How fast is the bandit?

Damien Lamberton, Gilles Pagès

How fast is the bandit? *

Damien Lamberton † Gilles Pagès ‡

Abstract

In this paper we investigate the rate of convergence of the so-called two-armed bandit algorithm in a financial context of asset allocation. The behaviour of the algorithm turns out to be highly non-standard: no CLT whatever the time scale, possible existence of two rate regimes.

Key words: Two-armed bandit algorithm, Stochastic Approximation, learning automata, asset allocation.

2001 AMS classification: 62L20, secondary 93C40, 91E40, 68T05, 91B32 91B32

Introduction

In a recent joint work with P. Tarrès (see [6]), we studied the convergence of the so-called two-armed bandit algorithm. In the terminology of learning theory (see *e.g.* [9, 10]) this algorithm is a Linear Reward Inaction (LRI) scheme. Viewed as a Markovian Stochastic Approximation (SA) recursive procedure, it appears as the simplest example of an algorithm having two possible limits – its *target* and a *trap* – both *noiseless*. In SA theory a target is a *stable* equilibrium of the Ordinary Differential Equation (ODE) associated to the *mean* function of the algorithm, a trap being an unstable one. Various results from SA theory show that an algorithm never "falls" into a noisy trap (see *e.g.* [8, 13, 2, 3, 14]. We established in [6] that the two-armed bandit algorithm can be either *infallible* (*i.e.* converging to its target with probability one, starting from any initial value except the trap itself) or *fallible*. This depends on the speed at which the (deterministic) *learning rate parameter* goes to 0.

Our aim on this paper is to investigate the rate of convergence of the algorithm, toward either of its limits. In fact, the algorithm behaves in a highly non standard way among SA procedures. In particular, this rate is never ruled by a Central Limit Theorem

*This work has benefitted from the stay of both authors at the Isaac Newton Institute on the program *Developments in Quantitative Finance*.

†Laboratoire d'analyse et de mathématiques appliquées, UMR 8050, Univ. Marne-la-Vallée, Cité Descartes, 5 Bld Descartes, Champs-sur-Marne, F-77454 Marne-la-Vallée Cedex 2. damien.lamberton@univ-mlv.fr

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

(CLT). Furthermore, this study will provide some new insight on the infallibility problem as it will be seen further on. However our motivations are not only theoretical but also practical in connection with the financial context in which the algorithm was presented in [6], namely a procedure for the optimal allocation of a fund between the two traders who manage it. Imagine that the owner of a fund can share his wealth between two traders, say A and B, and that, every day, he can evaluate the results of one of the traders and, subsequently, modify the percentage of the fund managed by both traders. Denote by X_n the percentage managed by trader A at time n. We assume that the owner selects the trader to be evaluated at random, in such a way that the probability that A is evaluated at time n is X_n, in order to select preferably the trader in charge of the greater part of the fund. In the LRI scheme, if the evaluated trader performs well, its share is increased by a fraction $\gamma_n \in (0,1)$ of the share of the other trader, and nothing happens if the evaluated trader performs badly. Therefore, the dynamics of the sequence $(X_n)_{n\geq 0}$ can be modelled as follows:

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

where $(U_n)_{n\geq 1}$ is an i.i.d. sequence of uniform random variables on the interval $[0,1]$, A_n (resp. B_n) is the event "trader A (resp. trader B) performs well at time n". We assume $\mathbb{P}(A_n) = p_A$, $\mathbb{P}(B_n) = p_B$, for $n \geq 1$, with p_A, $p_B \in (0,1)$, and independence between these events and the sequence $(U_n)_{n\geq 1}$. The point is that the owner of the fund does not know the parameters p_A, p_B. Note that this procedure is $[0,1]$-valued and that 0 and 1 are absorbing states. The γ_n parameter is the *learning* rate of the procedure (we will say from now on *reward* to take into account the modelling context).

This recursive learning procedure has been designed in order to assign progressively the whole fund to the best trader when $p_A \neq p_B$. From now on we will assume without loss of generality that $p_A > p_B$. This means that X_n is expected to converge toward its target 1 with probability 1 provided $X_0 \in (0,1)$ (and consequently never to get trapped in 0). However this "infallibility" property needs some very stringent assumption on the reward parameter γ_n: thus, if $\gamma_n = \left(\frac{C}{C+n}\right)^{\alpha}$, $n \geq 1$, with $0 < \alpha \leq 1$ and $C > 0$, it is shown in [6] (see Corollary 1(b)) that the algorithm is infallible if and only if $\alpha = 1$ and $C \leq \frac{1}{p_B}$.

In a standard SA framework, when an algorithm is converging to its target – *i.e.* a zero x^* of its mean function $h(x) = \frac{\mathbb{E}(X_{n+1}-X_n\,|\,X_n=x)}{\gamma_n}$, stable for the ODE $\dot{x} = h(x)$ – its rate is ruled by a CLT at a $\sqrt{\gamma_n}$-rate with an asymptotic variance $\sigma^2_{x^*}$ related to the asymptotic excitation of x^* by the noise (see [1, 5, 12]).

As concerns the two-armed bandit algorithm, there is no exciting noise at 1 (nor at 0 indeed). This is made impossible simply because both equilibrium points lie at the boundary of the state space $[0,1]$ of the algorithm (otherwise the algorithm would leave the unit interval when getting too close to its boundary). This same feature which causes the fallibility of the algorithm when γ_n goes to 0 too slowly also induces its non-standard rate of convergence.

To illustrate this behaviour and consider again the steps $\gamma_n = \frac{C}{C+n}$, $n \geq 1$, with $C > 0$. As a consequence of our main results, one obtains:

- If $C > \frac{1}{p_B}$ the algorithm is fallible with positive probability from any $x \in [0,1)$ and, when failing, it goes to 0 at a n^{-Cp_B}-rate. The rate of convergence to 1 may vary according to the parameters, see Section 4.

- If $\frac{1}{p_A - p_B} \leq C \leq \frac{1}{p_B}$ (this case requires that $2\,p_B \leq p_A$), the algorithm is infallible from any $x \in (0,1]$ and goes to 1 at a n^{-Cp_A}-rate.

- If $\frac{1}{p_A} < C < \frac{1}{p_A - p_B}$ then the algorithm is infallible (from any $x \in (0,1]$) and *two rates of convergence to* 1 *may occur with positive* $\mathbb{P}_x$*-probability*: a "slow" one – $n^{-C(p_A - p_B)}$ – and a "fast" one – n^{-Cp_A}.

- If $C \leq \frac{1}{p_A}$ then the algorithm is still infallible from any $x \in (0,1]$ but only the slowest rate of convergence "survives" *i.e.* $n^{-C(p_A - p_B)}$.

In fact the following rule holds true: *the greater the real constant C is, the faster the algorithm (X_n) converges, except that when C is too great, then the algorithm becomes fallible* which makes the two-armed bandit a very "moral" procedure. Furthermore, note that the "blind" choice – $C = 1$ – which ensures infallibility induces a *slow rate* of convergence $n^{-C(p_A - p_B)}$ since then $C \leq \frac{1}{p_A}$ (by contrast with the *fast* rate n^{-Cp_A}). Also note that this rate is precisely that of the *mean* algorithm $x_{n+1} = x_n + \gamma_n(p_A - p_B)x_n(1 - x_n)$. A last feature to be noticed is that the switching between rate regimes takes place "progressively" as the parameter C grows since it happens that two different rates coexist with positive probability.

For more exhaustive results, we refer to Section 4. If one thinks again of a practical implementation of the algorithm, the only reasonable choice for the reward parameter is $\gamma_n = \frac{1}{n+1}$: it ensures infallibility regardless of the (unknown) values of p_A and p_B. But when these two parameters become too close, the rate of convergence becomes too poor to remain really efficient. Unfortunately, this is more or less the standard situations: the daily performances of the traders are usually close and this can be extended to other fields where this procedure can be used (experimental psychology, clinical trials, industrial reliability, . . .). One clue to get rid of this dependency is to introduce a "fading" penalization in the procedure when an evaluated trader has unsatisfactory performances. (By fading we mean negligible with respect to the reward in order to preserve traders' motivation). This variant of the two-armed bandit algorithm which satisfies a pseudo-CLT at a (weak) $n^{-\frac{1}{2}}$-rate whatever the parameter p_A and p_B is described and investigated in [7].

The paper is organized as follows: Section 1 is devoted to some preliminary results and technical tools. Section 2 is devoted to the rate of convergence when the algorithm converges to its trap 0 whereas Section 3 deals with the rate of convergence toward its target 1. Section 4 proposes a summing up of the results for a natural parameterized family of reward parameter γ_n.

NOTATIONS: • 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 $a_n = b_n + o(b_n)$.
• The notation $\mathbb{P}_x$ is used in reference to $X_0 = x$.

1 Preliminary results

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

$$X_{n+1} = X_n + \gamma_{n+1}\left(\mathbb{1}_{\{U_{n+1}\le X_n\}\cap A_{n+1}}(1-X_n) - \mathbb{1}_{\{U_{n+1}>X_n\}\cap B_{n+1}}X_n\right), \quad n \in \mathbb{N}.$$

Here $(\gamma_n)_{n\ge1}$ is a sequence of nonnegative numbers satisfying

$$\gamma_n < 1 \quad \text{and} \quad \Gamma_n = \sum_{k=1}^{n} \gamma_k \to +\infty \quad \text{as} \quad n \to \infty,$$

$(U_n)_{n\ge1}$ 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 $0 < p_{_B} < p_{_A} < 1$, and the sequences $(U_n)_{n\ge1}$ and $(\mathbb{1}_{A_n}, \mathbb{1}_{B_n})_{n\ge1}$ are independent. The natural filtration of the sequence $(U_n, \mathbb{1}_{A_n}, \mathbb{1}_{B_n})_{n\ge1}$ is denoted by $(\mathcal{F}_n)_{n\ge0}$ and we set

$$\pi = p_{_A} - p_{_B} > 0.$$

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

$$X_{n+1} = X_n + \gamma_{n+1}\pi X_n(1-X_n) + \gamma_{n+1}\Delta M_{n+1}, \tag{1}$$

where $\Delta M_{n+1} = M_{n+1} - M_n$, and the sequence $(M_n)_{n\ge0}$ is the martingale defined by $M_0 = 0$ and

$$\Delta M_{n+1} = \mathbb{1}_{\{U_{n+1}\le X_n\}\cap A_{n+1}}(1-X_n) - \mathbb{1}_{\{U_{n+1}>X_n\}\cap B_{n+1}}X_n - \pi X_n(1-X_n).$$

One derives from (1) that (X_n) is a $[0,1]$-valued super-martingale. Hence it converges *a.s.* and in L^1 to a limit X_∞. Consequently

$$\sum_n \gamma_n X_n(1-X_n) < +\infty \qquad a.s.$$

which in turn shows that $X_\infty = 0$ or 1 with probability 1. One easily checks (see [6]) that 1 is a stable equilibrium of the so-called *mean ODE* $\equiv \dot{x} = \pi\, x(1-x)$ with attracting basin $(0,1]$ and 0 is a repulsive equilibrium of this *ODE* (whence the terminology: 1 is a target and 0 is a *trap*, see [6] for more details).

The conditional variance process of the martingale (M_n) will play a crucial role in our analysis, and we will often use the following estimates.

Proposition 1 *We have, for $n \ge 0$,*

$$p_{_B} X_n(1-X_n) \le \mathbb{E}\left(\Delta M_{n+1}^2 \mid \mathcal{F}_n\right) \le p_{_A} X_n(1-X_n).$$

PROOF: We have

$$\begin{aligned}
\mathbb{E}\left(\Delta M_{n+1}^2 \mid \mathcal{F}_n\right) &= p_{_A} X_n(1-X_n)^2 + p_{_B}(1-X_n)X_n^2 - \pi^2 X_n^2(1-X_n)^2 \\
&= X_n(1-X_n)\left(p_{_A}(1-X_n) + p_{_B} X_n - \pi^2 X_n(1-X_n)\right) \\
&\le X_n(1-X_n)\left(p_{_A}(1-X_n) + p_{_B} X_n\right) \\
&\le p_{_A} X_n(1-X_n),
\end{aligned}$$

where the last inequality follows from $p_{_B} \le p_{_A}$. For the lower bound, note that

$$\begin{aligned}
p_{_A}(1-X_n) + p_{_B} X_n - \pi^2 X_n(1-X_n) &= (1-X_n)(p_{_A} - \pi^2 X_n) + p_{_B} X_n \\
&\ge (1-X_n)(p_{_A} - \pi) + p_{_B} X_n = p_{_B},
\end{aligned}$$

where we have used $\pi X_n \le 1$. $\diamond$

2 Convergence to the trap

We first prove that, under rather general conditions, as soon as the sequence converges to the trapping state 0, it goes to it very fast in the sense that the series $\sum_n X_n$ is convergent.

Proposition 2 *If*

$$\liminf_n \frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n} > -\pi \tag{2}$$

then

$$\forall\, x \in (0,1), \qquad \{X_\infty = 0\} = \{\sum_n X_n < +\infty\} \qquad \mathbb{P}_x\text{-}a.s.$$

Note that (2) is satisfied if the sequence $(\gamma_n)_{n\ge 1}$ is nonincreasing (for large enough n).

PROOF OF PROPOSITION 2: Denote by E the event $\{X_\infty = 0\} \cap \{\sum_n X_n = +\infty\}$. We want to prove that $\mathbb{P}_x(E) = 0$. We first show that on E,

$$\liminf_{n\to\infty} \frac{X_n}{\gamma_n \sum_{k=1}^n X_{k-1}} > 0. \tag{3}$$

We deduce from (1) that

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

By summing up and setting $\gamma_0 = \gamma_1$, we derive

$$\frac{X_n}{\gamma_n} = \frac{x}{\gamma_1} + \sum_{k=1}^n \left(\frac{1}{\gamma_k} - \frac{1}{\gamma_{k-1}} + \pi(1-X_{k-1})\right) X_{k-1} + M_n.$$

From Proposition 1, we know that the conditional variance process of (M_n) satisfies

$$p_B \sum_{k=1}^{n} X_{k-1}(1-X_{k-1}) \leq \, <M>_n \, \leq p_A \sum_{k=1}^{n} X_{k-1}(1-X_{k-1}).$$

Therefore, on E, we have $<M>_\infty = +\infty$ a.s., and using the law of large numbers for martingales, we deduce that

$$\lim_{n\to\infty} \frac{M_n}{\sum_{k=1}^{n} X_{k-1}} = 0, \quad \text{a.s. on } E.$$

The estimate (3) then follows easily from the assumption (2).

Now let $S_n = \sum_{k=1}^{n} X_k$. Note that, on E, $S_n \sim \sum_{k=1}^{n} X_{k-1}$, so that, using (3),

$$\exists C > 0, \quad \forall n \geq 1, \quad \gamma_n \leq C \frac{X_n}{S_n}.$$

This implies

$$\sum_n \gamma_n^2 \leq C^2 \sum_n \frac{X_n^2}{S_n^2} \leq C^2 \sum_n \frac{X_n}{S_n^2} < +\infty,$$

where we have used $X_n \leq 1$. We also know from Proposition 9 of [6] (see (29) in particular) that, on the set $\{X_n \to 0\}$,

$$\limsup_{n\to\infty} \frac{X_n}{\sum_{k\geq n} \gamma_{k+1}^2} < +\infty \quad \text{a.s.}$$

Hence $X_n \leq C \sum_{k\geq n} \gamma_{k+1}^2$ for some $C > 0$, and, by plugging in the estimate $\gamma_{k+1} \leq C X_{k+1}/S_{k+1}$ we derive

$$\begin{aligned} X_n &\leq C \sum_{k\geq n} \frac{X_{k+1}^2}{S_{k+1}^2} \\ &\leq C \left(\sup_{k\geq n} X_{k+1} \right) \sum_{k\geq n} \frac{X_{k+1}}{S_{k+1}^2} \\ &\leq C \frac{\sup_{k\geq n} X_{k+1}}{S_n}. \end{aligned}$$

On the set E, we have $\lim_{n\to\infty} S_n = +\infty$, so, for n large enough, say $n \geq N$, we have

$$X_n \leq \frac{\sup_{k\geq n} X_{k+1}}{2}.$$

Now, by taking n to be the largest integer such that $X_n \geq X_N$ (which exists on $\{X_n \to 0\}$ because $X_N > 0$), we reach a contradiction, which proves that $\mathbb{P}_x(E) = 0$. $\diamond$

Our next result shows that under (3), there is essentially only one way for (X_n) to go to 0.

Proposition 3 *Assume (2).*
(a) Let $x \in (0,1)$. *Then*

$$\mathbb{P}_x(X_\infty = 0) > 0 \quad \Longleftrightarrow \quad \mathbb{P}\Big(\sum_{n\ge 1}\prod_{k=1}^{n}(1-\mathbf{1}_{B_k}\gamma_k) < +\infty\Big) > 0 \tag{4}$$

and, on the event $\{X_\infty = 0\}$, *there exists a (random) integer* $n_0 \ge 1$ *such that*

$$\forall\, n \ge n_0, \qquad X_n = X_{n_0}\prod_{k=n_0+1}^{n}(1-\mathbf{1}_{B_k}\gamma_k). \qquad a.s. \tag{5}$$

Note that, as a special case of (4),

$$\sum_{n\ge 1}\prod_{k=1}^{n}(1-p_B\gamma_k) < +\infty \quad \Longrightarrow \quad \mathbb{P}_x(X_\infty = 0) > 0. \tag{6}$$

(b) Furthermore, if $\sum_{n\ge 1}\gamma_n^2 < +\infty$, *(4) reads*

$$\mathbb{P}_x(X_\infty = 0) > 0 \quad \Longleftrightarrow \quad \sum_{n\ge 1}\prod_{k=1}^{n}(1-p_B\gamma_k) < +\infty$$

and moreover there is a random variable $\Xi_x > 0$ *such that*

$$X_n \sim \Xi_x \prod_{k=1}^{n}(1-p_B\gamma_k) \qquad a.s. \quad on \quad \{X_\infty = 0\}.$$

Remark 1 If $\sum_{n\ge 1}\gamma_n^2 = +\infty$, a weaker (but still tractable) sufficient condition for $\mathbb{P}_x(X_\infty = 0)$ is given by

$$\exists\, \rho \in (0, p_B(1-p_B)/2), \qquad \sum_{n\ge 1} e^{-\rho\Gamma_n^{(2)}}\prod_{k=1}^{n}(1-p_B\gamma_k) < +\infty$$

where $\Gamma_n^{(2)} = \sum_{1\le k\le n}\gamma_k^2$ (see the proof of Proposition 3). Then, on the set $\{X_n \to 0\}$, for every $\eta \in (0, p_B(1-p_B)/2)$,

$$X_n = o\left(e^{-(\frac{p_B(1-p_B)}{2}-\eta)\Gamma_n^{(2)}}\prod_{k=1}^{n}(1-p_B\gamma_k)\right).$$

Remark 2 Note that the condition in (4) which characterizes fallibility does not depend on x: if the algorithm is fallible for one $x \in (0,1)$ then it is for any such x.

Proof of Proposition 3: (a) It follows from Proposition 2 and the conditional Borel-Cantelli Lemma that $\mathbb{P}_x$-$a.s.$

$$\{X_n \to 0\} = \{\sum_{n\geq 0} \mathbf{1}_{\{U_{n+1}\leq X_n\}} < +\infty\} = \bigcup_{n\geq 0} \bigcap_{k\geq n} \{U_{k+1} > X_k\}\,. \tag{7}$$

The sequence of events $\left(\bigcap_{k\geq n} \{U_{k+1} > X_k\}\right)_{n\geq 1}$ being non-decreasing, we have

$$\mathbb{P}_x(X_n \to 0) = \lim_{n\to\infty} \mathbb{P}_x\left(\bigcap_{k\geq n} \{U_{k+1} > X_k\}\right),$$

and the left-hand side is positive if and only if, for some integer $n \geq 1$,

$$\mathbb{P}_x\left(\bigcap_{k\geq n} \{U_{k+1} > X_k\}\right) > 0.$$

From the definition of the sequence (X_n), we get (with the convention $\prod_\emptyset = 1$),

$$\bigcap_{k\geq n} \{U_{k+1} > X_k\} = \bigcap_{k\geq n} \left\{U_{k+1} > X_k \text{ and } X_k = X_n \prod_{\ell=n+1}^{k} (1 - \mathbf{1}_{B_\ell}\gamma_\ell)\right\} \tag{8}$$

$$= \bigcap_{k\geq n} \left\{U_{k+1} > X_n \prod_{\ell=n+1}^{k} (1 - \mathbf{1}_{B_\ell}\gamma_\ell)\right\}. \tag{9}$$

Note that (5) follows from (7) and (8). Now, denote by $\mathcal{B}_n$ the σ-field generated by the random variable X_n and the events, B_k, $k \geq n$. We have

$$\mathbb{P}_x\left(\bigcap_{k\geq n} \left\{U_{k+1} > X_n \prod_{\ell=n+1}^{k} (1 - \mathbf{1}_{B_\ell}\gamma_\ell)\right\} \mid \mathcal{B}_n\right) = \prod_{k=n}^{\infty} \left(1 - X_n \prod_{l=n+1}^{k} (1 - \mathbf{1}_{B_\ell}\gamma_\ell)\right),$$

and the infinite product is positive if and only if

$$\sum_k \prod_{l=n+1}^{k} (1 - \mathbf{1}_{B_\ell}\gamma_\ell) < +\infty.$$

This clearly implies (4). The sufficient condition (6) follows from the equality

$$\mathbb{E}\left(\sum_{n\geq 1} \prod_{1\leq k\leq n} (1 - \mathbf{1}_{B_k}\gamma_k)\right) = \sum_{n\geq 1} \prod_{1\leq k\leq n} (1 - p_B\gamma_k).$$

(b) (and proof of the remark) If $\sum_{n\geq 1} \gamma_n^2 < +\infty$, then, a straightforward argument (see [6], proof of Lemma 2) shows that

$$\prod_{k=1}^{n} \left(\frac{1 - \mathbf{1}_{B_k}\gamma_k}{1 - p_B\gamma_k}\right) \longrightarrow \xi > 0 \qquad \text{a.s.} \qquad n \to +\infty.$$

This proves claim (b).

When $\sum_{n\geq 1}\gamma_n^2=+\infty$, one checks that

$$\log\prod_{k=1}^{n}\left(\frac{1-\mathbf{1}_{B_k}\gamma_k}{1-p_B\gamma_k}\right)=M_n^B-\sum_{k=1}^{n}(\frac{1}{2}p_B(1-p_B)+\varepsilon_k)\gamma_k^2.$$

where ε_k is random variable bounded by $c\gamma_k$ (c real constant) and

$$M_n^B=\sum_{k=1}^{n}(\mathbf{1}_{B_k}-p_B)\gamma_k(1-\gamma_k/2)$$

is a martingale with bounded increments satisfying $<M^B>_n\sim p_B(1-p_B)\Gamma_n^{(2)}\to+\infty$. Then

$$M_n^B=o\left(\Gamma_n^{(2)}\right)$$

since $\frac{M_n^B}{<M^B>_n}\to 0$ as $n\to\infty$. Consequently, $\mathbb{P}$-$a.s.$, there exists a finite random variable ξ such that

$$\prod_{k=1}^{n}(1-\mathbf{1}_{B_k}\gamma_k)\leq\xi\exp\left(-(\frac{1}{2}p_B(1-p_B)+o(1))\Gamma_n^{(2)}\right)\prod_{k=1}^{n}(1-p_B\gamma_k)$$

where $o(1)$ denotes a random variable $\mathbb{P}$-$a.s.$ going to 0 as $n\to\infty$. The sufficient condition given in the remark follows straightforwardly as well as the rate of convergence of X_n. $\diamond$

3 Convergence to the target

In order to study the rate of convergence to 1, we first rewrite (1) as follows:

$$1-X_{n+1}=(1-X_n)(1-\gamma_{n+1}\pi X_n)-\gamma_{n+1}\Delta M_{n+1}. \tag{10}$$

Now let

$$\theta_n=\prod_{k=1}^{n}(1-\gamma_k\pi X_{k-1}),\quad Y_n=(1-X_n)/\theta_n,\quad n\in\mathbb{N}.$$

Proposition 4 (a) *The sequence $(Y_n)_{n\in\mathbb{N}}$ is a non-negative martingale.*
(b) *On the set $\{X_\infty=1\}$, we have*

$$\lim_{n\to\infty}\frac{1-X_n}{\prod_{k=1}^{n}(1-\pi\gamma_k)}=\xi Y_\infty$$

almost surely, where ξ is a finite positive random variable and $Y_\infty=\lim_{n\to\infty}Y_n$.

PROOF: The first assertion follows from the equality

$$Y_{n+1}=Y_n-\frac{\gamma_{n+1}}{\theta_{n+1}}\Delta M_{n+1},$$

and the fact that the sequence $(\theta_n)_{n\in\mathbb{N}}$ is predictable.

As a non-negative martingale, the sequence $(Y_n)_{n\in\mathbb{N}}$ has a limit Y_∞, which satisfies $Y_\infty \geq 0$ a.s. and $\mathbb{E}(Y_\infty) < +\infty$.

Recall that $\sum_n \gamma_n X_{n-1}(1-X_{n-1}) < +\infty$ almost surely. Therefore, on $\{X_\infty = 1\}$, we have $\sum_n \gamma_n(1-X_{n-1}) < +\infty$ a.s., which implies that the sequence $\prod_{k=1}^n \frac{1-\pi\gamma_k X_{k-1}}{1-\pi\gamma_k}$ has a positive and finite limit and the second assertion of the Proposition follows easily. $\diamondsuit$

Remark 3 Note that, with the notation $\Gamma_n = \sum_{k=1}^n \gamma_k$, we have $\prod_{k=1}^n(1-\pi\gamma_k) \leq e^{-\pi\Gamma_n}$. Therefore, we deduce from Proposition 4 that, on the set $\{X_\infty = 1\}$, $1 - X_n = O(e^{-\pi\Gamma_n})$ almost surely. If we have $\sum_n \gamma_n^2 < +\infty$, the sequence $\left(e^{\pi\Gamma_n}\prod_{k=1}^n(1-\pi\gamma_k)\right)$ converges to a positive limit, so that, on the set $\{X_\infty = 1\}$, we have $\lim_{n\to\infty} e^{\pi\Gamma_n}(1-X_n) = \xi' Y_\infty$, with $\xi' \in (0,+\infty)$ almost surely.

On the other hand, on $\{X_\infty = 0\}$, the sequence $(\theta_n)_{n\in\mathbb{N}}$ itself converges to an almost surely positive limit, so that $\{Y_\infty = 0\} \subset \{X_\infty = 1\}$.

Proposition 5 *(a) If $\sum_n \gamma_n^2 e^{\pi\Gamma_n} < +\infty$, the martingale $(Y_n)_{n\in\mathbb{N}}$ is bounded in L^2 and its limit satisfies $\mathbb{E}(X_\infty Y_\infty) > 0$. Moreover, on the set $\{Y_\infty = 0\}$, we have*

$$\limsup_{n\to\infty} \frac{Y_n}{\sum_{k\geq n}\gamma_{k+1}^2 e^{\pi\Gamma_{k+1}}} < +\infty \tag{11}$$

almost surely.

(b) If

$$\sum_n \gamma_n^2 e^{\pi\Gamma_n} = +\infty \quad \textit{and} \quad \sup_{n\geq 1}\gamma_n e^{\pi\Gamma_n} < +\infty, \tag{12}$$

then, for every $x \in (0,1)$,

$$\{X_\infty = 1\} = \{Y_\infty = 0\} \quad \mathbb{P}_x\text{-a.s.}$$

Remark 4 It follows from Proposition 5 and Remark 3 that, if $\sum_n \gamma_n^2 e^{\pi\Gamma_n} < +\infty$, on the set $\{X_\infty = 1\} \cap \{Y_\infty > 0\}$ (which has positive probability) the sequence $((1-X_n)e^{\pi\Gamma_n})_{n\in\mathbb{N}}$ converges to a positive limit almost surely.

Remark 5 We also derive from the inequality $(1-X_{n+1}) \geq (1-X_n)\left(1-\gamma_{n+1}\mathbb{1}_{\{U_{n+1}\leq X_n\}\cap A_{n+1}}\right)$ that

$$1 - X_n \geq (1-x)\prod_{k=1}^n(1-\gamma_k \mathbb{1}_{A_k}) \geq Ce^{-p_A\Gamma_n},$$

for some real constant $C > 0$, if $\sum_n \gamma_n^2 < +\infty$. Therefore, we deduce from Proposition 5 that if $\lim_{n\to\infty}\left(e^{p_B\Gamma_n}\sum_{k\geq n}\gamma_{k+1}^2 e^{\pi\Gamma_{k+1}}\right) = 0$, then $\mathbb{P}(Y_\infty = 0) = 0$. On the other hand, the second part of Proposition 5 shows that, in some cases, we may have $1-X_n = o(e^{-\pi\Gamma_n})$, and we need to investigate what the real rate of convergence is in such cases: see Proposition 7.

Proof of Proposition 5: (a) Assume $\sum_n \gamma_n^2 e^{\pi\Gamma_n} < +\infty$. In order to prove L_2-boundedness, we estimate the conditional variance process. Using Proposition 1, we have

$$\begin{aligned}
\mathbb{E}\left((Y_{n+1}-Y_n)^2 \mid \mathcal{F}_n\right) &= \frac{\gamma_{n+1}^2}{\theta_{n+1}^2}\mathbb{E}\left(\Delta M_{n+1}^2 \mid \mathcal{F}_n\right) \\
&\le \frac{\gamma_{n+1}^2}{\theta_{n+1}^2} p_{_A} X_n(1-X_n) \\
&= \frac{\gamma_{n+1}^2}{\theta_{n+1}^2} p_{_A} X_n \theta_n Y_n \\
&\le p_{_A} \frac{\gamma_{n+1}^2}{\theta_n(1-\pi\gamma_{n+1})^2} Y_n \\
&\le p_{_A} \frac{\gamma_{n+1}^2}{(1-\pi\gamma_{n+1})^2 \prod_{k=1}^n (1-\pi\gamma_k)} Y_n \\
&\le C\, p_{_A} \gamma_{n+1}^2 e^{\pi\Gamma_{n+1}} Y_n, \qquad (13)
\end{aligned}$$

where we have used the inequality $\theta_n \ge \prod_{k=1}^n (1-\pi\gamma_k)$ and the fact that, since we have $\sum_{n\ge 1}\gamma_n^2 < +\infty$, $\prod_{k=1}^n (1-\pi\gamma_k) \ge e^{-\pi\Gamma_n}/C$ for some $C>0$. Note that $\sup_{n\in\mathbb{N}} \mathbb{E}\, Y_n < +\infty$. Therefore, the convergence of the series $\sum_n \gamma_n^2 e^{\pi\Gamma_n}$ implies that $(Y_n)_{n\in\mathbb{N}}$ is bounded in L_2.

In order to prove $\mathbb{E}(X_\infty Y_\infty) > 0$, we consider the conditional covariance

$$\begin{aligned}
\mathbb{E}_x\left((1-X_n)X_n \,|\, \mathcal{F}_{n-1}\right) &= X_{n-1}(1-X_{n-1})\left(1+\pi\,\gamma_n(1-2X_{n-1}) + \pi\,\gamma_n^2 X_{n-1} - p_{_A}\gamma_n^2\right) \\
&\ge X_{n-1}(1-X_{n-1})\left(1-\pi\,\gamma_n X_{n-1} - p_{_A}\gamma_n^2\right)
\end{aligned}$$

so that

$$\begin{aligned}
\mathbb{E}_x\left(X_n Y_n \,|\, \mathcal{F}_{n-1}\right) &\ge X_{n-1}Y_{n-1}\left(1-\frac{p_{_A}\gamma_n^2}{1-\pi\gamma_n X_{n-1}}\right) \\
&\ge X_{n-1}Y_{n-1}\left(1-\frac{p_{_A}\gamma_n^2}{1-\pi\gamma_n}\right).
\end{aligned}$$

For n large enough (say $n \ge n_0$), we have $1 > \frac{p_A\gamma_n^2}{1-\pi\gamma_n}$ and, by induction, for $n \ge n_0$,

$$\mathbb{E}_x(X_n Y_n) \ge \mathbb{E}_x X_{n_0} Y_{n_0} \prod_{k=n_0+1}^{n}\left(1-\frac{p_{_A}\gamma_k^2}{1-\pi\gamma_k}\right).$$

Now, using that $Y_n \to Y_\infty$ and $X_n \to X_\infty$ in $L^2(\mathbb{P})$, and $\sum_n \gamma_n^2 < +\infty$, one finally gets $\mathbb{E}_x(X_\infty Y_\infty) > 0$. Note that this implies that $\mathbb{P}_x(X_\infty = 1, Y_\infty > 0) > 0$ since $X_\infty = \mathbf{1}_{\{X_\infty = 1\}}$.

The first step to establish (11) is to apply to the martingale $(Y_n)_{n\ge 1}$ an approach originally developed in [6] to establish the infallibility property for (X_n): for every $n \ge 1$,

$$\begin{aligned}
\mathbb{P}(Y_\infty = 0 \,|\, \mathcal{F}_n) &= \frac{1}{Y_n^2}\,\mathbb{E}_x(\mathbf{1}_{\{Y_\infty = 0\}}(Y_\infty - Y_n)^2 \,|\, \mathcal{F}_n) \\
&\le \frac{1}{Y_n^2}\sum_{k\ge n+1}\mathbb{E}_x((Y_k - Y_{k-1})^2 \,|\, \mathcal{F}_n).
\end{aligned}$$

Plugging (13) in the above inequality and using that $\mathbb{E}_x(Y_k \,|\, \mathcal{F}_n) = Y_n$ for every $k \geq n$ yield,

$$\mathbb{P}_x(Y_\infty = 0 \,|\, \mathcal{F}_n) \leq \frac{C p_A}{Y_n} \sum_{k \geq n+1} \gamma_k^2 e^{\pi \Gamma_k}.$$

On the other hand the martingale $\mathbb{P}_x(Y_\infty = 0 \,|\, \mathcal{F}_n)$ converges $\mathbb{P}_x$-*a.s.* toward $\mathbf{1}_{\{Y_\infty = 0\}}$. The announced result follows easily.

(b) We now assume $\sum_n \gamma_n^2 e^{\pi \Gamma_n} = +\infty$ and $\sup_n \gamma_n e^{\pi \Gamma_n} < +\infty$. Note that the latter condition implies $\gamma_n^2 \leq C \gamma_n e^{-\pi \Gamma_n}$ for some $C > 0$, so that $\sum_n \gamma_n^2 < +\infty$. On the other hand, we have

$$\begin{aligned} |Y_n - Y_{n-1}| &= \frac{\gamma_n}{\theta_n} |\Delta M_n| \\ &\leq \frac{\gamma_n}{\prod_{k=1}^n (1 - \pi \gamma_k)} |\Delta M_n| \leq C \gamma_n e^{\pi \Gamma_n} |\Delta M_n|, \end{aligned}$$

so that the martingale $(Y_n)_{n \geq 1}$ has bounded increments. Consequently the Law of Iterated Logarithm (cf. [4]) implies that $\liminf_n Y_n = -\infty$ on the event $\{< Y >_\infty = +\infty\}$, and, since $Y_n \geq 0$, we deduce thereof that $\{< Y >_\infty < +\infty\}$ almost surely. On the other hand, we have, using Proposition 1 and the inequality $\theta_n \leq e^{-\pi \Gamma_n}$,

$$\begin{aligned} \Delta\!< Y >_n &= \frac{\gamma_n^2}{\theta_n^2} \mathbb{E}\left(\Delta M_n^2 \mid \mathcal{F}_{n-1}\right) \\ &\geq \frac{\gamma_n^2}{\theta_n (1 - \pi \gamma_n)} p_B X_{n-1} Y_{n-1} \\ &\geq C X_{n-1} Y_{n-1} \gamma_n^2 e^{\pi \Gamma_n}. \end{aligned}$$

Therefore, the assumption (12) implies that $Y_\infty = 0$ on the event $\{X_\infty = 1\}$. ◊

In order to clarify what happens when $Y_\infty = 0$, we first observe that we have, up to null events,

$$\begin{aligned} \left\{ \sum_n (1 - X_n) < +\infty \right\} &= \left\{ \sum_n \mathbb{1}_{\{U_n > X_n\}} < +\infty \right\} \\ &\subset \bigcup_{m \geq 1} \bigcap_{n \geq m} \left\{ 1 - X_n = (1 - X_m) \prod_{k=m+1}^{n} (1 - \mathbb{1}_{A_k} \gamma_k) \right\}, \end{aligned}$$

so that, on the set $\{\sum_n (1 - X_n) < +\infty\}$, we have

$$1 - X_n \sim \xi \prod_{k=1}^{n} (1 - \mathbb{1}_{A_k} \gamma_k) \quad \text{a.s.},$$

where ξ is a positive random variable. Recall that, if $\sum_n \gamma_n^2 < +\infty$, $\prod_{k=1}^n (1 - \mathbb{1}_{A_k} \gamma_k) \sim \xi' e^{-p_A \Gamma_n}$, for some (random) $\xi' > 0$. We thus see that, on the set $\{\sum_n (1 - X_n) < +\infty\}$, we have a "fast" rate of convergence. The possibility of occurrence of this fast rate is characterized in the following Proposition.

Proposition 6 *We have, for all $x \in (0,1)$,*

$$\mathbb{P}_x(\sum_n (1 - X_n) < +\infty) > 0 \quad \Longleftrightarrow \quad \mathbb{P}(\sum_{n\geq 1} \prod_{k=1}^n (1 - \mathbf{1}_{A_k}\gamma_k) < +\infty) > 0.$$

Note that the condition $\sum_{n\geq 1} \prod_{k=1}^n (1-p_A\gamma_k) < +\infty$ implies $\mathbb{P}(\sum_{n\geq 1} \prod_{k=1}^n (1-\mathbf{1}_{A_k}\gamma_k) < +\infty) = 1$ and that if $\sum_n \gamma_n^2 < +\infty$, we have

$$\mathbb{P}_x(\sum_n (1 - X_n) < +\infty) > 0 \quad \Longleftrightarrow \quad \sum_{n\geq 1} e^{-p_A\Gamma_n} < +\infty.$$

The proof of Proposition 6 and of these comments is similar to that of the analogous statements concerning convergence to 0.

In the following Proposition, we give a sufficient condition for the fast rate to be achieved with probability one and a sufficient condition under which we have at most two rates with positive probability: $e^{-\pi\Gamma_n}$ and the fast rate $e^{-p_A\Gamma_n}$.

Proposition 7 *Let $\varepsilon_n = \frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n} - \pi$ for $n \geq 1$.*
(a) If $\sum_n \gamma_n\varepsilon_n^+ < +\infty$, we have $\sum_n (1 - X_n) < +\infty$ almost surely on the set $\{X_\infty = 1\}$.
(b) If $\liminf_n \varepsilon_n > 0$, then $\sum_n \gamma_n^2 e^{\pi\Gamma_n} < +\infty$, and, on the event $\{Y_\infty = 0\}$, we have $\sum_n (1 - X_n) < +\infty$ almost surely.

Note that the condition $\sum_n \gamma_n\varepsilon_n^+ < +\infty$ implies $\liminf_n \varepsilon_n^+ = 0$ and is satisfied in the following cases:

- the sequence (γ_n) is constant,
- $\gamma_n = \lambda n^{-\alpha}$ (for large enough n), with λ a positive constant and $0 < \alpha < 1$,
- $\gamma_n = C/(C + n)$, where the constant C satisfies $\pi C \geq 1$.

On the other hand, if $\gamma_n = C/(C + n)$, with $\pi C < 1$, we have $\liminf_n \varepsilon_n > 0$.

Before proving Proposition 7, we state and prove a lemma which will be useful for the proof of the second statement.

Lemma 1 *Assume that, for some positive integer n_0, $\forall n \geq n_0$, $\varepsilon_n \geq 0$. Then, the sequence $(Z_n)_{n\geq n_0}$, with $Z_n = (1-X_n)/\gamma_n$ is a submartingale, and we have $\sum_n (1-X_n) < +\infty$ a.s., on the set $\{X_\infty = 1\} \cap \{\sup_n \frac{1-X_n}{\gamma_{n+1}} < +\infty\}$.*

Remark 6 If $\inf_n \gamma_n e^{\pi\Gamma_n} > 0$, we have (on the event $\{X_\infty = 1\}$) $1 - X_n \leq Ce^{-\pi\Gamma_n}$ and $(1 - X_n)/\gamma_{n+1} \leq Ce^{-\pi\Gamma_{n+1}}/\gamma_{n+1}$. Then one can slightly relax the assumption in claim (b) since it follows from Lemma 1 that if $\varepsilon_n \geq 0$ for n large enough, $\sum_n (1 - X_n) < +\infty$ almost surely on $\{X_\infty = 1\}$.

Proof of Lemma 1: Starting from (10), we have

$$\begin{aligned}\frac{1-X_{n+1}}{\gamma_{n+1}} &= \frac{1-X_n}{\gamma_{n+1}} - \pi X_n(1-X_n) - \Delta M_{n+1}\\ &= (1-X_n)\left(\frac{1}{\gamma_n}+\varepsilon_n+\pi-\pi X_n\right)-\Delta M_{n+1}\\ &= \frac{1-X_n}{\gamma_n}\left(1+\varepsilon_n\gamma_n+\pi\gamma_n(1-X_n)\right)-\Delta M_{n+1}, \qquad (14)\end{aligned}$$

so that, for $n \geq n_0$, $Z_{n+1} \geq Z_n - \Delta M_{n+1}$, which proves that $(Z_n)_{n\geq n_0}$ is a submartingale. Now set $\tau_L := \min\{n \geq n_0 : 1-X_n > L\gamma_{n+1}\}$, $L > 0$. Then the stopped submartingale $(Z_n^{\tau_L})_{n\geq n_0}$ satisfies

$$(\Delta Z^{\tau_L}_{n+1})_+ \leq \mathbf{1}_{\{\tau_L \geq n+1\}}(\Delta Z_{n+1})_+ \leq L + \sup_n \|\Delta M_n\|_\infty.$$

Consequently the sub-martingale $(Z_n^{\tau_L})_{n\geq n_0}$ is bounded with bounded increments. Hence it converges ($\mathbb{P}_x$-$a.s.$ and in $L^1(\mathbb{P}_x)$) toward an integrable random variable ζ^L_∞. Furthermore (see [11]) the conditional variance increment process of its martingale part also converges to a finite random variable as $n \to +\infty$. This reads

$$\sum_{n=n_0+1}^{\tau_L} \mathbb{E}((\Delta M_n)^2 \,|\, \mathcal{F}_{n-1}) < +\infty \qquad \mathbb{P}_x\text{-}a.s..$$

But, we know from Proposition 1 that

$$\mathbb{E}((\Delta M_n)^2 \,|\, \mathcal{F}_{n-1}) \geq p_B X_{n-1}(1-X_{n-1}).$$

Consequently,

$$\{X_\infty = 1\} \cap \left(\cup_{p\in\mathbb{N}}\{\tau_p = +\infty\}\right) \subset \{\sum_n 1-X_n < +\infty\}.$$

We conclude by observing that $\cup_{p\in\mathbb{N}}\{\tau_p = +\infty\} = \{\sup_n \frac{1-X_n}{\gamma_{n+1}} < +\infty\}$. ◊

Proof of Proposition 7: We first assume that $\sum_n \gamma_n \varepsilon_n^+ < +\infty$. The proof is based, as in Lemma 1, on the study of the sequence $((1-X_n)/\gamma_n)$. We deduce from (14) that

$$\frac{1-X_{n+1}}{\gamma_{n+1}} \leq \frac{1-X_n}{\gamma_n}\left(1+\varepsilon_n^+\gamma_n+\pi\gamma_n(1-X_n)\right)-\Delta M_{n+1}. \qquad (15)$$

Hence

$$\mathbb{E}\left(\frac{1-X_{n+1}}{\gamma_{n+1}} \,|\, \mathcal{F}_n\right) \leq \frac{1-X_n}{\gamma_n}\left(1+\varepsilon_n^+\gamma_n+\pi\gamma_n(1-X_n)\right). \qquad (16)$$

We know from Proposition 4 that, on the set $\{X_\infty = 1\}$, we have $\sup_n (1-X_n)e^{\pi\Gamma_n} < +\infty$, so that $\gamma_n(1-X_n) \leq C\gamma_n e^{-\pi\Gamma_n} \leq C$ for some $C > 0$, and $\sum_n \gamma_n(1-X_n) < +\infty$. We now deduce from (16) and a supermartingale argument that, on $\{X_\infty = 1\}$, the sequence $((1-X_n)/\gamma_n)_{n\in\mathbb{N}}$ is almost surely convergent.

On the other hand, with the notation $Z_n = (1 - X_n)/\gamma_n$, we know from (15) that

$$\Delta M_{n+1} \le Z_n - Z_{n+1} + Z_n \left(\varepsilon_n^+ \gamma_n + \pi\gamma_n(1 - X_n)\right).$$

Therefore, on $\{X_\infty = 1\}$ the martingale M_n is bounded from above, and, since it has bounded jumps, we must have $<M>_\infty < +\infty$ almost surely. We know from Proposition 1 that $<M>_\infty \ge p_B \sum_n X_{n-1}(1 - X_{n-1})$. Hence $\sum_n (1 - X_n) < +\infty$ a.s. on $\{X_\infty = 1\}$.

We now assume that $\liminf \varepsilon_n > 0$, so that for n large enough (say $n \ge n_0$), we have

$$\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n} - \pi \ge \varepsilon, \tag{17}$$

for some $\varepsilon > 0$. In particular the sequence $(\gamma_n)_{n \ge n_0}$ is non-increasing and, for $n \ge n_0$,

$$\gamma_n - \gamma_{n+1} \ge (\pi + \varepsilon)\gamma_n \gamma_{n+1},$$

which implies $\sum_n \gamma_n^2 < +\infty$. We also have, for $n \ge n_0$,

$$\gamma_{n+1} \le \gamma_n (1 - (\pi + \varepsilon)\gamma_{n+1}) \le e^{-(\pi+\varepsilon)\gamma_{n+1}}.$$

Therefore, for $k \ge n \ge n_0$,

$$\gamma_k \le \gamma_n e^{-(\pi+\varepsilon)(\Gamma_k - \Gamma_n)},$$

and

$$\begin{aligned} \sum_{k \ge n} \gamma_k^2 e^{\pi \Gamma_k} &\le \sum_{k \ge n} \gamma_k \gamma_n e^{-(\pi+\varepsilon)(\Gamma_k - \Gamma_n)} e^{\pi \Gamma_k} \\ &= \gamma_n e^{(\pi+\varepsilon)\Gamma_n} \sum_{k \ge n} \gamma_k e^{-\varepsilon \Gamma_k} \\ &\le \gamma_n e^{(\pi+\varepsilon)\Gamma_n} \int_{\Gamma_{n-1}}^{\infty} e^{-\varepsilon x} dx \\ &\le \gamma_n \frac{e^{\varepsilon \gamma_n}}{\varepsilon} e^{\pi \Gamma_n}. \end{aligned}$$

We have thus proved not only that $\sum_n \gamma_n^2 e^{\pi \Gamma_n} < +\infty$, but also that

$$\sum_{k \ge n} \gamma_k^2 e^{\pi \Gamma_k} \le C \gamma_n e^{\pi \Gamma_n}$$

for some $C > 0$. It then follows from Proposition 5 that, on the set $\{Y_\infty = 0\}$, $(1 - X_n) \le C\theta_n \gamma_n e^{\pi \Gamma_n}$, and, using Remark 3, we get $\sup_n (1 - X_n)/\gamma_n < +\infty$ a.s. on $\{Y_\infty = 0\}$. We complete the proof by applying Lemma 1. ◇

Remark 7 Assume, with the notation of Proposition 7, that $\liminf \varepsilon_n^+ > 0$ and $\sum_n e^{-p_A \Gamma_n} < +\infty$. This is the case if $\gamma_n = C/(n + C)$, with $\pi C < 1 < p_A C$. Then, we deduce from Propositions 7 and 6 that $0 < \mathbb{P}(Y_\infty = 0) < 1$ and that, on $\{Y_\infty = 0\}$ the sequence $(1 - X_n)e^{p_A \Gamma_n}$ converges to a positive limit, whereas on $\{Y_\infty > 0\}$, $(1 - X_n)e^{\pi \Gamma_n}$ converges to a positive limit almost surely.

4 A parametric guide to the rates

In this section we will call *fast* a rate of the algorithm which induces that the error series converges *i.e.* $\sum_n 1 - X_n < +\infty$ when $X_n \to 1$ and $\sum_n X_n < +\infty$ when $X_n \to 0$. Other rates will be considered as *slow*.

Assume (at least for large enough n) that

$$\gamma_n = \left(\frac{C}{C' + n}\right)^\alpha, \quad \alpha \in (0,1], \quad C, C' > 0.$$

Then, the algorithm behaves as follows:

- If ($\alpha \in (0,1)$) or ($\alpha = 1$ & $Cp_B > 1$) then the algorithm is fallible with positive probability from any $x \in [0,1)$ (note that this probability is lower than 1 if $x \in (0,1)$). When failing, it always goes to 0 at a fast rate, (n^{-Cp_B} if $\alpha = 1$). This follows from Proposition 2.

- If $\alpha = 1$ and $C \leq \frac{1}{p_B}$, the algorithm is infallible from any $x \in (0,1]$. This follows from Proposition 3(b).

As concerns rates one has

- If $\alpha = 1$ and $C \geq \frac{1}{\pi}$ then the – fast – rate of convergence is n^{-Cp_A} on $\{X_n \to 1\}$. This follows from Proposition 7(a).

- If $\alpha = 1$ and $\frac{1}{p_A} < C < \frac{1}{\pi}$ then exactly *two rates of convergence occur with positive* $\mathbb{P}_x$*-probability* on $\{X_n \to 1\}$: a slow one – $n^{-C\pi}$ – and a fast one – n^{-Cp_A}. This follows from Proposition 6 and 7(b) (see remark 7).

- If $\alpha = 1$ and $C \leq \frac{1}{p_A}$ then (the algorithm is infallible from any $x \in (0,1]$) but only the slow rate of convergence survives *i.e.* $n^{-C\pi}$ on $\{X_n \to 1\}$. This follows from Proposition 6.

Note as corollaries that,

– when $2p_B \leq p_A$ (then $\frac{1}{\pi} \leq \frac{1}{p_B}$): it is possible to choose $C \in [\frac{1}{\pi}, \frac{1}{p_B}]$ so that the algorithm is simultaneously infallible and converging with a fast rate. This is possible because in some sense p_A and p_B are remote enough. The fastest achievable rate is $n^{-\frac{p_A}{p_B}}$ (with $C = \frac{1}{p_B}$). Of course such a specification is purely theoretical since p_A and p_B are supposed to be unknown.

– when $p_B < p_A < 2p_B$ (then $\frac{1}{p_B} < \frac{1}{\pi}$): there is no access to fast converging rates within infallibility, because p_A and p_B are too close to each other .

– in any case, when no information is available on the parameters p_A and p_B, the "blind" choice $C = 1 \leq \frac{1}{p_A}$ which ensures infallibility induces a slow rate of convergence, namely $n^{-\pi}$. In fact this rate can be very poor when p_A and p_B get close to each other.

At this point the conclusion can be the following: the higher the parameter C is, the faster the algorithm goes. But if C is too high, it may go wrong.

– One further point to be noticed is that what we called the slow rate – $e^{-\pi\,\Gamma_n}$ – for the algorithm is but *the rate of its mean deterministic version* (see [6] for details). So, even when it is infallible (that is converges to the same limit as its mean version), it always converges at least as fast as this deterministic procedure (which is of no practical interest since its implementation would require p_A and p_B to be known). When no information is available on the parameters p_A and p_B, this is the rate which is actually obtained.

As a conclusion, the convergence rate behaviour of this stochastic approximation algorithm is completely non-standard. Thus, from a mathematical viewpoint, one last feature to be noticed is the unusual "spectrum" of the rates since the switching from one rate to another takes place "progressively" with a range of values of the parameter C for the gain parameter for which two different rates are achieved with positive probability.

References

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

[2] O. Brandière, M. Duflo (1996), Les algorithmes stochastiques contournent-ils les pièges ?, *Ann. Inst. H. Poincaré Probab. Statist.*, **32**, pp. 395-427.

[3] J.C. Fort, G. Pagès (2002), Decreasing step stochastic algorithms: *a.s.* behaviour of weighted empirical measures, *Monte Carlo Methods and Applications*, **8**, $n^0 3$, pp. 221-320.

[4] P. Hall, C.C Heyde (1980), *Martingale Limit Theory and its Application*, Academic Press, New York, 1980, 308p.

[5] 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.

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

[7] D. Lamberton, G. Pagès (2005), A penalized bandit algorithm, pre-print LPMA 1019, Univ. Paris 6, and pre-print Univ. Marne-la-Vallée (France).

[8] V.A. Lazarev (1992), Convergence of stochastic approximation procedures in the case of a regression equation with several roots, (transl. from) *Problemy Pederachi Informatsii*, vol. **28**(1).

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

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

[11] J. Neveu (1975), *Discrete-parameter martingales*, North-Holland, New York.

[12] M. Pelletier (1998), Weak convergence rates for stochastic approximation with application to multiple targets and simulated annealing, *Annals of Applied Probability*, **8**(1) pp. 10-44.

[13] R. Pemantle (1990), Nonconvergence to unstable points in urn models and stochastic approximations, *Annals of Probability*, **18**(2), pp. 698-712.

[14] P. Tarrès, Pièges répulsifs, *C.R.A.S. Acad. Sc. de Paris*, Série I, **330**, pp.125-130, 2000.