}^n (1 - p\gamma_k)$, *for some* $p \in (0,1)$. *On the set* $\{\sup_n(\rho_n^{\alpha} Y_n) < \infty\}$, *we have*

$$\lim_{n\to\infty} \theta_n \rho_n^{\frac{\alpha-\eta}{2}-1} \sum_{k=1}^n \frac{\gamma_k}{\theta_k}\Delta M_k = 0 \ \textit{a.s.},$$

where η *satisfies (11).*

Proof of Theorem 3: We start from the following form of (2):

$$1 - X_{n+1} = (1-X_n)(1-\gamma_{n+1}\pi X_n) - \rho_{n+1}\gamma_{n+1}\kappa(X_n) - \gamma_{n+1}\Delta M_{n+1}.$$

We know that $\lim_{n\to\infty} X_n = 1$ a.s.. Therefore, given π^+ and π^-, with $0 < \pi^- < \pi < \pi^+ < 1$, there exists $l \in \mathbb{N}$ such that, for $n \geq l$,

$$1 - X_{n+1} \leq (1-X_n)(1-\gamma_{n+1}\pi^-) - \rho_{n+1}\gamma_{n+1}\kappa(X_n) - \gamma_{n+1}\Delta M_{n+1}$$

and

$$1 - X_{n+1} \geq (1-X_n)(1-\gamma_{n+1}\pi^+) - \rho_{n+1}\gamma_{n+1}\kappa(X_n) - \gamma_{n+1}\Delta M_{n+1},$$

so that, with the notation $\theta_n^+ = \prod_{k=l+1}^n (1-\pi^+\gamma_k)$ and $\theta_n^- = \prod_{k=l+1}^n (1-\pi^-\gamma_k)$,

$$\frac{1-X_{n+1}}{\theta_{n+1}^-} \leq \frac{1-X_n}{\theta_n^-} - \frac{\rho_{n+1}\gamma_{n+1}}{\theta_{n+1}^-}\kappa(X_n) - \frac{\gamma_{n+1}}{\theta_{n+1}^-}\Delta M_{n+1}$$

and

$$\frac{1-X_{n+1}}{\theta_{n+1}^+} \geq \frac{1-X_n}{\theta_n^+} - \frac{\rho_{n+1}\gamma_{n+1}}{\theta_{n+1}^+}\kappa(X_n) - \frac{\gamma_{n+1}}{\theta_{n+1}^+}\Delta M_{n+1}.$$

By summing up these inequalities, we get, for $n \geq l+1$,

$$\frac{1-X_n}{\theta_n^-} \leq (1-X_l) - \sum_{k=l+1}^n \frac{\rho_k\gamma_k}{\theta_k^-}\kappa(X_{k-1}) - \sum_{k=l+1}^n \frac{\gamma_k}{\theta_k^-}\Delta M_k$$

and

$$\frac{1-X_n}{\theta_n^+} \geq (1-X_l) - \sum_{k=l+1}^{n} \frac{\rho_k\gamma_k}{\theta_k^+}\kappa(X_{k-1}) - \sum_{k=l+1}^{n} \frac{\gamma_k}{\theta_k^+}\Delta M_k.$$

Hence

$$Y_n \leq \frac{\theta_n^-}{\rho_n}(1-X_l) - \frac{\theta_n^-}{\rho_n}\sum_{k=l+1}^{n} \frac{\rho_k\gamma_k}{\theta_k^-}\kappa(X_{k-1}) - \frac{\theta_n^-}{\rho_n}\sum_{k=l+1}^{n} \frac{\gamma_k}{\theta_k^-}\Delta M_k \tag{12}$$

and

$$Y_n \geq \frac{\theta_n^+}{\rho_n}(1-X_l) - \frac{\theta_n^+}{\rho_n}\sum_{k=l+1}^{n} \frac{\rho_k\gamma_k}{\theta_k^+}\kappa(X_{k-1}) - \frac{\theta_n^+}{\rho_n}\sum_{k=l+1}^{n} \frac{\gamma_k}{\theta_k^+}\Delta M_k. \tag{13}$$

We have, with probability 1, $\lim_{n\to\infty} \kappa(X_n) = \kappa(1) = -(1-p_A)$, and, since $\sum_{n=1}^{\infty} \rho_n\gamma_n = +\infty$,

$$\sum_{k=l+1}^{n} \frac{\rho_k\gamma_k}{\theta_k^-}\kappa(X_{k-1}) \sim -(1-p_A)\sum_{k=l+1}^{n} \frac{\rho_k\gamma_k}{\theta_k^-}. \tag{14}$$

On the other hand,

$$\begin{aligned}\sum_{k=l+1}^{n} \frac{\rho_k\gamma_k}{\theta_k^-} &= \frac{1}{\pi^-}\sum_{k=l+1}^{n} \rho_k\left(\frac{1}{\theta_k^-} - \frac{1}{\theta_{k-1}^-}\right) \\ &= \frac{1}{\pi^-}\left(\sum_{k=l+1}^{n} (\rho_{k-1}-\rho_k)\frac{1}{\theta_{k-1}^-} + \frac{\rho_n}{\theta_n^-} - \frac{\rho_l}{\theta_l^-}\right) \\ &\sim \frac{1}{\pi^-}\frac{\rho_n}{\theta_n^-},\end{aligned} \tag{15}$$

where we have used the condition $\rho_k - \rho_{k-1} = o(\rho_k\gamma_k)$ and $\sum_{k=l+1}^{\infty} \frac{\rho_k\gamma_k}{\theta_k^-} = +\infty$. We deduce from (14) and (15) that

$$\lim_{n\to\infty} \frac{\theta_n^-}{\rho_n}\sum_{k=1}^{n} \frac{\rho_k\gamma_k}{\theta_k^-}\kappa(X_{k-1}) = -\frac{1-p_A}{\pi^-}$$

and, also, that $\lim_{n\to\infty} \frac{\theta_n^-}{\rho_n} = 0$. By a similar argument, we get $\lim_{n\to\infty} \frac{\theta_n^+}{\rho_n} = 0$ and

$$\lim_{n\to\infty} \frac{\theta_n^+}{\rho_n}\sum_{k=1}^{n} \frac{\rho_k\gamma_k}{\theta_k^+}\kappa(X_{k-1}) = -\frac{1-p_A}{\pi^+}$$

It follows from Lemma 3, that given $\alpha \in [0,1]$, we have, on the set $E_\alpha := \{\sup_n(\rho_n^\alpha Y_n) < \infty\}$,

$$\lim_{n\to\infty} \rho_n^{\frac{\alpha-\eta}{2}-1}\theta_n^{\pm}\sum_{k=1}^{n} \frac{\gamma_k}{\theta_k^{\pm}}\Delta M_k = 0.$$

Together with (12) and (13) this implies

- $\lim_{n\to\infty} Y_n = (1-p_A)/\pi$ a.s., if $\frac{\alpha-\eta}{2} \leq 0$,

- $\lim_{n\to\infty} Y_n \rho_n^{\frac{\alpha-\eta}{2}} = 0$ a.s., if $\frac{\alpha-\eta}{2} > 0$.

We obviously have $\mathbb{P}(E_\alpha) = 1$ for $\alpha = 1$. We deduce from the previous argument that if $\mathbb{P}(E_\alpha) = 1$ and $\frac{\alpha-\eta}{2} > 0$, then $\mathbb{P}(E_{\alpha'}) = 1$, with $\alpha' = \frac{\alpha-\eta}{2}$. Set $\alpha_0 = 1$ and $\alpha_{k+1} = \frac{\alpha_k-\eta}{2}$. If $\alpha_0 \leq \eta$, we have $\lim_{n\to\infty} Y_n = (1-p_A)/\pi$ a.s. on E_{α_0}. If $\alpha_0 > \eta$, let j be the largest integer such that $\alpha_j > \eta$ (note that j exists because $\lim_{k\to\infty} \alpha_k = -\eta < 0$). We have $\mathbb{P}(E_{\alpha_{j+1}}) = 1$, and, on $E_{\alpha_{j+1}}$, $\lim_{n\to\infty} Y_n = (1-p_A)/\pi$ a.s., because $\frac{\alpha_{j+1}-\eta}{2} \leq 0$. ◊

We now turn to the proof of Lemma 3 which is based on the following classical martingale inequality (see [16], remark 1, p.14 for a proof in the case of i.i.d. random variables: the extension to bounded martingale increments is straightforward).

Lemma 4. *(Bernstein's inequality for bounded martingale increments) Let $(Z_i)_{1\leq i\leq n}$ be a finite sequence of square integrable random variables, adapted to the filtration $(\mathcal{F}_i)_{1\leq i\leq n}$, such that*

1. $\mathbb{E}(Z_i \mid \mathcal{F}_{i-1}) = 0$, $i = 1, \ldots, n$,
2. $\mathbb{E}(Z_i^2 \mid \mathcal{F}_{i-1}) \leq \sigma_i^2$, $i = 1, \ldots, n$,
3. $|Z_i| \leq \Delta_n$, $i = 1, \ldots, n$,

where σ_1^2, ..., σ_n^2, Δ_n are deterministic positive constants.
Then, the following inequality holds:

$$\forall \lambda > 0, \quad \mathbb{P}\left(\left|\sum_{i=1}^n Z_i\right| \geq \lambda\right) \leq 2\exp\left(-\frac{\lambda^2}{2\left(b_n^2 + \lambda\frac{\Delta_n}{3}\right)}\right),$$

with $b_n^2 = \sum_{i=1}^n \sigma_i^2$.

We will also need the following technical result.

Lemma 5. *Let $(\theta_n)_{n\geq 1}$ be a sequence of positive numbers such that $\theta_n = \prod_{k=1}^n (1 - p\gamma_k)$, for some $p \in (0,1)$ and let $(\xi_n)_{n\geq 1}$ be a sequence of non-negative numbers satisfying*

$$\gamma_n \xi_n - \gamma_{n-1}\xi_{n-1} = o(\gamma_n^2 \xi_n).$$

We have

$$\sum_{k=1}^n \frac{\gamma_k^2 \xi_k}{\theta_k^2} \sim \frac{\gamma_n \xi_n}{2p\theta_n^2}.$$

Proof: First observe that the condition $\gamma_n\xi_n - \gamma_{n-1}\xi_{n-1} = o(\gamma_n^2\xi_n)$ implies $\gamma_n\xi_n \sim \gamma_{n-1}\xi_{n-1}$ and that, given $\varepsilon > 0$, we have, for n large enough,

$$\begin{aligned} \gamma_n\xi_n - \gamma_{n-1}\xi_{n-1} &\geq -\varepsilon\gamma_n^2\xi_n \\ &\geq -\varepsilon\gamma_{n-1}\gamma_n\xi_n, \end{aligned}$$

where we have used the fact that the sequence (γ_n) is non-increasing. Since $\gamma_n\xi_n \sim \gamma_{n-1}\xi_{n-1}$, we have, for n large enough, say $n \geq n_0$,

$$\gamma_n\xi_n \geq \gamma_{n-1}\xi_{n-1}(1-2\varepsilon\gamma_{n-1}).$$

Therefore, for $n > n_0$,

$$\gamma_n\xi_n \geq \gamma_{n_0}\xi_{n_0}\prod_{k=n_0+1}^{n}(1-2\varepsilon\gamma_{k-1}).$$

From this, we easily deduce that $\lim\limits_{n\to\infty}\gamma_n\xi_n/\theta_n = \infty$ and that $\sum_n \gamma_n^2\xi_n/\theta_n^2 = \infty$.

Now, from

$$\frac{1}{\theta_k^2} - \frac{1}{\theta_{k-1}^2} = \frac{2\gamma_k p - \gamma_k^2 p^2}{\theta_k^2} \sim \frac{2\gamma_k p}{\theta_k^2},$$

we deduce (recall that $\lim\limits_{n\to\infty}\gamma_n = 0$)

$$\frac{\gamma_k^2\xi_k}{\theta_k^2} \sim \frac{\gamma_k\xi_k}{2p}\left(\frac{1}{\theta_k^2} - \frac{1}{\theta_{k-1}^2}\right),$$

and, since $\sum_n \gamma_n^2\xi_n/\theta_n^2 = \infty$,

$$\begin{aligned}
\sum_{k=1}^{n}\frac{\gamma_k^2\xi_k}{\theta_k^2} &\sim \frac{1}{2p}\sum_{k=1}^{n}\gamma_k\xi_k\left(\frac{1}{\theta_k^2} - \frac{1}{\theta_{k-1}^2}\right)\\
&= \frac{1}{2p}\left(\frac{\gamma_n\xi_n}{\theta_n^2} + \sum_{k=1}^{n}(\gamma_{k-1}\xi_{k-1} - \gamma_k\xi_k)\frac{1}{\theta_{k-1}^2}\right)\\
&= \frac{\gamma_n\xi_n}{2p\theta_n^2} + o\left(\sum_{k=1}^{n}\frac{\gamma_k^2\xi_k}{\theta_k^2}\right),
\end{aligned}$$

where, for the first equality, we have assumed $\xi_0 = 0$, and, for the last one, we have used again $\gamma_n\xi_n - \gamma_{n-1}\xi_{n-1} = o(\gamma_n^2\xi_n)$. ◊

Proof of Lemma 3: Given $\mu > 0$, let

$$\nu_\mu = \inf\{n \geq 0 \mid \rho_n^\alpha Y_n > \mu\}.$$

Note that $\{\sup_n \rho_n^\alpha Y_n < \infty\} = \bigcup_{\mu>0}\{\nu_\mu = \infty\}$.

On the set $\{\nu_\mu = \infty\}$, we have

$$\sum_{k=1}^{n}\frac{\gamma_k}{\theta_k}\Delta M_k = \sum_{k=1}^{n}\frac{\gamma_k}{\theta_k}\mathbf{1}_{\{k\leq\nu_\mu\}}\Delta M_k.$$

We now apply Lemma 4 with $Z_i = \frac{\gamma_i}{\theta_i}\mathbf{1}_{\{i\leq\nu_\mu\}}\Delta M_i$. We have, using Proposition 4,

$$\begin{aligned}
\mathbb{E}(Z_i^2 \mid \mathcal{F}_{i-1}) &= \frac{\gamma_i^2}{\theta_i^2}\mathbf{1}_{\{i\leq\nu_\mu\}}\mathbb{E}(\Delta M_i^2 \mid \mathcal{F}_{i-1})\\
&\leq \frac{\gamma_i^2}{\theta_i^2}\mathbf{1}_{\{i\leq\nu_\mu\}}\left(p_A\rho_{i-1}Y_{i-1} + \rho_i^2(1-p_B)\right)\\
&\leq \frac{\gamma_i^2}{\theta_i^2}\left(p_A\rho_{i-1}^{1-\alpha}\mu + \rho_i^2(1-p_B)\right),
\end{aligned}$$

where we have used the fact that, on $\{i \leq \nu_\mu\}$, $\rho_{i-1}^\alpha Y_{i-1} \leq \mu$. Since $\lim_{n\to\infty} \rho_n = 0$ and $\lim_{n\to\infty} \rho_n/\rho_{n-1} = 1$ (which follows from $\rho_n - \rho_{n-1} = o(\gamma_n \rho_n)$), we have

$$\mathbb{E}(Z_i^2 \mid \mathcal{F}_{i-1}) \leq \sigma_i^2,$$

with $\sigma_i^2 = C_\mu \frac{\gamma_i^2 \rho_i^{1-\alpha}}{\theta_i^2}$, for some $C_\mu > 0$, depending only on μ. Using Lemma 5 with $\xi_n = \rho_n^{1-\alpha}$, we have

$$\sum_{i=1}^n \sigma_i^2 \sim C_\mu \frac{\gamma_n \rho_n^{1-\alpha}}{2p\theta_n^2}.$$

On the other hand, we have, because the jumps ΔM_i are bounded,

$$|Z_i| \leq C \frac{\gamma_i}{\theta_i},$$

for some $C > 0$. Note that $\frac{\gamma_k/\theta_k}{\gamma_{k-1}/\theta_{k-1}} = \frac{\gamma_k}{\gamma_{k-1}(1-p\gamma_k)}$, and, since $\gamma_k - \gamma_{k-1} = o(\gamma_k^2)$ (take $\beta = 0$ in (10)), we have, for k large enough, $\gamma_k - \gamma_{k-1} \geq -p\gamma_k\gamma_{k-1}$, so that $\gamma_k/\gamma_{k-1} \geq 1 - p\gamma_k$, and the sequence (γ_n/θ_n) is non-decreasing for n large enough. Therefore, we have

$$\sup_{1\leq i\leq n} |Z_i| \leq \Delta_n,$$

with $\Delta_n = C\gamma_n/\theta_n$ for some $C > 0$. Now, applying Lemma 4 with $\lambda = \lambda_0 \rho_n^{1-\frac{\alpha-\eta}{2}}/\theta_n$, we get

$$\begin{aligned}
\mathbb{P}\left(\theta_n \left|\sum_{k=1}^n \frac{\gamma_k}{\theta_k} \mathbf{1}_{\{k\leq\nu_\mu\}} \Delta M_k\right| \geq \lambda_0 \rho_n^{1-\frac{\alpha-\eta}{2}}\right) &\leq 2\exp\left(-\frac{\lambda_0^2 \rho_n^{2-\alpha+\eta}}{2\theta_n^2 b_n^2 + 2\lambda_0 \theta_n \rho_n^{1-\frac{\alpha-\eta}{2}} \frac{\Delta_n}{3}}\right) \\
&\leq 2\exp\left(-\frac{C_1 \rho_n^{2-\alpha+\eta}}{C_2 \gamma_n \rho_n^{1-\alpha} + C_3 \gamma_n \rho_n^{1-\frac{\alpha-\eta}{2}}}\right) \\
&\leq 2\exp\left(-C_4 \frac{\rho_n^{1+\eta}}{\gamma_n}\right),
\end{aligned}$$

where the positive constants C_1, C_2, C_3 and C_4 depend on λ_0 and μ, but not on n. Using (11) and the Borel-Cantelli lemma, we conclude that, on $\{\nu_\mu = \infty\}$, we have, for n large enough,

$$\theta_n \left|\sum_{k=1}^n \frac{\gamma_k}{\theta_k} \Delta M_k\right| < \lambda_0 \rho_n^{1-\frac{\alpha-\eta}{2}}, \text{ a.s.},$$

and, since λ_0 is arbitrary, this completes the proof of the Lemma. $\diamond$

3 Weak convergence of the normalized algorithm

Throughout this section, we assume (in addition to the initial conditions on the sequence $(\gamma_n)_{n\in\mathbb{N}}$)

$$\gamma_n^2 - \gamma_{n-1}^2 = o(\gamma_n^2) \quad \text{and} \quad \frac{\gamma_n}{\rho_n} = g + o(\gamma_n), \tag{16}$$

where g is a positive constant. Note that a possible choice is $\gamma_n = ag/\sqrt{n}$ and $\rho_n = a/\sqrt{n}$, with $a > 0$.

Under these conditions, we have $\rho_n - \rho_{n-1} = o(\gamma_n^2)$, and we can write, as in the beginning of Section 2,

$$Y_{n+1} = Y_n \left(1 + \gamma_{n+1}\varepsilon_n - \pi_n\gamma_{n+1}X_n\right) - \gamma_{n+1}\kappa(X_n) - \frac{\gamma_{n+1}}{\rho_{n+1}}\Delta M_{n+1}, \tag{17}$$

where $\lim_{n\to\infty} \varepsilon_n = 0$ and $\lim_{n\to\infty} \pi_n = \pi$. As observed in Remark 2, we know that, under the assumptions (16), the sequence $(Y_n)_{n\geq 1}$ is tight. We will prove that it is convergent in distribution.

Theorem 4. *Under conditions (16), the sequence $(Y_n = (1 - X_n)/\rho_n)_{n\in\mathbb{N}}$ converges weakly to the unique stationary distribution of the Markov process on $[0, +\infty)$ with generator L defined by*

$$Lf(y) = p_{\scriptscriptstyle B}\, y\frac{f(y+g) - f(y)}{g} + (1 - p_{\scriptscriptstyle A} - p_{\scriptscriptstyle A}y)f'(y), \quad y \geq 0, \tag{18}$$

for f continuously differentiable and compactly supported in $[0, +\infty)$.

The method for proving Theorem 4 is based on the classical functional approach to central limit theorems for stochastic algorithms (see Bouton [3], Kushner [11], Duflo [7]). The long time behavior of the sequence (Y_n) will be elucidated through the study of a sequence of continuous-time processes $Y^{(n)} = (Y_t^{(n)})_{t\geq 0}$, which will be proved to converge weakly to the Markov process with generator L. We will show that this Markov process has a unique stationary distribution, and that this is the weak limit of the sequence $(Y_n)_{n\in\mathbb{N}}$.

The sequence $Y^{(n)}$ is defined as follows. Given $n \in \mathbb{N}$, and $t \geq 0$, set

$$Y_t^{(n)} = Y_{N(n,t)}, \tag{19}$$

where

$$N(n,t) = \min\left\{m \geq n \mid \sum_{k=n}^{m} \gamma_{k+1} > t\right\},$$

so that $N(n, 0) = n$, for $t \in [0, \gamma_{n+1})$, and, for $m \geq n+1$, $N(n,t) = m$ if and only if $\sum_{k=n+1}^{m} \gamma_k \leq t < \sum_{k=n+1}^{m+1} \gamma_k$.

Theorem 5. *Under the assumptions of Theorem 4, the sequence of continuous time processes $(Y^{(n)})_{n\in\mathbb{N}}$ converges weakly (in the sense of Skorokhod) to a Markov process with generator L.*

The proof of Theorem 5 is done in two steps: in section 3.1, we prove tightness, in section 3.2, we characterize the limit by a martingale problem. In section 3.3, we study the stationary distribution of the limit Markov process and we prove Theorem 4.

3.1 Tightness

It follows from (17) that the process $Y^{(n)}$ admits the following decomposition:

$$Y_t^{(n)} = Y_n + B_t^{(n)} + M_t^{(n)}, \tag{20}$$

with

$$B_t^{(n)} = -\sum_{k=n+1}^{N(n,t)} \gamma_k \left[\kappa(X_{k-1}) + (\pi_{k-1}X_{k-1} - \varepsilon_{k-1})Y_{k-1}\right]$$

and

$$M_t^{(n)} = -\sum_{k=n+1}^{N(n,t)} \frac{\gamma_k}{\rho_k}\Delta M_k.$$

The process $(M_t^{(n)})_{t\geq 0}$ is a square integrable martingale with respect to the filtration $(\mathcal{F}_t^{(n)})_{t\geq 0}$, with $\mathcal{F}_t^{(n)} = \mathcal{F}_{N(n,t)}$, and we have

$$\langle M^{(n)}\rangle_t = \sum_{k=n+1}^{N(n,t)} \left(\frac{\gamma_k}{\rho_k}\right)^2 \mathbb{E}(\Delta M_k^2 \mid \mathcal{F}_{k-1}).$$

We already know (see Remark 2) that the sequence $(Y_n)_{n\in\mathbb{N}}$ is tight. Recall that in order for the sequence $(M^{(n)})$ to be tight, it is sufficient that the sequence $(\langle M^{(n)}\rangle)$ is C-tight (see [8], Theorem 4.13, p. 358, chapter VI). Therefore, the tightness of the sequence $(Y^{(n)})$ in the sense of Skorokhod will follow from the following result.

Proposition 5. *Under the assumptions (16), the sequences $(B^{(n)})$ and $(\langle M^{(n)}\rangle)$ are C-tight.*

For the proof of this proposition,we will need the following lemma.

Lemma 6. *Define ν^l as in Lemma 1, for $l \in \mathbb{N}$. There exists a positive constant C such that, for all $l, n, N \in \mathbb{N}$ with $l \leq n \leq N$, we have*

$$\forall \lambda \geq 1, \quad \mathbb{P}\left(\sup_{n\leq j\leq N} |Y_j - Y_n| \geq \lambda\right) \leq \mathbb{P}(\nu^l < \infty) + C\frac{(1+\mathbb{E}\,Y_l)\left(\sum_{k=n+1}^{N}\gamma_k\right)}{\lambda}.$$

Proof: The function κ being bounded on $[0,1]$, it follows from (17) that there exist positive, deterministic constants a and b such that, for all $n \in \mathbb{N}$,

$$-\gamma_{n+1}(a+bY_n) - \frac{\gamma_{n+1}}{\rho_{n+1}}\Delta M_{n+1} \leq Y_{n+1} - Y_n \leq \gamma_{n+1}(a+bY_n) - \frac{\gamma_{n+1}}{\rho_{n+1}}\Delta M_{n+1}. \qquad (21)$$

We also know from Proposition 4 that

$$\mathbb{E}\left(\Delta M_{n+1}^2 \mid \mathcal{F}_n\right) \leq p_{_A}\rho_n Y_n + (1-p_{_B})\rho_{n+1}^2. \qquad (22)$$

From (21), we derive, for $j \geq n$,

$$|Y_j - Y_n| \leq \sum_{k=n+1}^{j} \gamma_k(a+bY_{k-1}) + \left|\sum_{k=n+1}^{j} \frac{\gamma_k}{\rho_k}\Delta M_k.\right|$$

Let $\widetilde{Y}_k = Y_k \mathbf{1}_{\{k \le \nu^l\}}$ and $\Delta\widetilde{M}_k = \mathbf{1}_{\{k \le \nu^l\}} \Delta M_k$ On the set $\nu^l = \infty$, we have $Y_{k-1} = \widetilde{Y}_{k-1}$ and $\Delta M_k = \Delta\widetilde{M}_k$. Hence

$$\begin{aligned}\mathbb{P}\left(\sup_{n\le j\le N} |Y_j - Y_n| \ge \lambda\right) \quad &\le \quad \mathbb{P}(\nu^l < \infty) + \mathbb{P}\left(\sum_{k=n+1}^{N} \gamma_k (a + b\widetilde{Y}_{k-1}) \ge \lambda/2\right) + \\ &\qquad \mathbb{P}\left(\sup_{n\le j\le N} \left|\sum_{k=n+1}^{j} \frac{\gamma_k}{\rho_k} \Delta\widetilde{M}_k\right| \ge \lambda/2\right).\end{aligned}$$

We have, using Markov's inequality and Lemma 1,

$$\begin{aligned}\mathbb{P}\left(\sum_{k=n+1}^{N} \gamma_k (a + b\widetilde{Y}_{k-1}) \ge \lambda/2\right) \quad &\le \quad \frac{2}{\lambda}\mathbb{E}\sum_{k=n+1}^{N} \gamma_k (a + b\widetilde{Y}_{k-1}) \\ &\le \quad \frac{2}{\lambda}\left(a + b\sup_{k\ge l}\mathbb{E}\left(Y_k \mathbf{1}_{\{\nu^l \ge k\}}\right)\right)\sum_{k=n+1}^{N}\gamma_k \\ &\le \quad \frac{2}{\lambda}\left(b\mathbb{E}\, Y_l + b\frac{||\kappa||_\infty}{\pi^-} + a\right)\sum_{k=n+1}^{N}\gamma_k.\end{aligned}$$

On the other hand, using Doob's inequality,

$$\begin{aligned}\mathbb{P}\left(\sup_{n\le j\le N} \left|\sum_{k=n+1}^{j} \frac{\gamma_k}{\rho_k} \Delta\widetilde{M}_k\right| \ge \lambda/2\right) \quad &\le \quad \frac{16}{\lambda^2}\mathbb{E}\sum_{k=n+1}^{N}\frac{\gamma_k^2}{\rho_k^2}\mathbb{E}\left(\Delta\widetilde{M}_k^2 \mid \mathcal{F}_{k-1}\right) \\ &\le \quad \frac{16}{\lambda^2}\mathbb{E}\sum_{k=n+1}^{N}\frac{\gamma_k^2}{\rho_k^2}\mathbf{1}_{\{k\le\nu^l\}}\left(p_A \rho_{k-1} Y_{k-1} + (1-p_B)\rho_k^2\right).\end{aligned}$$

Using $\lim_n(\gamma_n/\rho_n) = g$, $\rho_{k-1} \sim \rho_k$, $\lim_n \rho_n = 0$ and Lemma 1, we get, for some $C > 0$,

$$\mathbb{P}\left(\sup_{n\le j\le N} \left|\sum_{k=n+1}^{j} \frac{\gamma_k}{\rho_k} \Delta\widetilde{M}_k\right| \ge \lambda/2\right) \le C\frac{(1 + \mathbb{E}\, Y_l)\left(\sum_{k=n+1}^{N}\gamma_k\right)}{\lambda^2},$$

and, since we have assumed $\lambda \ge 1$, the proof of the lemma is completed. ◊

Proof of Proposition 5: Given s and t, with $0 \le s \le t$, we have, using the boundedness of κ,

$$|B_t^{(n)} - B_s^{(n)}| \le \sum_{k=N(n,s)+1}^{N(n,t)} \gamma_k (a + bY_{k-1})$$

for some $a, b > 0$.

Similarly, using (22), we have

$$\left|\langle M^{(n)}\rangle_t - \langle M^{(n)}\rangle_s\right| \le \sum_{k=N(n,s)+1}^{N(n,t)} \gamma_k (a' + b'Y_{k-1})$$

for some $a', b' > 0$. These inequalities express the fact that the processes $B^{(n)}$ and $\langle M^{(n)} \rangle$ are *strongly dominated* (in the sense of [8], definition 3.34) by a linear combination of the processes $X^{(n)}$ and $Z^{(n)}$, where $X_t^{(n)} = \sum_{k=n+1}^{N(n,t)} \gamma_k$ and $Z_t^{(n)} = \sum_{k=n+1}^{N(n,t)} \gamma_k Y_{k-1}$. Therefore, we only need to prove that the sequences $(X^{(n)})$ and $(Z^{(n)})$ are C-tight. This is obvious for the sequence $X^{(n)}$, which in fact converges to the deterministic process t. We now prove that $Z^{(n)}$ is C-tight. We have, for $0 \leq s \leq t \leq T$

$$\begin{aligned} |Z_t^{(n)} - Z_s^{(n)}| &\leq \left(\sup_{n \leq j \leq N(n,T)} Y_j \right) \sum_{k=N(n,s)+1}^{N(n,t)} \gamma_k \\ &\leq (t - s + \gamma_{N(n,s)+1}) \sup_{n \leq j \leq N(n,T)} Y_j \\ &\leq (t - s + \gamma_{n+1}) \sup_{n \leq j \leq N(n,T)} Y_j, \end{aligned}$$

where we have used $\sum_{k=n+1}^{N(n,t)} \gamma_k \leq t$ and $s \leq \sum_{k=n+1}^{N(n,s)+1} \gamma_k$ and the monotony of the sequence $(\gamma_n)_{n \geq 1}$.
Therefore, for $\delta > 0$, and n large enough so that $\gamma_{n+1} \leq \delta$,

$$\begin{aligned} \mathbb{P}\left(\sup_{0 \leq s \leq t \leq T, t-s \leq \delta} |Z_t^{(n)} - Z_s^{(n)}| \geq \eta \right) &\leq \mathbb{P}\left(\sup_{n \leq j \leq N(n,T)} Y_j \geq \frac{\eta}{\delta + \gamma_{n+1}} \right) \\ &\leq \mathbb{P}\left(\sup_{n \leq j \leq N(n,T)} Y_j \geq \frac{\eta}{2\delta} \right) \\ &\leq \mathbb{P}\left(Y_n \geq \frac{\eta}{4\delta} \right) \\ &\quad + \mathbb{P}\left(\sup_{n \leq j \leq N(n,T)} |Y_j - Y_n| \geq \frac{\eta}{4\delta} \right). \end{aligned}$$

For $l \leq n$, we have, from Lemma 6,

$$\begin{aligned} \mathbb{P}\left(\sup_{n \leq j \leq N(n,T)} |Y_j - Y_n| \geq \frac{\eta}{4\delta} \right) &\leq \mathbb{P}(\nu^l < \infty) + \frac{4C\delta}{\eta} (1 + \mathbb{E}\, Y_l) \sum_{k=n+1}^{N(n,T)} \gamma_k \\ &\leq \mathbb{P}(\nu^l < \infty) + \frac{4CT\delta}{\eta} (1 + \mathbb{E}\, Y_l). \end{aligned}$$

We easily conclude from these estimates and Lemma 1 that, given $T > 0$, $\varepsilon > 0$ and $\eta > 0$, we have for n large enough and δ small enough,

$$\mathbb{P}\left(\sup_{0 \leq s \leq t \leq T, t-s \leq \delta} |Z_t^{(n)} - Z_s^{(n)}| \geq \eta \right) < \varepsilon,$$

which proves the C-tightness of the sequence $(Z^{(n)})_{n \geq 0}$. ◇

3.2 Identification of the limit

Lemma 7. *Let f be a C^1 function with compact support in $[0,+\infty)$. We have*

$$\mathbb{E}\left(f(Y_{n+1}) - f(Y_n) \mid \mathcal{F}_n\right) = \gamma_{n+1} Lf(Y_n) + \gamma_{n+1} Z_n, \quad n \in \mathbb{N},$$

where the operator L is defined by

$$Lf(y) = p_{\scriptscriptstyle B}\, y \frac{f(y+g) - f(y)}{g} + (1 - p_{\scriptscriptstyle A} - p_{\scriptscriptstyle A} y) f'(y), \quad y \geq 0, \tag{23}$$

and the sequence $(Z_n)_{n\in\mathbb{N}}$ satisfies $\lim_{n\to\infty} Z_n = 0$ in probability.

Proof: From (17), we have

$$\begin{aligned}
Y_{n+1} &= Y_n + \gamma_{n+1}(-\kappa(1) - \pi Y_n) - \frac{\gamma_{n+1}}{\rho_{n+1}} \Delta M_{n+1} + \gamma_{n+1}\zeta_n \\
&= Y_n + \gamma_{n+1}(1 - p_{\scriptscriptstyle A} - \pi Y_n) - \frac{\gamma_{n+1}}{\rho_{n+1}} \Delta M_{n+1} + \gamma_{n+1}\zeta_n \\
&= Y_n + \gamma_{n+1}(1 - p_{\scriptscriptstyle A} - \pi Y_n) - g\Delta M_{n+1} + \gamma_{n+1}\zeta_n + \left(g - \frac{\gamma_{n+1}}{\rho_{n+1}}\right)\Delta M_{n+1},
\end{aligned} \tag{24}$$

where $\zeta_n = \kappa(1) - \kappa(X_n) + Y_n(\pi - (\pi_n X_n - \varepsilon_n))$, so that ζ_n is $\mathcal{F}_n$-measurable and, using the tightness of (Y_n), $\lim_{n\to\infty} \zeta_n = 0$ in probability. Going back to (3), we rewrite the martingale increment ΔM_{n+1} as follows:

$$\begin{aligned}
\Delta M_{n+1} = {} & -X_n \left(\mathbf{1}_{\{U_{n+1} > X_n\}\cap B_{n+1}} - p_{\scriptscriptstyle B}(1 - X_n)\right) + \rho_n Y_n \left(\mathbf{1}_{\{U_{n+1} \leq X_n\}\cap A_{n+1}} - p_{\scriptscriptstyle A} X_n\right) \\
& -\rho_{n+1}\left(X_n \mathbf{1}_{\{U_{n+1} \leq X_n\}\cap A^c_{n+1}} - (1 - X_n)\mathbf{1}_{\{U_{n+1} > X_n\}\cap B^c_{n+1}} + \kappa(X_n)\right).
\end{aligned}$$

Hence,

$$Y_{n+1} = Y_n + \gamma_{n+1}(1 - p_{\scriptscriptstyle A} - \pi Y_n + \zeta_n) + \xi_{n+1} + \Delta\hat{M}_{n+1},$$

where

$$\xi_{n+1} = g X_n \left(\mathbf{1}_{\{U_{n+1} > X_n\}\cap B_{n+1}} - p_{\scriptscriptstyle B}(1 - X_n)\right)$$

and

$$\begin{aligned}
\Delta\hat{M}_{n+1} = {} & \left(g - \frac{\gamma_{n+1}}{\rho_{n+1}}\right)\Delta M_{n+1} - g\rho_n Y_n \left(\mathbf{1}_{\{U_{n+1} \leq X_n\}\cap A_{n+1}} - p_{\scriptscriptstyle A} X_n\right) \\
& + g\rho_{n+1}\left(X_n \mathbf{1}_{\{U_{n+1} \leq X_n\}\cap A^c_{n+1}} - (1 - X_n)\mathbf{1}_{\{U_{n+1} > X_n\}\cap B^c_{n+1}} + \kappa(X_n)\right).
\end{aligned}$$

Note that, due to our assumptions on γ_n and ρ_n, we have, for some deterministic positive constant C,

$$\left|\Delta\hat{M}_{n+1}\right| \leq C\gamma_{n+1}(1 + Y_n), \quad n \in \mathbb{N}. \tag{25}$$

Now, let

$$\widetilde{Y}_n = Y_n + \gamma_{n+1}(1 - p_{\scriptscriptstyle A} - \pi Y_n + \zeta_n) \text{ and } \bar{Y}_{n+1} = \widetilde{Y}_n + \xi_{n+1},$$

so that $Y_{n+1} = \bar{Y}_{n+1} + \Delta\hat{M}_{n+1}$. We have

$$f(Y_{n+1}) - f(Y_n) = f(Y_{n+1}) - f(\bar{Y}_{n+1}) + f(\bar{Y}_{n+1}) - f(Y_n).$$

We will first show that

$$f(Y_{n+1})-f(\bar{Y}_{n+1})=f'(\widetilde{Y}_n)\Delta\hat{M}_{n+1}+\gamma_{n+1}T_{n+1},\ \text{where}\ \mathbb{P}\text{-}\lim_{n\to\infty}\mathbb{E}(T_{n+1}\mid\mathcal{F}_n)=0, \tag{26}$$

with the notation $\mathbb{P}$- lim for a limit in probability. Denote by w the modulus of continuity of f':

$$w(\delta)=\sup_{|x-y|\leq\delta}|f'(y)-f'(x)|,\quad \delta>0.$$

We have, for some (random) $\theta\in(0,1)$,

$$\begin{aligned} f(Y_{n+1})-f(\bar{Y}_{n+1}) &= f'(\bar{Y}_{n+1}+\theta\Delta\hat{M}_{n+1})\Delta\hat{M}_{n+1}\\ &= f'(\widetilde{Y}_n)\Delta\hat{M}_{n+1}+V_{n+1}, \end{aligned}$$

where $V_{n+1}=\Big(f'(\bar{Y}_{n+1}+\theta\Delta\hat{M}_{n+1})-f'(\widetilde{Y}_n)\Big)\,\Delta\hat{M}_{n+1}$. We have

$$\begin{aligned} |V_{n+1}| &\leq w\left(|\xi_{n+1}|+|\Delta\hat{M}_{n+1}|\right)|\Delta\hat{M}_{n+1}|\\ &\leq Cw\left(|\xi_{n+1}|+C\gamma_{n+1}(1+Y_n)\right)\gamma_{n+1}(1+Y_n), \end{aligned}$$

where we have used $\bar{Y}_{n+1}=\widetilde{Y}_n+\xi_{n+1}$ and (25). In order to get (26), it suffices to prove that $\lim_{n\to\infty}\mathbb{E}\left(w\left(|\xi_{n+1}|+C\gamma_{n+1}(1+Y_n)\right)\mid\mathcal{F}_n\right)=0$ in probability. On the set $\{U_{n+1}>X_n\}\cap B_{n+1}$, we have $|\xi_{n+1}|=gX_n\left(1-p_B(1-X_n)\right)\leq g$, and, on the complement, $|\xi_{n+1}|=gX_np_B(1-X_n)\leq g(1-X_n)$. Hence

$$\begin{aligned} \mathbb{E}\left(w\left(|\xi_{n+1}|+C\gamma_{n+1}(1+Y_n)\right)\mid\mathcal{F}_n\right) &\leq p_B(1-X_n)w\left(g+C\gamma_{n+1}(1+Y_n)\right)\\ &\quad +(1-p_B(1-X_n))w\left(\hat{Y}_n\right), \end{aligned}$$

where $\hat{Y}_n=g(1-X_n)+C\gamma_{n+1}(1+Y_n)$. Observe that $\lim_{n\to\infty}\hat{Y}_n=0$ in probability (recall that $\lim_{n\to\infty}X_n=1$ almost surely). Therefore, we have (26).

We deduce from $\mathbb{E}(\Delta\hat{M}_{n+1}\mid\mathcal{F}_n)=0$ that

$$\mathbb{E}\left(f(Y_{n+1})-f(Y_n)\mid\mathcal{F}_n\right)=\gamma_{n+1}\mathbb{E}(T_{n+1}\mid\mathcal{F}_n)+\mathbb{E}\left(f(\bar{Y}_{n+1})-f(Y_n)\mid\mathcal{F}_n\right),$$

so that the proof will be completed when we have shown

$$\mathbb{P}\text{-}\lim_{n\to\infty}\mathbb{E}\left(\frac{f(\bar{Y}_{n+1})-f(Y_n)-\gamma_{n+1}Lf(Y_n)}{\gamma_{n+1}}\mid\mathcal{F}_n\right)=0. \tag{27}$$

We have

$$\begin{aligned} \mathbb{E}\left(f(\bar{Y}_{n+1})\mid\mathcal{F}_n\right) &= \mathbb{E}\left(f(\widetilde{Y}_n+\xi_{n+1})\mid\mathcal{F}_n\right)\\ &= p_B(1-X_n)f(\widetilde{Y}_n+gX_n(1-p_B(1-X_n)))\\ *[.4em] &\quad +(1-p_B(1-X_n))f(\widetilde{Y}_n-gX_np_B(1-X_n))\\ &= p_B\rho_nY_nf(\widetilde{Y}_n+gX_n(1-p_B(1-X_n)))\\ *[.4em] &\quad +(1-p_B\rho_nY_n)f(\widetilde{Y}_n-gX_np_B(1-X_n)). \end{aligned}$$

Hence

$$\mathbb{E}\left(f(\bar{Y}_{n+1}) - f(Y_n) \mid \mathcal{F}_n\right) = F_n + G_n,$$

with

$$F_n = p_B \rho_n Y_n \left(f(\widetilde{Y}_n + gX_n(1 - p_B(1 - X_n))) - f(Y_n)\right)$$

and

$$G_n = (1 - p_B \rho_n Y_n)\left(f(\widetilde{Y}_n - gX_n p_B(1 - X_n)) - f(Y_n)\right).$$

For the behavior of F_n as n goes to infinity, we use

$$\mathbb{P}\text{-}\lim_{n\to\infty}\left(\widetilde{Y}_n + gX_n(1 - p_B(1 - X_n)) - Y_n - g\right) = 0,$$

and $\lim_{n\to\infty} \rho_n/\gamma_{n+1} = 1/g$, so that

$$\mathbb{P}\text{-}\lim_{n\to\infty}\left(\frac{F_n}{\gamma_{n+1}} - p_B Y_n \frac{f(Y_n + g) - f(Y_n)}{g}\right) = 0.$$

For the behavior of G_n, we write, using $\lim_{n\to\infty} \rho_n/\gamma_{n+1} = 1/g$ again,

$$\begin{aligned}
\widetilde{Y}_n - gX_n p_B(1 - X_n) &= Y_n + \gamma_{n+1}(1 - p_A - \pi Y_n + \zeta_n) - g p_B X_n \rho_n Y_n \\
&= Y_n + \gamma_{n+1}(1 - p_A - p_A Y_n) + \gamma_{n+1}\eta_n,
\end{aligned}$$

with $\mathbb{P}\text{-}\lim_{n\to\infty} \eta_n = 0$, so that, using the fact that f is C^1 with compact support and the tightness of (Y_n),

$$\mathbb{P}\text{-}\lim_{n\to\infty}\left(\frac{G_n}{\gamma_{n+1}} - (1 - p_A - p_A Y_n) f'(Y_n)\right) = 0,$$

which completes the proof of (27). ◇

Proof of Theorem 5: As mentioned before, it follows from Proposition 5 that the sequence of processes $(Y^{(n)})$ is tight in the Skorokhod sense.

On the other hand, it follows from Lemma 7 that, if f is a C^1 function with compact support in $[0, +\infty)$, we have

$$f(Y_n) = f(Y_0) + \sum_{k=1}^{n} \gamma_k Lf(Y_{k-1}) + \sum_{k=1}^{n} \gamma_k Z_{k-1} + M_n,$$

where (M_n) is a martingale and (Z_n) is an adapted sequence satisfying $\mathbb{P}\text{-}\lim_{n\to\infty} Z_n = 0$. Therefore,

$$f(Y_t^{(n)}) - f(Y_0^{(n)}) = M_t^{(n)} + \sum_{k=N(n,0)+1}^{N(n,t)} \gamma_k (Lf(Y_{k-1}) + Z_{k-1}),$$

where $M_t^{(n)} = M_{N(n,t)} - M_{N(n,0)}$. It is easy to verify that $M^{(n)}$ is a martingale with respect to $\mathcal{F}^{(n)}$.

We also have

$$\int_0^t Lf(Y_s^{(n)})ds = \sum_{k=n+1}^{N(n,t)} \gamma_k Lf(Y_{k-1}) + \left(t - \sum_{k=n+1}^{N(n,t)} \gamma_k\right) f(Y_t^{(n)}).$$

Therefore

$$f(Y_t^{(n)}) - f(Y_0^{(n)}) - \int_0^t Lf(Y_s^{(n)})ds = M_t^{(n)} + R_t^{(n)},$$

where $\mathbb{P}\text{-}\lim_{n\to\infty} R_t^{(n)} = 0$. It follows that any weak limit of the sequence $(Y^{(n)})_{n\in\mathbb{N}}$ solves the martingale problem associated with L. From this, together with the study of the stationary distribution of L (see Section 3.3), we will deduce Theorem 4 and Theorem 5. $\diamond$

3.3 The stationary distribution

Theorem 6. *The Markov process $(Y_t)_{t\geq 0}$, on $[0,+\infty)$, with generator L has a unique stationary probability distribution ν. Moreover, ν has a density on $[0,+\infty)$, which vanishes on $(0, r_A]$ (where $r_A = (1-p_A)/p_A$), and is positive and continuous on the open interval $(r_A, +\infty)$. The stationary distribution ν also satisfies the following property: for every compact set K in $[0,+\infty)$, and every bounded continuous function f, we have*

$$\lim_{t\to\infty} \sup_{y\in K} \left| \mathbb{E}_y(f(Y_t)) - \int f\, d\nu \right| = 0, \tag{28}$$

where $\mathbb{E}_y$ refers to the initial condition $Y_0 = y$.

Before proving Theorem 6, we will show how Theorem 4 follows from (28).

Proof of Theorem 4: Fix $t > 0$. For n large enough, we have $\gamma_n \leq t < \sum_{k=1}^n \gamma_k$, so that there exists $\bar{n} \in \{1, \ldots, n-1\}$ such that

$$\sum_{k=\bar{n}+1}^{n} \gamma_k \leq t < \sum_{k=\bar{n}}^{n} \gamma_k.$$

Let $t_n = \sum_{k=\bar{n}+1}^n \gamma_k$. We have

$$0 \leq t - t_n < \gamma_{\bar{n}} \quad \text{and} \quad Y_{t_n}^{(\bar{n})} = Y_n.$$

Since t is fixed, the condition $\sum_{k=\bar{n}+1}^n \gamma_k \leq t$ implies $\lim_{n\to\infty} \bar{n} = \infty$ and $\lim_{n\to\infty} t_n = t$.

Now, given $\varepsilon > 0$, there is a compact set K such that for every weak limit μ of the sequence $(Y_n)_{n\in\mathbb{N}}$, $\mu(K^c) < \varepsilon$. Using (28), we choose t such that

$$\sup_{y\in K} \left| \mathbb{E}_y(f(Y_t)) - \int f d\nu \right| < \varepsilon.$$

Now take a weakly convergent subsequence $(Y_{n_k})_{k\in\mathbb{N}}$. By another subsequence extraction, we can assume that the sequence $(Y^{(\overline{n}_k)})$ converges weakly to a process $Y^{(\infty)}$ which satisfies the martingale problem associated with L. We then have, due to the quasi-left continuity of $Y^{(\infty)}$,

$$\lim_{k\to\infty} \mathbb{E}f(Y^{(\overline{n}_k)}_{t_{n_k}}) = \mathbb{E}f(Y^{(\infty)}_t),$$

for every bounded continuous function f (keep in mind that the functional tightness of $(M^{(n)})$ follows from Theorem 1.13 in [8] which in turn relies on the so-called Aldous criterion; any weak limiting process of such a sequence in the Skorokhod sense is then quasi-left continuous and so is $Y^{(\infty)}$ since every weak limit of the sequence $(B^{(n)})$ is pathwise continuous). Hence $\lim\limits_{k\to\infty} \mathbb{E}f(Y_{n_k}) = \mathbb{E}f(Y^{(\infty)}_t)$. Observe that the law of $Y^{(\infty)}_0$ is a weak limit of the sequence Y_n, so that $\mathbb{P}(Y^{(\infty)}_0 \in K^c) < \varepsilon$. Now we have

$$\mathbb{E}f(Y_{n_k}) - \int f d\nu = \mathbb{E}f(Y_{n_k}) - \mathbb{E}f(Y^{(\infty)}_t) + \mathbb{E}f(Y^{(\infty)}_t) - \int f d\nu,$$

so that, if μ denotes the law of $Y^{(\infty)}_0$,

$$\begin{aligned}\limsup_{k\to\infty}\left|\mathbb{E}f(Y_{n_k}) - \int f d\nu\right| &\leq \left|\mathbb{E}f(Y^{(\infty)}_t) - \int f d\nu\right| \\ &= \left|\int \mathbb{E}_y(f(Y_t))d\mu(y) - \int f d\nu\right| \\ &\leq \varepsilon + 2||f||_\infty \mu(K^c) \\ &\leq \varepsilon(1+2||f||_\infty).\end{aligned}$$

It follows that any weak limit of the sequence $(Y_n)_{n\in\mathbb{N}}$ is equal to ν, which completes the proof of Theorem 4. $\diamondsuit$

For the proof of Theorem 6, we first observe that the generator L depends in an affine way on the state variable y. This *affine* structure suggests that the Laplace transform $\mathbb{E}_y e^{-pY_t}$ has the form $e^{\varphi_p(t)+y\psi_p(t)}$, for some functions φ_p and ψ_p. Affine models have been recently extensively studied in connection with interest rate modelling (see for instance [5] or [6]). The following proposition gives a precise description of the Laplace transform.

Proposition 6. *Let $(Y_t)_{t\geq 0}$ be the Markov process with generator L on $[0,+\infty)$. We have, for $p>0$, $y\in[0,+\infty)$,*

$$\mathbb{E}_y\, e^{-pY_t} = \exp\left(\varphi_p(t) + y\psi_p(t)\right), \tag{29}$$

where ψ_p is the unique solution, on $[0,+\infty)$ of the differential equation

$$\psi' = p_B\frac{e^{g\psi}-1}{g} - p_A\psi, \ \ \textit{with } \psi(0) = -p,$$

and

$$\varphi_p(t) = (1-p_A)\int_0^t \psi_p(s)ds.$$

Before proving the Proposition, we study the involved ordinary differential equation.

Lemma 8. *Given $\psi_0 \in (-\infty, 0]$, the ordinary differential equation*

$$\psi' = p_B \frac{e^{g\psi} - 1}{g} - p_A \psi \tag{30}$$

has a unique solution on $[0, +\infty)$ satisfying the initial condition $\psi(0) = \psi_0$. Moreover, we have

$$\forall t \geq 0, \quad \psi(0) \leq \psi(t)e^{\pi t} \leq 0.$$

PROOF: Existence and uniqueness of a local solution follows from the Cauchy-Lipschitz theorem. In order to prove non-explosion, observe that if ψ solves (30), we have, using the inequality $(e^{g\psi} - 1)/g \geq \psi$,

$$\psi' + \pi\psi \geq 0.$$

Therefore, the function $t \mapsto \psi(t)e^{\pi t}$ is non-decreasing, so that $\psi(0) \leq \psi(t)e^{\pi t}$. Since 0 is an equilibrium of the equation, we have $\psi(t) \leq 0$ if $\psi(0) \leq 0$, and the inequality is strict unless $\psi(0) = 0$. Hence $\psi(0) \leq \psi(t)e^{\pi t} \leq 0$ and the lemma follows easily. ◇

PROOF OF PROPOSITION 6: Let $u_p(t, y) = \exp(\varphi_p(t) + y\psi_p(t))$, where ψ_p and φ_p are defined as in the statement of the Proposition. The existence of ψ_p follows from Lemma 8. An easy computation shows that $\frac{\partial u_p}{\partial t} - Lu_p = 0$ on $[0, +\infty) \times [0, +\infty)$, so that, for $T > 0$, the process $(u_p(T - t, Y_t))_{0 \leq t \leq T}$ is a martingale, and $\mathbb{E}\, u_p(T, Y_0) = \mathbb{E}\, u_p(0, Y_T)$, and the Proposition follows easily. ◇

PROOF OF THEOREM 6:

• Uniqueness of the invariant distribution. We deduce from Lemma 8 that, with the notation of Proposition 6, $|\psi_p(t)| \leq e^{-\pi t}$ and $\lim_{t\to\infty} \varphi_p(t) = (1 - p_A) \int_0^{+\infty} \psi_p(s)ds$. Therefore

$$\lim_{t\to\infty} \mathbb{E}_y(e^{-pY_t}) = \exp\left((1 - p_A) \int_0^{\infty} \psi_p(s)ds\right),$$

and the convergence is uniform on compact sets. This implies the uniqueness of the stationary distribution as well as (28). We also have the Laplace transform of ν:

$$\int_{\mathbb{R}^+} e^{-py}\nu(dy) = \exp\left((1 - p_A) \int_0^{\infty} \psi_p(s)ds\right).$$

Note that, since $\psi_p \leq 0$ and $\psi_p' = p_B \frac{e^{g\psi_p} - 1}{g} - p_A\psi_p$, we have $\psi_p' + p_A\psi_p \leq 0$. Therefore, $\psi_p(t) \leq -pe^{-p_A t}$, and

$$\forall p \geq 0, \quad \int e^{-py}\nu(dy) \leq \exp(-p(1 - p_A)/p_A) = \exp(-pr_A).$$

This yields $\int_{[0,r_A)} e^{p(r_A - y)}\nu(dy) \leq 1$, so that (by taking $p \to +\infty$), $\nu([0, r_A)) = 0$.

• Further properties of the invariant distribution ν. The stationary distribution satisfies $\int Lf d\nu = 0$ for any continuously differentiable function f with compact support in $[0, +\infty)$. This reads

$$\forall f \in C_K^1, \quad \int \left(ry\frac{f(y + g) - f(y)}{g} + (r_A - y)f'(y)\right)\nu(dy) = 0, \tag{31}$$

where $r = p_B/p_A$ and $r_A = (1-p_A)/p_A$.

We first show that $\nu(\{r_A\}) = 0$. Let φ be a non-negative continuously differentiable function satisfying $\varphi = 1$ in a neighbourhood of the origin and $\varphi = 0$ outside the interval $[-1,1]$. For $n \geq 1$ let

$$f_n(y) = \varphi(n(y - r_A)), \quad y \in \mathbb{R}.$$

We have $f_n(y) = 0$ if $|y - r_A| \geq 1/n$. In particular, the support of f_n lies in $[0, +\infty)$, for n large enough. Applying (31) with $f = f_n$, we get

$$\int \left(ry \frac{f_n(y+g) - f_n(y)}{g} + (r_A - y) n \varphi'(n(y - r_A)) \right) \nu(dy) = 0.$$

Observe that $\lim_{n\to\infty} f_n = \mathbf{1}_{\{r_A\}}$ so that

$$\lim_{n\to\infty} \int y(f_n(y+g) - f_n(y))\nu(dy) = (r_A - g)\nu(\{r_A - g\}) - r_A \nu(\{r_A\}) = -r_A \nu(\{r_A\}),$$

where we have used $\nu(-\infty, r_A) = 0$. On the other hand, we have $|(r_A - y) n\, \varphi'(n(y - r_A))| \leq \sup_{u\in\mathbb{R}}(u\varphi'(u))$, and $\lim_{n\to\infty}(n\varphi'(n(y - r_A))) = 0$, so that, by dominated convergence,

$$\lim_{n\to\infty} \int (r_A - y) n\, \varphi'(n(y - r_A))\nu(dy) = 0.$$

Hence $\nu(\{r_A\}) = 0$.

We now study the measure ν on the open interval $(r_A, +\infty)$. Denote by $\mathcal{D}$ the set of all infinitely differentiable functions with compact support in $(r_A, +\infty)$. We deduce from (31) that, for $f \in \mathcal{D}$,

$$\frac{r}{g} \int \nu(dy) y f(y+g) - \frac{r}{g} \int \nu(dy) y f(y) + \int \nu(dy)(r_A - y) f'(y) = 0. \tag{32}$$

Denote by ν_g the measure defined by $\int \nu_g(dy) f(y) = \int \nu(dy) f(y+g)$. We deduce from (32) that ν satisfies the following equation in the sense of distributions:

$$(y - r_A)\nu' + (1 - (r/g)y)\nu = -\frac{r}{g}(y - g)\nu_g,$$

or

$$\nu' + \frac{1 - (r/g)y}{y - r_A} \nu = -\frac{r}{g} \frac{y - g}{y - r_A} \nu_g. \tag{33}$$

Denote by F the function defined by

$$F(y) = e^{ry/g}(y - r_A)^{d-1}, \quad y > r_A, \tag{34}$$

where $d = r\, r_A/g$. We have

$$F'(y) = -\frac{1 - (r/g)y}{y - r_A} F(y),$$

so that the equation satisfied by ν reads

$$\left(\frac{1}{F}\nu\right)' = \frac{G}{F}\nu_g, \tag{35}$$

where the function G is defined by $G(y) = -\frac{r}{g}\frac{y-g}{y-r_A}$.

On the set (r_A, r_A+g), the measure ν_g vanishes, so that $\nu = \lambda_0 F$ for some non-negative constant λ_0. At this point, we know that the restriction of the measure ν to the set $(0, r_A+g)$ has a density which vanishes on $(0, r_A)$ and is given by $\lambda_0 F$ on (r_A, r_A+g).

We will prove by induction that the distribution ν coincides with a continuous function on (r_A, r_A+ng), which is infinitely differentiable on $(r_A+(n-1)g, r_A+ng)$. The claim has been proved for $n=1$. Assume that it is true for n. On the set $(r_A, r_A+(n+1)g)$, the distributional derivative of $(1/F)\nu$ coincides with the function $y \mapsto (G(y)/F(y))\nu(y-g)$, which is locally integrable on (r_A, r_A+ng+g), continuous on (r_A+g, r_A+ng+g), and infinitely differentiable on (r_A+ng, r_A+ng+g), due to the induction hypothesis (there may be a discontinuity at r_A+g if $d<1$). It follows that $(1/F)\nu$ is a continuous (resp. infinitely differentiable) function, and so is ν on $(r_A, r_A+(n+1)g)$ (resp. (r_A+ng, r_A+ng+g)). We have proved that ν has a continuous density on $(r_A, +\infty)$, which is infinitely differentiable on the open set $\bigcup_{n=1}^{\infty}(r_A+(n-1)g, r_A+ng)$.

Finally, we prove that the density of ν is positive on $(r_A, +\infty)$. Note that $G(y)<0$ if $y>g$ and that the density vanishes at $y-g$ if $y<g$. Therefore $\left(\frac{1}{F}\nu\right)' \le 0$, so that the function $y \mapsto \nu(y)/F(y)$ is non-decreasing. It follows that λ_0 cannot be zero (otherwise ν would be identically zero). Hence $\nu(y)>0$ for $y \in (r_A, r_A+g)$. Now, if $\nu(y)>0$ for $y \in (r_A+ng-g, r_A+ng)$, the function ν/F is strictly decreasing on (r_A+ng, r_A+ng+g) and, therefore, cannot vanish. So, by induction, the density is positive on $(r_A, +\infty)$. This completes the proof of Theorem 6. $\diamond$

Additional remarks • The proof of Theorem 6 provides a bit more information on the invariant distribution ν. Let $g>0$ and let ϕ_g denote its continuous density on $(r_A, +\infty)$: the function ϕ_g is $\mathcal{C}^\infty$ on $[r_A, +\infty) \setminus (r_A + g\,\mathbb{N})$ and it follows from (34) and the definitions of r and r_A (and $d = rr_A/g$, see the proof of Theorem 6) that

$$\phi_g(r_A) = +\infty \ \text{ if } \ g > g^*, \quad \phi_g(r_A) \in (0, +\infty) \ \text{ if } \ g = g^* \ \text{ and } \ \phi_g(r_A) = 0 \ \text{ if } \ g < g^*$$

where $g^* = \frac{p_B(1-p_A)}{p_A^2} \in (0, \frac{1-p_A}{p_A})$. As concerns the regularity of the density ϕ_g at points $y \in r_A + g\,\mathbb{N}$, one easily derives from Equation (33) that for every m, $k \in \mathbb{N}$,

– ϕ_g is C^{m+k} at $r_A + kg$ as soon as $g < \frac{g^*}{m+1}$,

– the $(m+k)^{th}$ derivative $\phi_g^{(m+k)}$ is only right and left continuous at $r_A + kg$ if $g = \frac{g^*}{m+1}$.

• One can characterize the finite positive exponential moments of ν by slightly extending the proof of Proposition 6 (Laplace transform). For every $y>1$, let $\theta(y)$ denote the unique (strictly) positive solution of the equation

$$\frac{e^\theta - 1}{\theta} = y.$$

Note that $\log y < \theta(y) < 2(y-1)$ and that $\lim_{y\to 1}\frac{\theta(y)}{2(y-1)} = 1$ and $\lim_{y\to\infty}\frac{\theta(y)}{\log y} = 1$. The result is as follows

$$\int e^{py}\nu(dy) < +\infty \quad \text{ if and only if } \quad p < p_g^* := g\,\theta(p_A/p_B). \tag{36}$$

With the notations of Proposition 6, it follows from Fatou's Lemma that

$$\forall\, p>0, \qquad \int e^{py}\nu(dy) \le \liminf_{t\to\infty} \mathbb{E}_y(e^{pY_t}). \tag{37}$$

We know that

$$\mathbb{E}_y(e^{pY_t}) = e^{\widetilde{\varphi}_p(t)+y\widetilde{\psi}_p(t)}$$

with $\widetilde{\varphi}_p(t) = (1-p_A)\int_0^t \widetilde{\psi}_p(s)ds$ and $\widetilde{\psi}_p$ is solution on the non-negative real line (if any) of

$$\psi'(t) = G(\psi(t)), \quad \psi(0) = p \quad \text{with} \quad G(u) = -p_A u + \frac{p_B}{g}(e^{gu}-1).$$

The function G is convex on $\mathbb{R}_+$ and satisfies $G(0) = G(p_g^*) = 0$, $G((0,p_g^*)) \subset (-\infty,0)$.
Let $p\in (0,p_g^*)$. The convexity of G implies

$$\forall\, u\in[0,p], \quad \frac{G(u)}{u} \le \frac{G(p)}{p} < 0.$$

It follows that $\widetilde{\psi}_p$ does exist on $\mathbb{R}_+$ and satisfies $0\le \widetilde{\psi}_p(t) \le pe^{\frac{G(p)t}{p}}$ (hence it goes to 0 when t goes to infinity). One derives that

$$\lim_{t\to+\infty} \widetilde{\varphi}_p(t) = (1-p_A)\int_0^{+\infty} \widetilde{\psi}_p(t)dt \le -(1-p_A)\frac{p^2}{G(p)}.$$

Combining this with (37) yields

$$\int e^{py}\nu(dy) \le e^{-(1-p_A)\frac{p^2}{G(p)}} < +\infty.$$

On the other hand if $p=p_g^*$, $\widetilde{\psi}_p(t) = p_g^*$ and $\widetilde{\varphi}_p(t) = (1-p_A)p_g^* t$. Consequently

$$\forall\, t\ge 0, \qquad \int e^{p_g^* y}\nu(dy) = \int \mathbb{E}_y(e^{p_g^* Y_t})\nu(dy) = e^{(1-p_A)p_g^* t}\int e^{p_g^* y}\nu(dy).$$

Now the right hand side of this equality goes to ∞ as t goes to infinity since $(1-p_A)p_g^*>0$ which shows that $\int e^{p_g^* y}\nu(dy) = +\infty$ (since it cannot be 0).

- One has, in accordance with the convergence rate result obtained for $\rho_n = o(\gamma_n)$, that

$$\int y\,\nu(dy) = \frac{1-p_A}{\pi}.$$

To prove this claim, one first notes, using the definition (18) of the generator L, that $L(Id)(y) = 1-p_A-\pi\, y$. Hence the above claim will follow from $\int L(Id)(y)\nu(dy) = 0$. Let $\varphi:\mathbb{R}_+\to\mathbb{R}_+$ denote a continuously differentiable function such that $\varphi(y) = y$ if $y\in[0,1]$, $\varphi(y)=0$ if $y\ge 2$ and φ' is bounded on $\mathbb{R}_+$. Set $\varphi_n(y) = n\varphi(y/n)$, $n\ge 1$. One checks that $L(\varphi_n)\to L(Id)$ as n

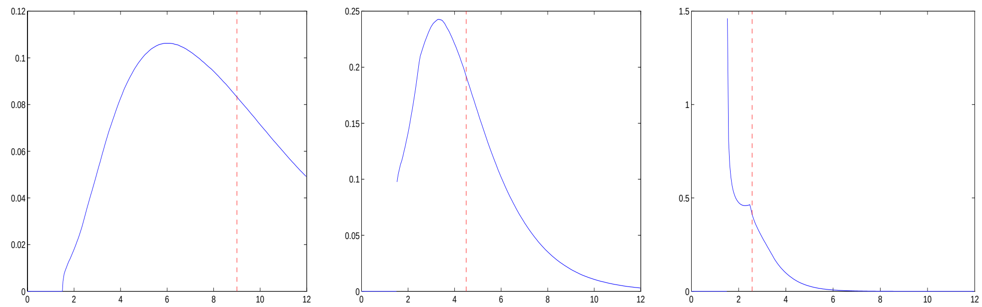

Figure 1: *Graphs of the p.d.f* ϕ_g*,* $p_{\!{}_A} = 2/5$*,* $g = 1$*; the vertical dotted line shows the mean* $\frac{1-p_A}{\pi}$ *of* ν*.* Left: $p_{\!{}_B} = 1/3$ *(*$g^* > g = 1$*).* Center: $p_{\!{}_B} = 4/15$ *(*$g^* = g = 1$*).* Right: $p_{\!{}_B} = 1/6$ *(*$g^* < g = 1$*).*

goes to infinity and $|L(\varphi_n)(y)| \le ay + b$ for some positive real constants a, b. One derives by the dominated convergence theorem that

$$\int L(Id)(y)\nu(dy) = \lim_n \int L(\varphi_n)(y)\nu(dy) = 0$$

where we used that the function φ_n has compact support on $[0, +\infty)$. One shows similarly that $\displaystyle\int L(u \mapsto u^2)(y)\nu(dy) = 0$ to derive that

$$\int \left(y - \frac{1-p_{\!{}_A}}{\pi}\right)^2 \nu(dy) = g\,\frac{p_{\!{}_B}(1-p_{\!{}_A})}{2\pi^2}.$$

Note that, as one could expect, this variance goes to 0 as $g \to 0$. As a conclusion, we present in Figure 1 three examples of shape for ϕ_g. They were obtained from an exact simulation of the Markov process $(Y_t)_{t\ge 0}$ (associated to the generator L) at its jump times: we approximated the p.d.f. by a histogram method using Birkhoff's ergodic Theorem.

A final remark about the case $\pi = 0$ and $\gamma_n = g\,\rho_n$. In that setting (see Remark 1) the asymptotics of the algorithm cannot be elucidated by using the *ODE* approach since it holds in a weak sense. Setting $Y_n = 1 - 2X_n$ one checks that $Y_n \in [-1, 1]$ and

$$Y_{n+1} = Y_n(1 - 2g\rho_{n+1}^2(1-p_{\!{}_A})) - 2g\,\rho_{n+1}\Delta M_{n+1}$$

and that $\mathbb{E}((\Delta M_{n+1})^2\,|\mathcal{F}_{n+1}) = \frac{p_A}{4}(1 - Y_n^2) + O(\rho_{n+1}^2)$. Then, a similar approach as that developed in this section (but significantly less technical since (Y_n) is bounded by 1) shows that Y_n converges in distribution to the invariant distribution μ of the Brownian diffusion with generator $\mathcal{L}f(y) = -2g(1-p_{\!{}_A})yf'(y) + \frac{1}{2}g^2 p_{\!{}_A}(1-y^2)f''(y)$. In that case, it is well-known that μ has a density function for which a closed form is available (see [9]), namely

$$\mu(dy) = m(y)dy \quad \text{with} \quad m(y) = C_{g,r_{\!{}_A}}(1-y^2)^{\frac{2r_A}{g}-1}\mathbf{1}_{(-1,1)}(y).$$

Note that when $g = 2r_{\!{}_A} = 2(1/p_{\!{}_A} - 1) > 0$, μ is but the uniform distribution over $[-1, 1]$.

References

[1] Z.-D. Bai, F. Hu (2005), Asymptotics in randomized urn models, *Annals of Applied Probability*, **15**, 914-940. MR2114994

[2] M. Benaïm (1999), Dynamics of Stochastic approximation Algorithms, *Séminaire de Probabilités XXXIII*, J. Azéma, M. Émery, M. Ledoux, M. Yor éds., Lecture Notes in Mathematics n^{0}1709, pp.1-68. MR1767993

[3] C. Bouton (1988), Approximation gaussienne d'algorithmes stochastiques à dynamique markovienne, *Ann. Inst. Henri Poincaré, Probab. Stat.*, **24**(1), 131-155. MR0937959

[4] L. Dubins and D. Freedman (1965), A sharper form of the Borel-Cantelli lemma and the strong law, *Ann. of Math. Stat.*, **36**, 800-807. MR0182041

[5] D. Duffie, J. Pan, K. Singleton (2000), Transform Analysis and Asset Pricing for Affine Jump-Diffusions, *Econometrica*, **68**, 1343-1376. MR1793362

[6] D. Duffie, D. Filipovic, W. Schachermayer (2003), Affine processes and applications in finance. *Ann. Appl. Probab.*, **13**(3), 984-1053 MR1994043

[7] M. Duflo (1996), Algorithmes stochastiques, coll. *Mathématiques & Applications*, **23**, Springer-Verlag, Berlin, 319p. MR1612815

[8] J. Jacod, A.N. Shiryaev (2003), Limit Theorems for Stochastic Processes, 2^{nd} edition, *Fundamental Principles of Mathematical Sciences*, **28**, Springer-Verlag, Berlin, 661p. MR1943877

[9] S. Karlin, H.M. Taylor (1981), *A second course in stochastic processes*, Academic Press, New-York. MR0611513

[10] H.J. Kushner, D.S. Clark (1978), Stochastic Approximation for Constrained and Unconstrained Systems, *Applied Math. Science Series*, **26**, Springer-Verlag, New York. MR0499560

[11] H.J. Kushner, G.G. Yin (2003), Stochastic approximation and recursive algorithms and applications, 2^{nd} edition, *Applications of Mathematics, Stochastic Modelling and Applied Probability*, **35**, Springer-Verlag, New York. MR1993642

[12] S. Lakshmivarahan (1979), ε-optimal Learning Algorithms-Non-Absorbing Barrier Type, Univ. Oklahoma, School of Electrical Engineering and Computing Science, Techn. Report EECS 7901.

[13] S. Lakshmivarahan (1981), *Learning Algotithms: Theory and Applications*, New York, Springer-verlag. MR0666244

[14] D. Lamberton, G. Pagès (2005), How fast is the bandit?, pre-print LPMA-1018, Univ. Paris 6, forthcoming in *Stochastic Analysis and Applications*.

[15] D. Lamberton, G. Pagès, P. Tarrès (2004), When can the two-armed bandit algorithm be trusted?, *Annals of Applied Probability*, **14**(3), 1424-1454. MR2071429

[16] P. Massart (2003), *St-Flour Lecture Notes*, Cours de l'école d'été de Saint-Flour 2003, pre-print, Univ. Paris-Sud (France), http://www.math.u-psud.fr/ massart/flour.pdf.

[17] K.S. Narendra, M.A.L. Thathachar (1974), Learning Automata - A survey, *IEEE Trans. Systems, Man., Cybernetics*, S.M.C-4, 323-334. MR0469583

[18] K.S. Narendra, M.A.L. Thathachar (1989), *Learning Automata - An introduction*, Prentice Hall, Englewood Cliffs, NJ, 476p.

[19] M.F. Norman (1968), On linear Models with Two Absorbing Barriers, *J. of Mathematical Psychlogy*, **5**, 225-241. MR0226954

[20] R. Pemantle (1990), Non-convergence to unstable points in urn models and stochastic approximations, *Annals of Probability*, **18**, n^0 2, 698-712. MR1055428

[21] I.J. Shapiro, K.S. Narendra (1969), Use of Stochastic Automata for Parameter Self-Optimization with Multi-Modal Perfomance Criteria, *IEEE Trans. Syst. Sci. and Cybern.*, SSC-5, 352-360.

[22] P. Tarrès, *Algorithmes stochastiques et marches aléatoires renforcées*, Thèse de l'ENS Cachan (France), novembre 2001.

[23] P. Tarrès, P. Vandekerkhove, On the ergodic two armed-bandit algorithm, *preprint*, 2006.