The Annals of Probability
2012, Vol. 40, No. 5, 1980–2007
DOI: 10.1214/11-AOP666

SUFFICIENT CONDITIONS OF STANDARDNESS FOR FILTRATIONS OF STATIONARY PROCESSES TAKING VALUES IN A FINITE SPACE

BY GAËL CEILLIER

IUT Robert Schumann

Let X be a stationary process with finite state-space A. Bressaud et al. [*Ann. Probab.* **34** (2006) 1589–1600] recently provided a sufficient condition for the natural filtration of X to be standard when A has size 2. Their condition involves the conditional laws $p(\cdot|x)$ of X_0 conditionally on the whole past $(X_k)_{k\leq -1} = x$ and controls the strength of the influence of the "old" past of the process on its present X_0. It involves the maximal gaps between $p(\cdot|x)$ and $p(\cdot|y)$ for infinite sequences x and y which coincide on their n last terms. In this paper, we first show that a slightly stronger result holds for any finite state-space. Then, we provide sufficient conditions for standardness based on average gaps instead of maximal gaps.

1. Introduction.

1.1. *Setting.* In this paper we study stationary processes $X = (X_n)_{n\in\mathbb{Z}}$ indexed by the integer line $\mathbb{Z}$ and with values in a finite set A. We assume that X is defined recursively as follows: for every $n \in \mathbb{Z}$, X_n is a function of the "past" $X_{n-1}^{\triangleleft} = (X_k)_{k\leq n-1}$ of X and of a "fresh" random variable U_n, which brings in some "new" randomness. In particular the process $U = (U_n)_{n\in\mathbb{Z}}$ is independent. To be more specific, we introduce some notations and definitions about σ-algebras.

All σ-fields are assumed to be complete. For every process $\xi = (\xi_n)_{n\in\mathbb{Z}}$ and every $n \in \mathbb{Z}$, let $\xi_n^{\triangleleft} = (\xi_k)_{k\leq n}$ and $\mathcal{F}_n^{\xi} = \sigma(\xi_n^{\triangleleft})$. The natural filtration of ξ is the nondecreasing sequence $\mathcal{F}^{\xi} = (\mathcal{F}_n^{\xi})_{n\in\mathbb{Z}}$. Furthermore, $\mathcal{F}_{\infty}^{\xi} = \sigma(\xi_k; k \in \mathbb{Z})$ and $\mathcal{F}_{-\infty}^{\xi}$ is the tail σ-algebra $\mathcal{F}_{-\infty}^{\xi} = \bigcap_{k\in\mathbb{Z}} \mathcal{F}_k^{\xi}$.

We say that a process U is a *governing process* for X, or that U *governs* X if, for every $n \in \mathbb{Z}$, (i) U_{n+1} is independent of $\mathcal{F}_n^{X,U}$, and (ii) X_{n+1} is measurable with respect to $\sigma(U_{n+1}) \vee \mathcal{F}_n^X$. In particular any governing process is independent. If moreover the U_n are uniform on $[0, 1]$, the process (U, X) is—according to Schachermayer's definition [9] and up to a time reversal—a *parametrization* of the process X.

Likewise, we say that a process U is a *generating process* for X, or that U *generates* X if, for every $n \in \mathbb{Z}$, X_n is measurable with respect to $\mathcal{F}_n^U$. This is

Received April 2010; revised March 2011.
MSC2010 subject classifications. 60A99.
Key words and phrases. Filtration, parametrization, non-Markovian processes.

equivalent to the condition that $\mathcal{F}_n^X \subset \mathcal{F}_n^U$ for every $n \in \mathbb{Z}$, a property which, from now on, we write as $\mathcal{F}^X \subset \mathcal{F}^U$.

One could be led to believe that when $\mathcal{F}_{-\infty}^X$ is trivial, any process governing X generates X as well. But, although notoriously used by Wiener and Kallianpur in [6] (not published, but see a discussion in [8]), this argument is false. As a simple counterexample, assume that X is i.i.d., that every X_n is uniform on $\{-1, 1\}$ and set $U_n = X_n X_{n-1}$ for every $n \in \mathbb{Z}$. Then $\mathcal{F}_{-\infty}^X$ is trivial, and U governs X, but X_0 is independent of $\mathcal{F}_\infty^U$; hence U does not generate X.

Governing and generating processes are related to *immersions* of filtrations. Recall that the filtration $\mathcal{F}^X$ is *immersed* in the filtration $\mathcal{F}^U$ if $\mathcal{F}^X \subset \mathcal{F}^U$ and if, for every $n \in \mathbb{Z}$, $\mathcal{F}_{n+1}^X$ and $\mathcal{F}_n^U$ are independent conditionally on $\mathcal{F}_n^X$. Roughly speaking, this means that $\mathcal{F}_n^U$ gives no further information on X_{n+1} than $\mathcal{F}_n^X$ does. Equivalently, $\mathcal{F}^X$ is immersed in $\mathcal{F}^U$ if every $\mathcal{F}^X$-martingale is an $\mathcal{F}^U$-martingale. The following easy fact holds (see a proof in Section 5.2).

LEMMA 1.1. *If U is a governing and generating process for X, then $\mathcal{F}^X$ is immersed in $\mathcal{F}^U$.*

Another notable property of filtrations is *standardness*. Recall that $\mathcal{F}^X$ is *standard* if, modulo an enlargement of the probability space, one can immerse $\mathcal{F}^X$ in a filtration generated by an i.i.d. process. Vershik introduced standardness in the context of ergodic theory. Examples of nonstandard filtrations include the filtrations of $[T, T^{-1}]$ transformations, introduced in [5]. Split-word processes, inspired by Vershik's (r_n)-adic sequences of decreasing partitions [11] and studied in [10] and [7], for instance, also provide nonstandard filtrations.

Obviously, Lemma 1.1 above implies that if X has a generating and governing process, then $\mathcal{F}^X$ is standard. Whether the converse holds is not known.

Necessary and sufficient conditions for standardness include Vershik's self-joining criterion and Tsirelson's notion of I-cosiness. Both notions are discussed in [3] and are based on conditions which are subtle and not easy to use nor to check in specific cases.

Our goal in this paper is to provide sufficient conditions of standardness that are easier to use than the ones mentioned above. Each of our conditions involves a measure of the influence of the "old" past of the process on its present. We introduce them in the next section.

1.2. *Statement of the results.* We now introduce some measures of the influence of the past of a process on its present. To conveniently state these definitions and, later on, our results, we first introduce some notations.

Recall that X is a stationary process indexed by the integer line $\mathbb{Z}$ with values in some finite set A and with natural filtration $\mathcal{F}^X$.

NOTATION 1. (1) Slabs: For any sequence $(\xi_n)_{n\in\mathbb{Z}}$ in $A^{\mathbb{Z}}$, deterministic or random, and any integers $i \le j$, $\xi_{i:j}$ is the $(j-i+1)$-uple $(\xi_n)_{i\le n\le j}$ in A^{j-i+1}.

(2) Shifts: If $k-i=\ell-j$, $\xi_{i:k}=\zeta_{j:\ell}$ means that $\xi_{i+n}=\zeta_{j+n}$ for every integer n such that $0\le n\le k-i$.

Infinite case: Let $A^{\lhd}$ denote the space of sequences $(\xi_n)_{n\le -1}$. For every i in $\mathbb{Z}$, a sequence $(\xi_n)_{n\le i}$ is also considered as an element of $A^{\lhd}$ since, similarly to the finite case, one identifies $\xi_i^{\lhd}=(\xi_n)_{n\le i}$ and $\zeta_j^{\lhd}=(\zeta_n)_{n\le j}$ if $\xi_{i+n}=\zeta_{j+n}$ for every integer $n\le 0$.

(3) Concatenation: For all $i\ge 0$, $j\ge 0$, $x=(x_n)_{1\le n\le i}$ in A^i and $y=(y_n)_{1\le n\le j}$ in A^j, xy denotes the concatenation of x and y, defined as

$$xy=(x_1,\dots,x_i,y_1,\dots,y_j),\qquad xy\in A^{i+j}.$$

Infinite case: $i\ge 0$, $y=(y_n)_{1\le n\le i}$ in A^i and $x=(x_n)_{n\le -1}$ in $A^{\lhd}$, xy denotes the concatenation of x and y, defined as

$$xy=(\dots,x_{-2},x_{-1},y_1,\dots,y_i),\qquad xy\in A^{\lhd}.$$

NOTATION 2. For each $n\ge 0$, $x\in A^n$ and $a\in A$, set

$$p(a|x)=\mathbb{P}(X_0=a|X_{-n:-1}=x)$$

with the convention

$$p(a|x)=\mathbb{P}(X_0=a)\qquad \text{if } \mathbb{P}[X_{-n:-1}=x]=0.$$

In the following,

$$p(\cdot|x)=\mathbb{P}(X_0=\cdot|X_{-1}^{\lhd}=x),\qquad x\in A^{\lhd},$$

denotes a regular version of the conditional law of X_0 given $X_{-1}^{\lhd}$.

We now introduce three quantities γ_n, α_n and δ_n measuring the pointwise influence at distance n.

DEFINITION 1. For every $n\ge 0$, let

$$\gamma_n=1-\inf\left\{\frac{p(a|xz)}{p(a|yz)};a\in A,x\in A^{\lhd},y\in A^{\lhd},z\in A^n,p(a|yz)>0\right\},$$

$$\alpha_n=1-\inf_{z\in A^n}\sum_{a\in A}\inf\{p(a|yz);y\in A^{\lhd}\},$$

$$\delta_n=\sup\{\|p(\cdot|xz)-p(\cdot|yz)\|;x\in A^{\lhd},y\in A^{\lhd},z\in A^n\},$$

where, for all probabilities μ and ν on A, $\|\mu-\nu\|$ is the distance in total variation between μ and ν, defined as

$$\|\mu-\nu\|=\frac{1}{2}\sum_{a\in A}|\mu(a)-\nu(a)|=\sum_{a\in A}[\mu(a)-\nu(a)]_+.$$

Note that the definitions of γ_n, α_n and δ_n depend on the choice of the regular version $(p(\cdot|x))_{x\in A^{\lhd}}$ of the conditional law of X_0 given $X_{-1}^{\lhd}$. One needs a "good" version to get small influences for applying the theorems below.

The sequences $(\gamma_n)_{n\geq 0}$, $(\alpha_n)_{n\geq 0}$ and $(\delta_n)_{n\geq 0}$ are nonincreasing, $[0,1]$-valued and $\delta_n \leq \gamma_n$, $\delta_n \leq \alpha_n$ for every $n \geq 0$; see the proof in Section 5.1.

For every $[0,1]$-valued sequence $(\varepsilon_n)_{n\geq 0}$, we consider the condition

$$(\mathcal{H}(\varepsilon)) \qquad \sum_{k=0}^{+\infty}\prod_{n=0}^{k}(1-\varepsilon_n) = +\infty.$$

For instance, $\mathcal{H}(\gamma)$ and $\mathcal{H}(2\delta)$ are, respectively,

$$\sum_{k=0}^{+\infty}\prod_{n=0}^{k}(1-\gamma_n) = +\infty \quad \text{and} \quad \sum_{k=0}^{+\infty}\prod_{n=0}^{k}(1-2\delta_n) = +\infty.$$

Observe that if two $[0,1]$-valued sequences $(\varepsilon_n)_{n\geq 0}$ and $(\zeta_n)_{n\geq 0}$ are such that $\varepsilon_n \leq \zeta_n$ for every $n \geq 0$, then $\mathcal{H}(\zeta)$ implies $\mathcal{H}(\varepsilon)$. Hence condition $(\mathcal{H}(\varepsilon))$ asserts that $(\varepsilon_n)_{n\geq 0}$ is "small enough" in a way.

The definition of $(\gamma_n)_{n\geq 0}$ and the assumption $\mathcal{H}(\gamma)$ are both stated in [1]. The main result of [1] is the following.

THEOREM 1 (Bressaud et al. [1]). *Assume that the size of A is 2, then $\mathcal{H}(\gamma)$ implies that $\mathcal{F}^X$ is standard.*

The scope of Theorem 1 is restricted by the following three conditions. First, the size of A must be 2. Second, one must control the *ratios* of probabilities which define γ_n. Third, $\mathcal{H}(\gamma)$ implies that $\gamma_0 < 1$; therefore one can show that $\mathcal{H}(\gamma)$ implies the existence of $c > 0$ such that $p(a|x) \geq c$ for every x in $A^{\lhd}$ and a in A such that $\mathbb{P}[X_0 = a] > 0$; see the proof in Section 5.4.

Our first result allows us to get rid of the first two restrictions.

THEOREM 2. (1) *Assume that A is finite, that $2\delta_0 < 1$ and that $\mathcal{H}(2\delta)$ holds. Then $\mathcal{F}^X$ is standard.*

(2) *If the size of A is 2, $\mathcal{H}(\delta)$ alone implies that $\mathcal{F}^X$ is standard.*

Theorem 2 generalizes and improves on Theorem 1 of [1], since $\delta_n \leq \gamma_n$ for every n. Note that the straight adaptation of the proof of [1] to sizes of A at least 3 leads to the more stringent condition $\mathcal{H}(2\gamma)$.

Another measure of influence, based on the quantities α_n defined before, is introduced and used in [2] (actually the notation there is $a_n = 1 - \alpha_n$). The authors show that if $\mathcal{H}(\alpha)$ holds, there exists a perfect sampling algorithm for the process X, a result which implies that $\mathcal{F}^X$ is standard. But since $\delta_n \leq \alpha_n$ for every $n \geq 0$, the result of [2] does not imply Theorem 1.

Theorems 1 and 2 and the exact sampling algorithm of [2] all require an upper bound of some pointwise influence sequence. Our next result uses a less restrictive hypothesis based on some average influences η_n, defined below.

DEFINITION 2. For every $n \geq 0$, let η_n denote the average influence at distance n, defined as

$$\eta_n = \sum_{z \in A^n} \mathbb{E}\big[\big\| p(\cdot|z) - p(\cdot|X_{-n-1}^{\triangleleft} z)\big\|\big] \cdot \mathbb{P}[X_{-n:-1} = z],$$

and call $\mathcal{H}'(\eta)$ the condition

$$(\mathcal{H}'(\eta)) \qquad \sum_{k=0}^{+\infty} \eta_k < +\infty.$$

Note that η_n is also

$$\eta_n = \mathbb{E}\big[\big\| p(\cdot|Y_{-n:-1}) - p(\cdot|X_{-n-1}^{\triangleleft} Y_{-n:-1})\big\|\big],$$

where Y is an independent copy of X.

DEFINITION 3 (Priming condition). We say that the process X fulfills the priming condition if for every a in A, $p(a|X_{-1}^{\triangleleft}) > 0$ almost surely.

THEOREM 3. *Assume that A is finite and that X fulfills the priming condition. Then, $\mathcal{H}'(\eta)$ implies that $\mathcal{F}^X$ is standard.*

The sequence $(\eta_n)_{n\geq 0}$ is $[0, 1]$-valued. If $\eta_n < 1$ for every $n \leq 0$, then $\mathcal{H}'(\eta)$ clearly implies $\mathcal{H}(\eta)$. Yet, since $\eta_n \leq \delta_n$ for every $n \geq 0$ (see the proof in Section 5.1), the condition $\mathcal{H}'(\eta)$ cannot be compared to the conditions $\mathcal{H}(\delta)$ and $\mathcal{H}(2\delta)$.

Theorem 3 gives a remarkable result for chains with memory of variable length. These chains, studied notably in [4] and widely used for mathematical models, are stationary processes X taking values in a finite alphabet A, such that the distribution of X_0 given the past $X_{-1}^{\triangleleft}$ depends only on a past $X_{-\ell:-1}$ of length ℓ, where ℓ is random and measurable with respect to $\mathcal{F}_{-1}^X$.

More precisely, for $x \in A^{\triangleleft}$, let

$$\begin{aligned}\ell(x) &= \inf\{n \geq 0; y \mapsto p(\cdot|yx_{-n:-1}) \text{ is constant on } A^{\triangleleft}\}\\ &= \inf\{n \geq 0; \forall y \in A^{\triangleleft}, p(\cdot|yx_{-n:-1}) = p(\cdot|x)\}.\end{aligned}$$

Then X is a variable length Markov chain if $\ell(X_{-1}^{\triangleleft})$ is almost surely finite. The following result holds.

COROLLARY 1.2. *If X fulfills the priming condition and if $\ell(X_{-1}^{\triangleleft})$ is integrable, then the natural filtration $\mathcal{F}^X$ is standard.*

Once again we refer the reader to Section 5.3 for the proof.

Here is a plan of the rest of the paper. In Section 2, we prove Theorem 2. In Section 3, we prove Theorem 3. In Section 4, we compare Theorems 2 and 3 through examples. Finally in Section 5, we prove some facts stated without proof in the Introduction, namely Lemma 1.1, Corollary 1.2, a consequence of the assumption $\mathcal{H}(\gamma)$ and some inequalities involving the quantities α_n, γ_n, δ_n and η_n.

2. Pointwise influence.

2.1. *Construction of a governing sequence.* We construct a governing sequence with values in the standard simplex on $\#A$ vertices.

NOTATION 3. Let H be the hyperplane in $\mathbb{R}^A$ defined by

$$H = \Big\{x = (x_a)_{a\in A} \in \mathbb{R}^A : \sum_{a\in A} x_a = 1\Big\}.$$

Let S be the simplex in H defined by

$$S := (\mathbb{R}_+)^A \cap H = \Big\{x \in (\mathbb{R}_+)^A : \sum_{a\in A} x_a = 1\Big\}.$$

In other words, $S = \mathrm{Conv}(E_A)$ is the convex enveloppe of the canonical basis $E_A = (E_a)_{a\in A}$ of $\mathbb{R}^A$.

Let λ denote the Lebesgue measure on H and $\mu = (\mathbf{1}_S/\lambda(S))\lambda$ the uniform distribution on S.

NOTATION 4. For any probability p on A, let

$$G(p) = (p(a))_{a\in A} = \sum_{a\in A} p(a)E_a, \qquad G(p) \in S.$$

For a in A, denote by $f_a(\cdot, p)$ the affine map from H to H which sends E_a on $G(p)$ and lets invariant E_b for every b in A, $b \neq a$. Let

$$S_a(p) = f_a(S, p) = \mathrm{Conv}\big(\{G(p)\} \cup E_A \setminus \{E_a\}\big).$$

A short computation yields the interpretation of $p(a)$ below.

LEMMA 2.1. *For any a in A, $\det(f_a(\cdot, p)) = p(a)$. Therefore, for any measurable $B \subset S$,*

$$\lambda[f_a(B, p)] = \lambda[B]p(a).$$

In particular $\lambda(S_a(p)) = \lambda(S)p(a)$, hence $\mu(S_a(p)) = p(a)$.

We now characterize $S_a(p)$.

By convention, for every $r > 0$, we set $r/0 = \infty$, and $0/0 = 0$.

LEMMA 2.2. *Let p be a probability on A, a in A, then*

$$S_a(p) = \left\{x = (x_a)_{a\in A} \in S : \frac{x_a}{p(a)} = \min_{b\in A} \frac{x_b}{p(b)}\right\}$$
$$= \left\{x \in S : \forall bb \in A, \frac{x_a}{p(a)} \le b \in A \frac{x_b}{p(b)}\right\}.$$

COROLLARY 2.3. *S is the union of the simplices $S_a(p)$, with a in A and that, if $a \neq b$, the simplices $S_a(p)$ and $S_b(p)$ meet only at their boundary.*

Proof: It's a straight corollary of the Lemmas 2.1 and 2.2.

PROOF. Call $\Sigma_a(p)$ the right-hand side. Since $\Sigma_a(p)$ is a convex polyedron and contains the points $G(p)$ and E_b for every $b \neq a$, $S_a(p) \subset \Sigma_a(p)$.

As regards the other inclusion, let $x = (x_a)_{a\in A}$ in $\Sigma_a(p)$. Then $x_a/p(a)$ is finite, and

$$x = \frac{x_a}{p(a)} G(p) + \sum_{b\neq a}\left(x_b - p(b)\frac{x_a}{p(a)}\right) E_b.$$

From the definition of $\Sigma_a(p)$, $x_b/p(b) \ge x_a/p(a)$ for every $b \neq a$, hence one has $x_b - p(b)x_a/p(a) \ge 0$ for every $b \neq a$. Furthermore,

$$\frac{x_a}{p(a)} + \sum_{b\neq a}\left(x_b - p(b)\frac{x_a}{p(a)}\right) = \frac{x_a}{p(a)} + \sum_{b\in A}\left(x_b - p(b)\frac{x_a}{p(a)}\right) = \sum_{b\in A} x_b = 1,$$

hence x is indeed a barycenter of the points $G(p)$ and E_b for $b \neq a$. This concludes the proof. □

One knows that the simplices $(S_a(p))_{a\in A}$ cover S and intersect only on a set of measure zero. Hence, for almost every s in S, there exists a unique a in A such that $s \in S_a(p)$. Our next definition deals with the tie cases.

DEFINITION 4. Fix once and for all a total ordering of A. For every s in S and every probability p on A with full support, define

$$g(s, p) = \min\{a \in A : s \in S_a(p)\}.$$

LEMMA 2.4. *Let U denote a random variable uniformly distributed on S. Then the distribution of $g(U, p)$ is p.*

Indeed, up to negligible events, $\{g(U, p) = a\} = \{U \in S_a(p)\}$, hence

$$\mathbb{P}[g(U, p) = a] = \mu(S_a(p)) = p(a).$$

The following lemma is our main tool to construct governing sequences.

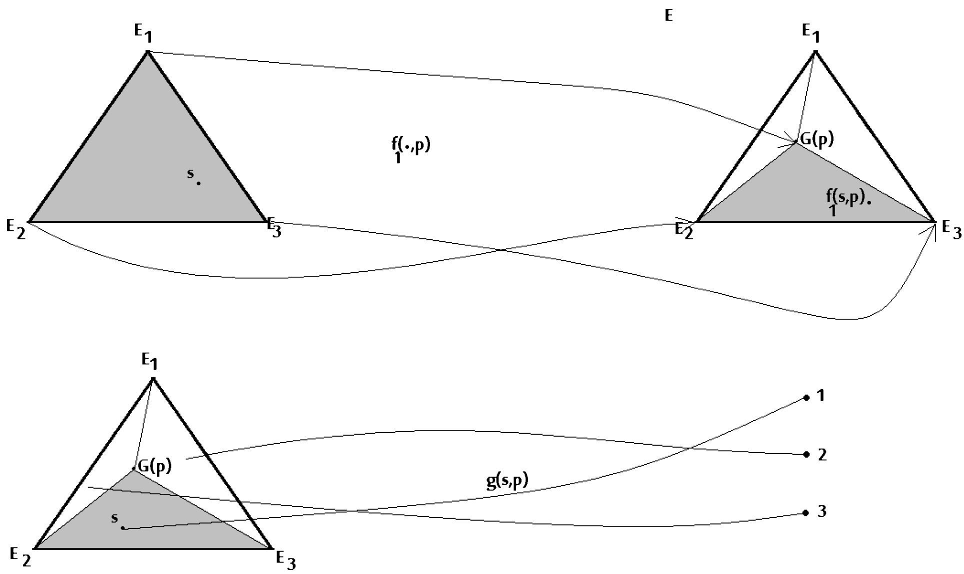

FIG. 1. $f(\cdot, p)$ *and* $g(\cdot, p)$.

LEMMA 2.5. *Let X be a random variable with distribution p on A. Let W be a random variable with uniform distribution on S and independent of X. Introduce*

$$U = f_X(W, p) = \sum_{a \in A} f_a(W, p) \mathbf{1}_{\{X=a\}}.$$

Then U is uniformly distributed on S and $X = g(U, p)$ almost surely.

PROOF. Since $U \in f_X(S, p) = S_X(p)$, $X = g(U, p)$ almost surely. We now prove that U is uniformly distributed on S.

The sets $S_a(p)$ for a in A cover S and their pairwise intersections are negligible for λ. Hence, for every Borel subset B of S,

$$\begin{aligned}\mathbb{P}[U \in B] &= \sum_{a \in A} \mathbb{P}[X = a; f_a(W, p) \in B] = \sum_{a \in A} \mathbb{P}[X = a] \cdot \mathbb{P}[W \in f_a(\cdot, p)^{-1}(B)] \\ &= \sum_{a \in A} p(a) \frac{\lambda(f_a(\cdot, p)^{-1}(B) \cap S)}{\lambda(S)} = \sum_{a \in A} \frac{\lambda(B \cap S_a)}{\lambda(S)} = \frac{\lambda(B)}{\lambda(S)} = \mu(B),\end{aligned}$$

where the second equality stems from the independence of X and W, and the fourth equality stems from Lemma 2.1. This concludes the proof. □

2.2. *Upper bound of the error.* In this section we study the dependence of the random variable $g(U, p)$ with respect to p. The following result will be used twice.

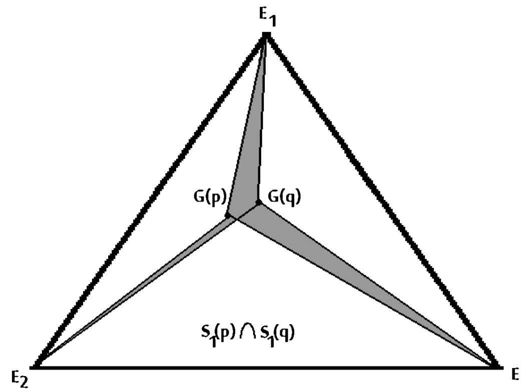

FIG. 2. *Computation of* $\mathbb{P}[g(U,p) \neq g(U,q)]$. *The gray area shows the* $s \in S$ *such that* $g(s,p) \neq g(s,q)$.

PROPOSITION 2.6 (Upper bound of the error). *Let U be a random variable uniformly distributed on S. Let p and q be two probabilities on A. Then,*

$$\mathbb{P}[g(U,p) \neq g(U,q)] \leq 2\|p-q\|.$$

In the special case $\#A = 2$,

$$\mathbb{P}[g(U,p) \neq g(U,q)] = \|p-q\|.$$

REMARK 1. The better result when $\#A = 2$ is the reason why Theorem 2 involves weaker hypotheses on $(\delta_n)_n$ in this case.

PROOF OF PROPOSITION 2.6. Assume without loss of generality that U is constructed from i.i.d. random variables $(\varepsilon_a)_{a \in A}$ exponentially distributed with parameter 1, as follows. For every a in A,

$$U_a = \frac{\varepsilon_a}{\sum_{b \in A} \varepsilon_b}.$$

The event $\{g(U,p) \neq g(U,q)\}$ depends on $(\varepsilon_a)_a$, as follows. By definition of g, up to negligible events,

$$\{g(U,p) = g(U,q)\} = \bigcup_{a \in A} C_a \qquad \text{with } C_a = \{U \in S_a(p) \cap S_a(q)\}.$$

Furthermore, since for every $a \in A$, $\mathbb{P}[C_a] = 0$ if $p(a) = 0$ or $q(a) = 0$, and since $\mu(C_a \cap C_b) = 0$ for $a \neq b$, one gets

$$\mathbb{P}[g(U,p) = g(U,q)] = \sum_{a \in A} \mathbb{P}[C_a] \mathbf{1}_{\{p(a)>0, q(a)>0\}}.$$

For every $a \in A$ such that $p(a) > 0$ and $q(a) > 0$, Lemma 2.2 gives

$$\begin{aligned} C_a &= \left\{ \frac{\varepsilon_a}{p(a)} = \min_b \frac{\varepsilon_b}{p(b)}; \frac{\varepsilon_a}{q(a)} = \min_b \frac{\varepsilon_b}{q(b)} \right\} \\ &= \left\{ \varepsilon_a \le \min_b \left(p(a) \frac{\varepsilon_b}{p(b)}, q(a) \frac{\varepsilon_b}{q(b)} \right) \right\}, \end{aligned}$$

hence

$$C_a = \bigcap_{b \neq a} \{\varepsilon_b \ge \lambda_{b/a} \varepsilon_a\}, \qquad \lambda_{b/a} = \max\left(\frac{p(b)}{p(a)}, \frac{q(b)}{q(a)} \right).$$

Conditioning on ε_a and using that the random variables $(\varepsilon_b)_{b \neq a}$ are i.i.d., exponentially distributed and independent of ε_a, one gets

$$\mathbb{P}[C_a | \varepsilon_a] = \mathbb{P}\left[\bigcap_{b \neq a} \{\varepsilon_b \ge \lambda_{b/a} \varepsilon_a\} \middle| \varepsilon_a \right] = \prod_{b \neq a} \exp(-\lambda_{b/a} \varepsilon_a),$$

hence

$$\mathbb{P}[C_a] = \mathbb{E}\left(\exp\left(-\left(\sum_{b \neq a} \lambda_{b/a} \right) \varepsilon_a \right) \right) = \frac{1}{1 + \sum_{b \neq a} \lambda_{b/a}}.$$

Therefore

$$\mathbb{P}[g(U, p) = g(U, q)] = \sum_{a \in A} \mathbb{P}[C_a] \mathbf{1}_{\{p(a) > 0, q(a) > 0\}} = \sum_{a \in A} \frac{\mathbf{1}_{\{p(a) > 0, q(a) > 0\}}}{1 + \sum_{b \neq a} \lambda_{b/a}}.$$

This last expression is not so easy to compute because each $\lambda_{b/a}$ is defined as a maximum. However,

$$\sum_{a \in A} \frac{\mathbf{1}_{\{p(a) > 0\}}}{1 + \sum_{b \neq a} p(b)/p(a)} = \sum_{a \in A} \frac{p(a) \mathbf{1}_{\{p(a) > 0\}}}{p(a) + \sum_{b \neq a} p(b)} = \sum_a p(a) = 1.$$

Subtracting the expression for $\mathbb{P}[g(U, p) = g(U, q)]$ to this, one gets

$$\mathbb{P}[g(U, p) \neq g(U, q)] = \sum_{a \in A} \frac{\mathbf{1}_{\{p(a) > 0\}}}{1 + \sum_{b \neq a} p(b)/p(a)} - \sum_{a \in A} \frac{\mathbf{1}_{\{p(a) > 0, q(a) > 0\}}}{1 + \sum_{b \neq a} \lambda_{b/a}}.$$

Coming back to the definition of $\lambda_{b/a}$ and using simple algebraic manipulations, one gets for any $a \in A$ such that $p(a) > 0$ and $q(a) > 0$,

$$\begin{aligned} &\frac{1}{1 + \sum_{b \neq a} p(b)/p(a)} - \frac{1}{1 + \sum_{b \neq a} \lambda_{b/a}} \\ &\quad = \frac{\sum_{b \neq a} (\lambda_{b/a} - p(b)/p(a))}{(1 + \sum_{b \neq a} p(b)/p(a))(1 + \sum_{b \neq a} \lambda_{b/a})} \\ &\quad = p(a) \frac{r(a)}{q(a) + r(a)}, \end{aligned}$$

where

$$r(a) = \sum_b [q(b)p(a) - p(b)q(a)]_+.$$

Furthermore, for any $a \in A$ such that $p(a) > 0$ and $q(a) = 0$, one gets

$$\frac{1}{1 + \sum_{b \neq a} p(b)/p(a)} = p(a) = p(a)\frac{r(a)}{q(a) + r(a)}.$$

Summing on every a, one gets finally

$$\mathbb{P}[g(U, p) \neq g(U, q)] = \sum_{a \in A} p(a)\frac{r(a)}{q(a) + r(a)}.$$

If $A = \{a, a'\}$ and, for example, $q(a) < p(a)$, then $r(a) = p(a) - q(a)$ and $r(a') = 0$, hence

$$\mathbb{P}[g(U, p) \neq g(U, q)] = r(a) = p(a) - q(a) = \|p - q\|.$$

In the general case, note that

$$q(a) + r(a) \geq q(a) + \sum_b \bigl(q(b)p(a) - p(b)q(a)\bigr) = p(a),$$

hence

$$\begin{aligned}\mathbb{P}[g(U, p) \neq g(U, q)] &\leq \sum_{a \in A} r(a) = \sum_{a \in A} \sum_{b \neq a} [p(a)q(b) - q(a)p(b)]_+ \\ &\leq \sum_{a \in A} \sum_{b \neq a} p(a)\bigl[[q(b) - p(b)]_+ + p(b)[p(a) - q(a)]_+\bigr],\end{aligned}$$

where the last inequality stems from the fact that $(u + v)_+ \leq (u)_+ + (v)_+$ for every u and v. Finally, the last double sum is at most $2\|p - q\|$, which ends the proof in the general case. □

Recall that if p and q are two *fixed* probabilities on A, then for every random variables Z_p and Z_q with laws p and q defined on the same probability space,

$$\mathbb{P}[Z_p \neq Z_q] \geq \|p - q\|.$$

Conversely, a standard construction in coupling theory provides random variables Z_p and Z_q with laws p and q such that $\mathbb{P}[Z_p \neq Z_q] = \|p - q\|$.

The interest of Proposition 2.6 is to provide a global coupling of *all* probabilities on A. One can wonder whether the constant 2 in this proposition can be improved. Our next result (not used in the sequel) shows that the constant 2 is optimal for the coupling $(g(U, p))_p$, and that it is not possible to do much better with any other global coupling.

PROPOSITION 2.7 (Optimality of the upper bound of the error). *If* $\#A \geq 3$, *the constant* 2 *in the inequality* $\mathbb{P}[g(U,p) \neq g(U,q)] \leq 2\|p-q\|$ *of Proposition* 2.6 *is optimal.*

Furthermore, if $(Z_p)_p$ *is a family of random variables indexed by probabilities on* A, *where each* Z_p *follows the law* p, *then there exist two probabilities* $p \neq q$ *such that*

$$\mathbb{P}[Z_p \neq Z_q] \geq 2(1 - 1/\#A)\|p-q\|.$$

PROOF. The first part of the proposition follows from the explicit example where $\{a,b,c\} \subset A$, $p(a) = q(a) = 1 - \varepsilon$ and $p(b) = q(c) = \varepsilon$ in the limit $\varepsilon \to 0$.

With regard to the second part, let $N = \#A$ and for every a in A, let Z^a denote the random variable of $(Z_p)_p$ with uniform distribution on $A \setminus \{a\}$. Choose $a_0 \in A$, and consider the random set D of the elements a of A such that $Z^a = Z^{a_0}$. For every $a, b \in A$,

$$\mathbf{1}_{[Z^a \neq Z^b]} \geq \mathbf{1}_{[a \in D, b \notin D]} + \mathbf{1}_{[a \notin D, b \in D]}.$$

By summing over $a, b \in A$ and by taking expectations, one gets

$$\sum_{a,b} \mathbb{P}[Z^a \neq Z^b] \geq \mathbb{E}[2\#D(N - \#D)].$$

Of course $a_0 \in D$, and $Z^{a_0} \notin D$, since $Z^a \neq a$ almost surely for every a. Thus $1 \leq \#D \leq N - 1$. Hence,

$$\sum_{a,b} \mathbb{P}[Z^a \neq Z^b] \geq 2(N-1).$$

There are at most $N(N-1)$ nonzero terms in the sum above, hence there exist $a \neq b$ such that

$$\mathbb{P}[Z^a \neq Z^b] \geq 2/N.$$

Since $\|p_a - p_b\| = 1/(N-1)$, this yields $\mathbb{P}[Z^a \neq Z^b] \geq 2(1 - 1/N)\|p_a - p_b\|$, which ends the proof. $\square$

2.3. *Proof of Theorem* 2. Let $W = (W_n)_{n \in \mathbb{Z}}$ be an i.i.d. sequence of random variables, uniformly distributed on S, independent of the process X. Applying the construction of governing sequences in Section 2.1, we introduce, for every n in $\mathbb{Z}$,

$$U_n = f_{X_n}(W_n, P_{n-1}) \qquad \text{with } P_{n-1} = p(\cdot | X_{n-1}^{\triangleleft}).$$

Let $n \in \mathbb{Z}$. Thanks to the stationarity of the process X and to the independence of X and W, P_n is the conditional law of X_{n+1} given $\mathcal{F}_n^{X,W}$. Since W_{n+1} is independent of $\mathcal{F}_n^{X,W}$ and X_{n+1}, Lemma 2.5 yields that:

(1) U_{n+1} is independent of $\mathcal{F}_n^{X,W}$, and therefore of $\mathcal{F}_n^{X,U}$;
(2) U_{n+1} is uniformly distributed on S;

(3) $X_{n+1} = g(U_{n+1}, P_n)$ almost surely.

For every T in $\mathbb{Z}$, we now define a process X^T which is a function of $(U_n)_{n\geq T+1}$ in such a way that X^T approximates X when $T \to -\infty$.

Let $X_n^T = a_0$ for $n \leq T$ with $a_0 \in A$ fixed, and assume that X_n^T is defined up to time $n \geq T$. Define

$$X_{n+1}^T = g(U_{n+1}, P_n^T) \qquad \text{where } P_n^T = p(\cdot|(X^T)_n^{\lhd}).$$

Proposition 2.6 implies that for $n \geq T$,

$$\mathbb{P}[X_{n+1} \neq X_{n+1}^T | \mathcal{F}_n^{X,U}] = \mathbb{P}[g(U_{n+1}, P_n) \neq g(U_{n+1}, P_n^T) | \mathcal{F}_n^{X,U}] \leq 2\|P_n - P_n^T\|,$$

because P_n and P_n^T are measurable for $\mathcal{F}_n^{X,U}$ and U_{n+1} is independent of $\mathcal{F}_n^{X,U}$.

For n in $\mathbb{Z}$, let L_n^T count the number of consecutive times before n such that X^T and X coincide, that is,

$$L_n^T = \max\{k \geq 0 : X_{n-k+1:n}^T = X_{n-k+1:n}\}.$$

On the event $\{L_n^T = \ell\}$, the sequences $X_n^{\lhd}$ and $(X^T)_n^{\lhd}$ coincide on their last ℓ terms. Hence, on the event $\{L_n^T = \ell\}$,

$$\|P_n - P_n^T\| \leq \sup\{\|p(\cdot|xz) - p(\cdot|yz)\|; x \in A^{\lhd}, y \in A^{\lhd}, z \in A^{\ell}\} = \delta_\ell.$$

One gets

$$\mathbb{P}[X_{n+1} \neq X_{n+1}^T | \mathcal{F}_n^{X,U}] \leq 2\delta_{L_n^T}.$$

The end of our proof follows the method in [1]: consider a $\mathbb{Z}^+$-valued Markov chain, $Z = (Z_n)_{n\geq 0}$ starting from $Z_0 = 0$, with transition probabilities

$$p_{i,i+1} = 1 - 2\delta_i, \qquad p_{i,0} = 2\delta_i \qquad \text{for every } i \geq 0.$$

For any $n \geq T$, it happens that L_n^T dominates stochastically Z_{n-T}, in the sense of the following lemma.

LEMMA 2.8. *For every $k \geq 0$ and $n \geq T$, $\mathbb{P}[L_n^T \geq k] \geq \mathbb{P}[Z_{n-T} \geq k]$.*

PROOF. The result is obvious for $n = T$ since $Z_0 = 0$. Assume that the result holds for $n \geq T$. Then,

$$\begin{aligned}
\mathbb{P}[L_{n+1}^T \geq k+1] &= \mathbb{P}[L_n^T \geq k, X_{n+1} = X_{n+1}^T] \\
&= \mathbb{E}[\mathbf{1}_{\{L_n^T \geq k\}} \mathbb{P}[X_{n+1} = X_{n+1}^T | \mathcal{F}_n^{X,U}]] \\
&\geq \mathbb{E}[\mathbf{1}_{\{L_n^T \geq k\}} (1 - 2\delta_{L_n^T})].
\end{aligned}$$

Since $(\delta_i)_i$ is nonincreasing and since $2\delta_i < 1$ for every i, the sequence indexed by i of general term $\mathbf{1}_{\{i\geq k\}}(1 - 2\delta_i)$ is nondecreasing. By induction, one obtains

$$\begin{aligned}
\mathbb{P}[L_{n+1}^T \geq k+1] &\geq \mathbb{E}[\mathbf{1}_{\{Z_{n-T} \geq k\}}(1 - 2\delta_{Z_{n-T}})] \\
&= \mathbb{P}[Z_{n-T} \geq k, Z_{n-T+1} = Z_{n-T} + 1] \\
&= \mathbb{P}[Z_{n-T+1} \geq k+1],
\end{aligned}$$

which ends the recurrence over $n \geq T$ and the proof of the lemma. $\square$

Using this to estimate $\mathbb{P}[L_n^T = 0]$, one gets

$$\mathbb{P}[X_n^T \neq X_n] = \mathbb{P}[L_n^T = 0] \leq \mathbb{P}[Z_{n-T} = 0].$$

Let μ be the measure defined on $\mathbb{Z}^+$ by

$$\nu(k) = \prod_{n=0}^{k} (1 - 2\delta_n)$$

for every $k \geq 0$. the hypothesis of Theorem 2 ensure that μ has infinite mass.

If $\prod_{n=0}^{+\infty}(1 - 2\delta_n) = 0$, then μ is invariant and the state 0 is recurrent. But the chain Z is irreducible since the hypothesis of Theorem 2 forces the positivity of the probabilities $(1 - 2\delta_n)$. Hence Z is null recurrent.

If $\prod_{n=0}^{+\infty}(1 - 2\delta_n) > 0$, then Z is transient since for every i, the probability of never returning to i from i is $\prod_{n=i}^{+\infty}(1 - 2\delta_n) > 0$. In both cases

$$\mathbb{P}[X_n^T \neq X_n] \to 0 \qquad \text{when } T \to -\infty.$$

In other words, X_n^T converges in probability to X_n when $T \to -\infty$; in particular X_n is measurable for $\mathcal{F}_n^U$, which proves that U generate X. Using Lemma 1.1 one gets that the filtration $\mathcal{F}^X$ is immersed in $\mathcal{F}^U$, and therefore $\mathcal{F}^X$ is standard. This ends the proof of Theorem 2.

3. Average influences.

This section is devoted to the proof of Theorem 3.

3.1. *Priming lemma.* Recall that the governing sequence U with values in S and based on Lemma 2.5 is defined by

$$U_n = f_{X_n}(W_n, P_{n-1}) \qquad \text{where } P_{n-1} = p(\cdot | X_{n-1}^{\triangleleft}),$$

where $W = (W_n)_{n \in \mathbb{Z}}$ is a sequence of i.i.d. random variables uniform on S, independent of X. Recall also that, from Lemma 2.5, $X_n = g(U_n, P_{n-1})$ almost surely for every $n \geq 0$.

Let $\ell > 0$. Let us show that with probability close to 1, for each x in A^ℓ, $X_{1:\ell} = x$ as soon as $U_{1:\ell} \in B_x$ where B_x is a measurable subset of S^ℓ with μ-measure independent of x.

Recall that X satisfies the priming condition if for every a in A, $p(a | X_{-1}^{\triangleleft}) > 0$ almost surely.

LEMMA 3.1 (Priming lemma). *Set $\ell > 0$. If X verifies the priming condition, then for every $\varepsilon \in]0, 1[$, there exist a real number $\beta_\ell > 0$ and a collection $(B_x)_{x \in A^\ell}$ of Borel sets of S^ℓ such that for every $x \in A^\ell$,*

$$\mu^{\otimes \ell}[B_x] = \beta_\ell \quad \textit{and} \quad \mathbb{P}[X_{1:\ell} = x | U_{1:\ell} \in B_x] \geq 1 - \varepsilon.$$

Therefore if Y is a random variable valued in A^ℓ independent of $(X_n, U_n)_{n \in \mathbb{Z}}$,

$$\mathbb{P}[U_{1:\ell} \in B_Y] = \beta_\ell \quad \textit{and} \quad \mathbb{P}[X_{1:\ell} = Y | U_{1:\ell} \in B_Y] \geq 1 - \varepsilon.$$

PROOF. For every $n \in \mathbb{Z}$, let $P_n = p(\cdot|X_n^{\triangleleft})$. Thanks to the stationarity of the process X, P_n is the conditional law of X_{n+1} given $\mathcal{F}_n^X$ and the priming condition ensures that the support of P_n is A almost surely.

Let $\varepsilon \in]0,1[$ and $\ell > 0$. For any fixed $x \in A^\ell$ let us construct by induction Borel sets $B_1^x, \dots, B_\ell^x$ of S with positive measure such that for every $m \in \{1, \dots, \ell\}$,

$$\mathbb{P}(C_m) \geq \left(1 - \frac{\varepsilon}{\ell}\right)\mu(B_m^x)\mathbb{P}(C_{m-1}) > 0,$$

where $C_0 = \Omega$ and for every $m \in \{1, \dots, \ell\}$,

$$C_m = \{X_{1:m} = x_{1:m}; U_{1:m} \in B_1^x \times \cdots \times B_m^x\}.$$

Let $m \in \{1, \dots, \ell\}$. Assume that $B_1^x, \dots, B_{m-1}^x$ are constructed verifying the induction hypothesis. Since $\mathbb{P}[C_{m-1}] > 0$, one gets, thanks to the priming condition,

$$\mathbb{P}[P_{m-1}(x_m) = 0|C_{m-1}] = 0.$$

Therefore one can choose a real number $q \in]0, 1]$ such that

$$\mathbb{P}[P_{m-1}(x_m) \leq q|C_{m-1}] < \frac{\varepsilon}{\ell}.$$

Set $U_m = (U_{m,1}, \dots, U_{m,N})$. Since $X_m = g(U_m, P_{m-1})$ almost surely, one gets up to negligible events,

$$\begin{aligned}
\{X_m = x_m\} &\supset \{X_m = x_m; P_{m-1}(x_m) > q\} \\
&= \{g(U_m, P_{m-1}) = x_m; P_{m-1}(x_m) > q\} \\
&= \left\{\frac{U_{m,x_m}}{P_{m-1}(x_m)} = \min_{k \in A} \frac{U_{m,k}}{P_{m-1}(k)}; P_{m-1}(x_m) > q\right\} \\
&\supset \left\{\frac{U_{m,x_m}}{q} \leq \min_{k \neq x_m} U_{m,k}; P_{m-1}(x_m) > q\right\}.
\end{aligned}$$

Set

$$B_m^x = \left\{(y_1, \dots, y_N) \in S; \frac{y_{x_m}}{q} \leq \min_{k \neq x_m} y_k\right\}.$$

Then $\mu(B_m^x) > 0$ and

$$\{X_m = x_m; U_m \in B_m^x\} \supset \{U_m \in B_m^x; P_{m-1}(x_m) > q\}.$$

Since

$$C_m = \{X_m = x_m; U_m \in B_m^x\} \cap C_{m-1},$$

the independence of U_m and $\mathcal{F}_{m-1}^{X,U}$ and the choice of q yield

$$\begin{aligned}
\mathbb{P}[C_m] &\geq \mathbb{P}[U_m \in B_m^x; P_{m-1}(x_m) > q; C_{m-1}] \\
&= \mu(B_m^x)\mathbb{P}[P_{m-1}(x_m) > q; C_{m-1}] \\
&\geq \mu(B_m^x)\left(1 - \frac{\varepsilon}{\ell}\right)\mathbb{P}[C_{m-1}].
\end{aligned}$$

Therefore $\mathbb{P}[C_m] > 0$.

By reducing the Borel set B_ℓ^x at the last step of the induction, one can make the measure $\mu^{\otimes \ell}[B_1^x \times \cdots \times B_\ell^x]$ independent of $x \in A^\ell$. Denote by β_ℓ this measure, and then set $B_x = B_1^x \times \cdots \times B_\ell^x$. One gets

$$\mathbb{P}[X_{1:\ell} = x | U_{1:\ell} \in B_x] = \frac{\mathbb{P}[X_{1:\ell} = x; U_{1:\ell} \in B_x]}{\mathbb{P}[U_{1:\ell} \in B_x]}.$$

By independence,

$$\mathbb{P}[X_{1:\ell} = x, U_{1:\ell} \in B_x] \geq \prod_{k=1}^{\ell} \mu(B_k^x)\left(1 - \frac{\varepsilon}{\ell}\right) \quad \text{and} \quad \mathbb{P}[U_{1:\ell} \in B_x] = \prod_{k=1}^{\ell} \mu(B_k^x),$$

hence

$$\mathbb{P}[X_{1:\ell} = x | U_{1:\ell} \in B_x] \geq \left(1 - \frac{\varepsilon}{\ell}\right)^\ell \geq 1 - \varepsilon,$$

which ends the proof. □

3.2. *Approximation until a given time.* Choose $\varepsilon > 0$ and $\ell \geq 1$ such that $\sum_{n \geq \ell} \eta_n \leq \varepsilon$, and let $J = [s, t]$ be an interval of integers such that $t - s + 1 = \ell$.

Then, let Y be a random variable taking values in A^ℓ, independent of $(X_n, U_n)_{n \in \mathbb{Z}}$ and distributed like X_J.

Lemma 3.1 provides a real number β_ℓ and Borel sets $(B_x)_{x \in A^\ell}$, such that

$$\mathbb{P}[X_J = Y | U_J \in B_Y] \geq 1 - \varepsilon \quad \text{and} \quad \mathbb{P}[U_J \in B_Y] = \beta_\ell.$$

Using Y and the governing sequence $(U_n)_{n \geq t+1}$, let us construct random variables $(X'_n)_{n \geq s}$ by taking $X'_J = Y$ and for every $n > t$

$$X'_n = g(U_n, P'_{n-1}) \qquad \text{where } P'_{n-1} = p(\cdot | X'_{s:n-1}).$$

The random variable Y is useful in the proof of our following result.

LEMMA 3.2. *For every $n \geq s$, the law of $X'_{s:n}$ is the law of $X_{s:n}$.*

PROOF. For every $n \geq t + 1$, $y \in A^{n-s}$ and all $x \in A$,

$$\mathbb{P}[X'_n = x | X'_{s:n-1} = y] = p(x|y) = \mathbb{P}[X_n = x | X_{s:n-1} = y].$$

Since the law of $X'_J = Y$ is the same as the law of X_J, the result follows by induction. □

LEMMA 3.3. *One has $\mathbb{P}[X' \neq X \text{ on } [s, +\infty[| U_J \in B_Y] \leq 3\varepsilon$.*

PROOF. Since $X_n = g(U_n, P_{n-1})$ and $X'_n = g(U_n, P'_{n-1})$, Proposition 2.6 yields for $n > t$,

$$\mathbb{P}[X'_n \neq X_n | \mathcal{F}^{X,U}_{n-1} \vee \sigma(Y)] \leq 2\|P'_{n-1} - P_{n-1}\|.$$

Let

$$(\star) = \mathbb{P}[X'_n \neq X_n;\, X'_{s:n-1} = X_{s:n-1};\, U_J \in B_Y].$$

Since

$$\{X'_{s:n-1} = X_{s:n-1};\, U_J \in B_Y\} \in \mathcal{F}^{X,U}_{n-1} \vee \sigma(Y),$$

one gets

$$\begin{aligned}
(\star) &\leq \mathbb{E}\big[2\|P_{n-1} - P'_{n-1}\|\mathbf{1}_{\{X'_{s:n-1}=X_{s:n-1}\}}\mathbf{1}_{\{U_J \in B_Y\}}\big] \\
&= 2 \sum_{\substack{y \in A^\ell \\ z \in A^{n-t-1}}} \mathbb{E}\big[\|p(\cdot|X^{\lhd}_{s-1}yz) - p(\cdot|yz)\|\mathbf{1}_{\{yz=X'_{s:n-1}=X_{s:n-1}\}}\mathbf{1}_{\{U_J \in B_y\}}\big] \\
&\leq 2\sum_{y,z} \mathbb{E}\big[\|p(\cdot|X^{\lhd}_{s-1}yz) - p(\cdot|yz)\|\mathbf{1}_{\{U_J \in B_y\}}\mathbf{1}_{\{yz=X'_{s:n-1}\}}\big] \\
&= 2\sum_{y,z} \mathbb{E}\big[\|p(\cdot|X^{\lhd}_{s-1}yz) - p(\cdot|yz)\|\big]\mu^{\otimes \ell_k}(B_y)\mathbb{P}[X'_{s:n-1} = yz] \\
&= 2\beta_\ell \sum_{x \in A^{n-s}} \mathbb{E}\big[\|p(\cdot|X^{\lhd}_{s-1}x) - p(\cdot|x)\|\big]\mathbb{P}[X_{s:n-1} = x] \\
&= 2\beta_\ell \eta_{n-s},
\end{aligned}$$

where the last three equations stem from the independence of $X^{\lhd}_{s-1}$, U_J, $U_{t+1:n-1}$ and Y, from Lemma 3.2 and from the definition of η_n. Hence,

$$\mathbb{P}[X'_n \neq X_n;\, X'_{s:n-1} = X_{s:n-1} | U_J \in B_Y] \leq 2\eta_{n-s},$$

therefore,

$$\mathbb{P}[X'_{s:n} = X_{s:n} | U_J \in B_Y] \geq \mathbb{P}[X'_{s:n-1} = X_{s:n-1} | U_J \in B_Y] - 2\eta_{n-s}.$$

By induction, one gets

$$\mathbb{P}[X'_{s:n} = X_{s:n} | U_J \in B_Y] \geq \mathbb{P}[X'_J = X_J | U_J \in B_Y] - 2\sum_{m=\ell}^{n-s} \eta_m.$$

Since $X'_J = Y$ and $\mathbb{P}[X_J = Y | U_J \in B_Y] \geq 1 - \varepsilon$, this yields

$$\mathbb{P}\big[X' = X \text{ on } [s, +\infty[| U_J \in B_Y\big] \geq 1 - \varepsilon - 2\sum_{m=\ell}^{\infty} \eta_m \geq 1 - 3\varepsilon,$$

which ends the proof. $\square$

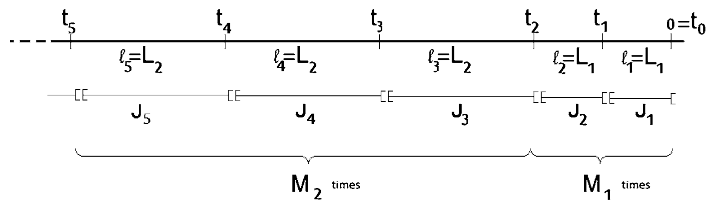

FIG. 3. *Splitting $\mathbb{Z}_-^*$ in intervals of times.*

3.3. *Successive approximations and end of the proof of Theorem 3.* Our next step in the proof of Theorem 3 is to approach the random variable X_0 by measurable functions of the governing sequence. To this aim, we group the innovations by intervals of times. For every $m > 0$ one chooses L_m such that

$$\sum_{n \geq L_m} \eta_n \leq 1/m.$$

For each m, Lemma 3.1 (the priming lemma) applied to $\ell = L_m$ and $\varepsilon = 1/m$ provides a real number $\beta_{L_m} > 0$ and Borel sets $(B_x)_{x \in A^{L_m}}$ of S^{L_m} with measure β_{L_m} such that

$$\mathbb{P}[X_{1:L_m} = x | U_{1:L_m} \in B_x] \geq 1 - 1/m.$$

Choose an integer $M_m \geq 1/\beta_{L_m}$. Split $\mathbb{Z}_-^*$ into M_1 intervals of length L_1, M_2 intervals of length $L_2, \ldots$. More precisely set, for every $n \geq 1$,

$$\ell_n = L_m \qquad \text{if } M_1 + \cdots + M_{m-1} < n \leq M_1 + \cdots + M_m$$

and

$$\varepsilon_n = 1/m \qquad \text{if } M_1 + \cdots + M_{m-1} < n \leq M_1 + \cdots + M_m.$$

Therefore, for every $k \geq 0$ one gets

$$\sum_{n \geq \ell_k} \eta_n \leq \varepsilon_k.$$

At last, for every $k \geq 0$, set

$$t_k = -\sum_{1 \leq n \leq k} \ell_n;$$

that is to say $t_0 = 0$ and $t_k = t_{k-1} - \ell_k$ for $k \geq 1$. Define, for $k \geq 0$, the interval of integers

$$J_k = [t_k, t_k + \ell_k - 1] = [t_k, t_{k-1} - 1] \quad \text{and} \quad X_{J_k} = X_{t_k : t_{k-1} - 1}.$$

Let $Y = (Y_k)_{k \geq 1}$ be a sequence of random variables, independent of $(X_n, U_n)_{n \in \mathbb{Z}}$ and such that for every $k \geq 1$, the law of Y_k is the law of $X_{1:\ell_k}$.

For every $k \geq 0$, let us use the construction of Section 3.2: set $X^k_{J_k} = Y_k$, and then for every $n \geq t_k + \ell_k = t_{k-1}$,

$$X^k_n = g(U_n, P^k_{n-1}) \qquad \text{where } P^k_{n-1} = p(\cdot|X^k_{t_k:n-1}).$$

Therefore Lemma 3.3 yields the inequality

$$\mathbb{P}[X_{t_k:0} \neq X^k_{t_k:0}|U_{J_k} \in B_{Y_k}] \leq 3\varepsilon_k$$

and

$$\mathbb{P}[X^k_0 \neq X_0|U_{J_k} \in B_{Y_k}] \to 0 \qquad \text{when } k \to +\infty.$$

Moreover each event $\{U_{J_k} \in B_{Y_k}\}$ is independent of the others (indeed they are functions of random variables U_k for disjoint sets of indices k) and

$$\sum_{k\geq 1} \mathbb{P}[U_{J_k} \in B_{Y_k}] = \sum_{k\geq 1} \beta_{\ell_k} = \sum_{m=1}^{+\infty} M_m \beta_{L_m} = +\infty$$

since $M_m \beta_{L_m} \geq 1$ by choice of M_m.

Lemma 3.4, stated below, provides a deterministic increasing function θ such that

$$\sum_{k\geq 1} \mathbb{P}\bigl[X^{\theta(k)}_0 \neq X_0; U_{J_{\theta(k)}} \in B_{Y_{\theta(k)}}\bigr] < +\infty$$

and

$$\sum_{k\geq 1} \mathbb{P}\bigl[U_{J_{\theta(k)}} \in B_{Y_{\theta(k)}}\bigr] = +\infty.$$

Using the Borel–Cantelli lemma, one deduces that:

- $\{X^{\theta(k)}_0 \neq X_0\} \cap \{U_{J_{\theta(k)}} \in B_{Y_{\theta(k)}}\}$ is realized for a finite number of k only, a.s.
- $\{U_{J_{\theta(k)}} \in B_{Y_{\theta(k)}}\}$ is realized for an infinite number of k a.s.

Thus, for every $a \in A$,

$$\{X_0 = a\} = \limsup_{k\to\infty} \{U_{J_{\theta(k)}} \in B_{Y_{\theta(k)}}\} \cap \{X^{\theta(k)}_0 = a\}.$$

Therefore, $\{X_0 = a\}$ belongs to $\mathcal{F}^U_0 \vee \sigma(Y)$. Since the sequence $Y = (Y_k)_{k\geq 0}$ is independent of $\mathcal{F}^{U,X}_0$, one gets

$$\mathbb{P}[X_0 = a|\mathcal{F}^U_0] = \mathbb{P}[X_0 = a|\mathcal{F}^U_0 \vee \sigma(Y)] = \mathbf{1}_{\{X_0=a\}} \qquad \text{a.s.},$$

therefore $\{X_0 = a\} \in \mathcal{F}^U_0$. By stationarity of the process (X, U), one gets the inclusion of the filtration $\mathcal{F}^X$ into the filtration $\mathcal{F}^U$. Therefore Lemma 1.1 yields that $\mathcal{F}^X$ is immersed in $\mathcal{F}^U$, which ends the proof.

LEMMA 3.4. *Let $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$ denote two bounded sequences of nonnegative real numbers such that the series $\sum_n b_n$ diverges and such that $a_n \ll b_n$. Then there exists an increasing function $\theta : \mathbb{N} \to \mathbb{N}$ such that the series $\sum_n a_{\theta(n)}$ converges, and the series $\sum_n b_{\theta(n)}$ diverges.*

4. Examples. In this section we study some examples showing the advantages and the limitations of our results.

- Our first example (Section 4.1) is a chain with memory of variable length which fulfills the hypotheses of Theorem 3, but not those of Theorem 1 nor Theorem 2. Its natural filtration is standard.
- Our second example (Section 4.2) is derived from the well-known $[T, T^{-1}]$ transformation. It provides a stationary process with values in a finite space, whose natural filtration is not standard. This example does not fulfill any of the two conditions of Theorem 3 (viz., the priming condition and the summability of the gaps).
- Our third example (Section 4.3) is a slight adaptation of the second one, where the filtration of the process is still nonstandard, although the priming condition is fulfilled.
- Our fourth and last example (Section 4.4) is another adaptation of the second example in which the filtration is standard, although the condition of summability of the gaps is not fulfilled and the related conditional probabilities are close to those of the second example.

4.1. *First example*: *Parity of the number of* 1 *in a row.* This example provides a setting where one proves standardness using Theorem 3.

Let $(X_n)_{n\in\mathbb{Z}}$ be a stationary process taking values in $\{0, 1\}$ such that

$$\mathbb{P}[X_0 = 0 | \mathcal{F}^X_{-1}] = \tfrac{1}{3} + \tfrac{1}{3}\mathbf{1}_{\{T \text{ is even or } T=-\infty\}} \qquad \text{where } T := \sup\{k < 0 : X_k = 0\}.$$

The existence of such a process is ensured by Proposition 2.10 in [4]. A simple computation gives, for every $n \geq 0$,

$$\gamma_n = \tfrac{1}{2}, \delta_n = \tfrac{1}{3}, \qquad \alpha_n = \tfrac{1}{3}, \qquad \eta_n \leq \tfrac{1}{3}\mathbb{P}[T < -n] \leq \left(\tfrac{2}{3}\right)^{n+1}.$$

Therefore this process fulfills the hypotheses of Theorem 3 (and its Corollary 1.2), but neither those of Theorem 1 nor those of Theorem 2. The filtration $\mathcal{F}^X$ is standard.

4.2. *Second example*: *Random walk in random scenery.* The following is a process whose filtration is not standard.

Let $X = (X_n)_{n\in\mathbb{Z}}$ and $C = (C_s)_{s\in\mathbb{Z}}$ be two independent sequences of i.i.d. random variables with uniform law on $\{-1, 1\}$. Set

$$S_n = X_1 + \cdots + X_n \qquad \text{if } n \geq 0,$$
$$S_n = -X_{n+1} - \cdots - X_0 \qquad \text{if } n < 0.$$

Therefore $S_{n+1} = S_n + X_{n+1}$ for every $n \in \mathbb{Z}$. Set $C_{S_n} = Y_n$. The stationary process Z, defined by $Z_n = (X_n, Y_n)$ for every $n \in \mathbb{Z}$ and taking values in $A = \{-1, 1\}^2$, is called random walk in random scenery.

This process is derived from the process $((X_{n+\cdot}, C_{S_n+\cdot}))_{n\in\mathbb{Z}}$ where $X_{n+\cdot} = (X_{n+m})_{m\leq 0}$ is the trajectory of X until time n and $C_{S_n+\cdot} = (C_{S_n+s})_{s\in\mathbb{Z}}$ is the scenery seen from S_n. It is easy to prove that the processes Z and $(X_{n+\cdot}, C_{S_n+\cdot})_{n\in\mathbb{Z}}$ generate the same filtration. Indeed, given (X_k, C_{S_k}) for every $k \leq n$, one knows the trajectory $X_{n+\cdot}$ and one can deduce the increments $(S_n - S_k)_{k\leq n}$. Since those increments visit almost surely every integer, one can recover the scenery seen from S_n.

The process $((X_{n+\cdot}, C_{S_n+\cdot}))_{n\in\mathbb{Z}}$ is the most famous $[T, T^{-1}]$ process. Indeed the $[T, T^{-1}]$ transformation is the application from $\{-1, 1\}^{\mathbb{Z}_-} \times \{-1, 1\}^{\mathbb{Z}}$ into itself defined by

$$[T, T^{-1}]((x_n)_{n\leq 0}, (c_s)_{s\in\mathbb{Z}}) = ((x_{n-1})_{n\leq 0}, (c_{s-x_0})_{s\in\mathbb{Z}}).$$

One checks that for every $n \leq 0$

$$(X_{n+\cdot}, C_{S_n+\cdot}) = [T, T^{-1}]^{-n}(X, C).$$

According to [5], the natural filtration of the process $((X_{n+\cdot}, C_{S_n+\cdot}))_{n\in\mathbb{Z}}$ is not standard though its asymptotic σ-field at $-\infty$ is trivial. Therefore the same holds for the natural filtration of Z.

Let $n \geq 0$. Let us study the probabilities $p(a|z)$ for $a \in A$ and $z \in A^n$. Note $z = (z_{-n}, \dots, z_{-1})$, $z_k = (x_k, y_k)$ and $a = (x_0, y_0)$.

First, note that some of the events

$$\{Z_{-n:-1} = z\} = \{X_{-n:-1} = x_{-n:-1}; Y_{-n:-1} = y_{-n:-1}\}$$

are impossible. Indeed, by definition of the process Y for $i < j$, $Y_i = Y_j$ on the event $X_{i+1} + \cdots + X_j = 0$. When the event $\{Z_{-n:-1} = z\}$ is impossible, one says that z is not admissible. Note that to compute η_n, one only needs to consider probabilities $p(\cdot|z)$ and $p(\cdot|wz)$ for admissible $z \in A^n$ and $w \in A^{\lhd}$. Yet, wz may be nonadmissible, even if z and w are admissible.

Assume that z is admissible. Then $\mathbb{P}[X_0 = x_0|Z_{-n:-1} = z] = 1/2$. If for some $i \in \{-n, \dots, -1\}$, $x_{i+1} + \cdots + x_0 = 0$, then the conditions $Z_{-n:-1} = z$ and $X_0 = x_0$ imply that $Y_0 = y_i$. Otherwise, the color Y_0 is independent of $Z_{-n:-1}$ and X_0. Thus for any admissible word $z \in A^n$,

$$p(a|z) = \begin{cases} 1/2, & \text{if there exists } i \text{ such that } x_{i+1} + \cdots + x_0 = 0 \text{ and } y_0 = y_i, \\ 0, & \text{if there exists } i \text{ such that } x_{i+1} + \cdots + x_0 = 0 \text{ and } y_0 \neq y_i, \\ 1/4, & \text{otherwise.} \end{cases}$$

Therefore for every $n \geq 0$, $\gamma_n = 1$ and $\delta_n = \alpha_n = 1/2$.

Furthermore, for almost every admissible word $w = (x_n, y_n)_{n<0}$ in $A^{\lhd}$, there exists $t < 0$ such that $x_{t+1} + \cdots + x_0 = 0$ and the same argument gives that

$$p(a|w) = \tfrac{1}{2}\mathbf{1}_{\{y_t = y_0\}}.$$

For nonadmissible $w \in A^{\lhd}$, the value of $p(a|w)$ can be chosen arbitrarily. Set

$$p(a|w) = \tfrac{1}{2}\mathbf{1}_{\{y_d = y_0\}} \qquad \text{where } d = \sup\{t \leq -1 : x_{t+1} + \cdots + x_0 = 0\},$$

if d is well defined, and $p(a|w) = 1/4$ otherwise.

With this convention, one gets that, for almost any admissible $w \in A^{\triangleleft}$ and $z = ((x_{-n}, y_{-n}), \dots, (x_{-1}, y_{-1})) \in A^n$,

$$|p(a|z) - p(a|wz)| = \begin{cases} 0, & \text{if there exists } t \in [-n, -1] \\ & \text{such that } x_t + \cdots + x_0 = 0, \\ 1/4, & \text{otherwise.} \end{cases}$$

Therefore, for every $n \geq 0$,

$$\begin{aligned}\eta_n &= \tfrac{1}{4}(\mathbb{P}[\forall i \in \{-n, \dots, -1\}, X_{i+1} + \cdots + X_{-1} \geq 0] \\ &\qquad + \mathbb{P}[\forall i \in \{-n, \dots, -1\}, X_{i+1} + \cdots + X_{-1} \leq 0]) \\ &= \tfrac{1}{2}\mathbb{P}[\forall i \in \{-n, \dots, -1\}, X_{i+1} + \cdots + X_{-1} \geq 0].\end{aligned}$$

One sees that $\eta_n \sim C/\sqrt{n}$ with $C \in \mathbb{R}_*^+$. Hence the process Z does not fulfill any of the hypotheses of Theorem 3.

4.3. *Third example*: *Random walk in random scenery with misreading*. We construct a variant of the random walk in random scenery which fulfills the priming condition but whose natural filtration is not standard.

Construct $Z_n = (X_n, Y_n)$ as in Section 4.2. Fix $q \in]0, 1/2[$. Let $(\xi_n)_{n \in \mathbb{Z}}$ be a sequence of i.i.d. random variables taking values in $\{-1, 1\}$, independent of $\mathcal{F}^Z$ and such that $\mathbb{P}[\xi_0 = 1] = 1 - q$. Define a process $(Z'_n)_{n \in \mathbb{Z}}$ by

$$Z'_n = (X_n, Y_n \xi_n).$$

The process $(Z'_n)_{n \in \mathbb{Z}}$ is a random walk in random scenery in which at each time, one misreads the color of the site Y_n with probability q.

The processes Z' and $(Z_n, \xi_n)_{n \in \mathbb{Z}}$ generate the same filtration. Indeed, the random variables ξ_m associated to the times $m < n$ where $S_m = S_n$ are independent and take the value 1 with probability $1 - q > 1/2$, therefore the color Y_n is the most common color among the colors $Y_m \xi_m$ seen at those times. Therefore, for almost every $z \in A^{\triangleleft}$ and $a \in A$, the corresponding conditional probability $p'(a|z)$ is equal to $q/2$ or to $(1 - q)/2$, depending on these colors.

Moreover, by independent enlargement, $\mathcal{F}^Z$ is immersed into $\mathcal{F}^{Z,\xi} = \mathcal{F}^{Z'}$. Since $\mathcal{F}^Z$ is nonstandard, one deduces that $\mathcal{F}^{Z'}$ is not standard either.

By a short calculation, one gets for every $n > 0$,

$$\gamma_n = \frac{q}{1 - q}, \qquad \delta_n = \frac{1 - 2q}{4}, \qquad \alpha_n = 1 - 2q.$$

Since the probabilities $p'(a|z)$ related to this process satisfy $p'(a|z) \geq q/2$, for every $a \in A$ and $z \in A^{\triangleleft}$, the priming condition is fulfilled. The exact value of

$p'(a|z)$ for $z \in A^n$ is difficult to compute, but the corresponding gaps η'_n verify

$$\begin{aligned}\eta'_n &\geq \left(\frac{1}{4} - \frac{q}{2}\right)(\mathbb{P}[\forall i \in \{-l, \dots, -1\}, X_{i+1} + \cdots + X_{-1} \geq 0] \\ &\qquad\qquad + \mathbb{P}[\forall i \in \{-l, \dots, -1\}, X_{i+1} + \cdots + X_{-1} \leq 0]) \\ &= (1-2q)\eta_n.\end{aligned}$$

Therefore for $q < 1/2$, the sequence $(\eta'_n)_n$ is not summable; thus the process Z' does not verify the condition of summability of the gaps.

4.4. *Fourth example*: *Random walk in renewed random scenery.* We construct another variant of the random walk in random scenery in which the natural filtration is standard although the condition of summability of the gaps of Theorem 3 is not fulfilled.

We consider a variant of the process $(X_n, C_{S_n+\cdot})_{n\in\mathbb{Z}}$ in which at each time n, the color at 0 of the scenery seen from S_n is changed with probability $q \in]0, 1/2[$. For every $g \in \{-1, 1\}^{\mathbb{Z}}$, denote $\overline{g} \in \{-1, 1\}^{\mathbb{Z}}$ the application defined by

$$\overline{g}(s) = g(s) \qquad \text{for } s \neq 0 \text{ and } \overline{g}(0) = -g(0).$$

Let (X_n, G_n) be a stationary Markov chain with values in $\{-1, 1\} \times \{-1, 1\}^{\mathbb{Z}}$, with transition probabilities

$$p((x, g), (x', g')) = \begin{cases} (1-q)/2, & \text{if } g' = g(x' + \cdot), \\ q/2, & \text{if } g' = \overline{g}(x' + \cdot). \end{cases}$$

The random walk in renewed random scenery is the process $Z'' = (Z''_n)_{n\in\mathbb{Z}}$ defined by $Z''_n = (X_n, G_n(0))$.

The corresponding probabilities $p''(a|z)$ are close to the probabilities $p(a|z)$. Indeed,

$$p''(a|z) = \begin{cases} (1-q)/2, & \text{if } p(a|z) = 1/2, \\ q/2, & \text{if } p(a|z) = 0, \\ 1/4, & \text{if } p(a|z) = 1/4. \end{cases}$$

Therefore the corresponding gaps verify

$$\eta''_n = (1-2q)\eta_n.$$

To show that the filtration $\mathcal{F}^{Z''}$ is standard, one can use the following trick: instead of changing the color at 0 of the scenery G_n with probability q, one draws at random this color with probability $2q$. One needs a random variable ε_n, taking the value 1, if this drawing occurs and 0 otherwise, and a random variable κ_n, giving the color obtained if the drawing occurs.

To construct these random variables, consider two independent sequences of random variables $(\beta_n)_{n\in\mathbb{Z}}$ and $(V_n)_{n\in\mathbb{Z}}$, independent of $\mathcal{F}^{X,G}$ such that:

- the β_n are i.i.d. Bernoulli variables of parameter $(1-2q)/(1-q)$;

- the V_n are i.i.d. and uniform on $\{-1, 1\}$.

Let, for every $n \in \mathbb{Z}$,

$$\varepsilon_n = 1 - \beta_n \mathbf{1}_{\{G_n(0)=G_{n-1}(X_n)\}} \quad \text{and} \quad \kappa_n = G_n(0)\mathbf{1}_{\{\varepsilon_n=1\}} + V_n \mathbf{1}_{\{\varepsilon_n=0\}}.$$

Let us show that the random variables $U_n = (X_n, \varepsilon_n, \kappa_n)$ constitute a governing sequence for the process Z''. Given Z''_{n-1} and U_n, one deduces Z''_n, thanks to the equalities

$$G_n(s) = G_{n-1}(X_n + s) \qquad \text{if } \varepsilon_n = 0 \text{ or } s \neq 0,$$
$$G_n(0) = \kappa_n \qquad \text{if } \varepsilon_n = 1.$$

It remains to check that U_n is independent of the σ-field $\mathcal{G}_{n-1} = \mathcal{F}^{X,G,\beta,V}_{n-1}$ and a fortiori of $\mathcal{F}^{Z'',U}_{n-1}$. Thanks to the independence of the processes β, V and (X, G) one gets for every $x \in \{-1, 1\}$,

$$\mathbb{P}[X_n = x; G_n = G_{n-1}(X_n + \cdot)|\mathcal{G}_{n-1}] = (1-q)/2,$$
$$\mathbb{P}[X_n = x; G_n \neq G_{n-1}(X_n + \cdot)|\mathcal{G}_{n-1}] = q/2.$$

Therefore, for every c and x in $\{-1, 1\}$,

$$\begin{aligned}\mathbb{P}[\varepsilon_n = 1; \kappa_n = c; X_n = x|\mathcal{G}_{n-1}] &= \mathbb{P}[\varepsilon_n = 1; G_n(0) = c; X_n = x|\mathcal{G}_{n-1}] \\ &= (1) + (2) + (3)\end{aligned}$$

with

$$(1) = \mathbb{P}[\beta_n = 0; G_n \neq G_{n-1}(X_n + \cdot); G_{n-1}(X_n) = -c; X_n = x|\mathcal{G}_{n-1}],$$
$$(2) = \mathbb{P}[\beta_n = 0; G_n = G_{n-1}(X_n + \cdot); G_{n-1}(X_n) = c; X_n = x|\mathcal{G}_{n-1}],$$
$$(3) = \mathbb{P}[\beta = 1; G_n \neq G_{n-1}(X_n + \cdot); G_{n-1}(X_n) = -c; X_n = x|\mathcal{G}_{n-1}].$$

One gets

$$\begin{aligned}(1) &= \frac{q}{1-q} \times \mathbb{P}[G_n \neq G_{n-1}(X_n + \cdot); G_{n-1}(x) = -c; X_n = x|\mathcal{G}_{n-1}] \\ &= \frac{q}{2} \times \frac{q}{1-q} \times \mathbb{P}[G_{n-1}(x) = -c|\mathcal{G}_{n-1}] \\ &= \frac{q}{2} \times \frac{q}{1-q} \times \mathbf{1}_{\{G_{n-1}(x)=-c\}},\end{aligned}$$

$$\begin{aligned}(2) &= \frac{q}{1-q} \times \mathbb{P}[G_n = G_{n-1}(X_n + \cdot); G_{n-1}(x) = c; X_n = x|\mathcal{G}_{n-1}] \\ &= \frac{1-q}{2} \times \frac{q}{1-q} \times \mathbb{P}[G_{n-1}(x) = c|\mathcal{G}_{n-1}] \\ &= \frac{q}{2} \times \mathbf{1}_{\{G_{n-1}(x)=c\}}\end{aligned}$$

and

$$
\begin{aligned}
(3) &= \frac{1-2q}{1-q} \times \mathbb{P}[G_n \neq G_{n-1}(X_n + \cdot); G_{n-1}(x) = -c; X_n = x | \mathcal{G}_{n-1}] \\
&= \frac{q}{2} \times \frac{1-2q}{1-q} \times \mathbb{P}[G_{n-1}(x) = -c | \mathcal{G}_{n-1}] \\
&= \frac{q}{2} \times \frac{1-2q}{1-q} \times \mathbf{1}_{\{G_{n-1}(x)=-c\}}.
\end{aligned}
$$

Thus, for every c and x in $\{-1, 1\}$,

$$
\mathbb{P}[\varepsilon_n = 1; \kappa_n = c; X_n = x | \mathcal{G}_{n-1}] = \frac{q}{2}.
$$

Moreover, by independence of β_n, V_n and $\mathcal{G}_{n-1}$,

$$
\begin{aligned}
&\mathbb{P}[\varepsilon_n = 0; \kappa_n = c; X_n = x | \mathcal{G}_{n-1}] \\
&\qquad = \mathbb{P}[\beta_n = 1; G_n = G_{n-1}(X_n + \cdot); V_n = c; X_n = x | \mathcal{G}_{n-1}] \\
&\qquad = \frac{1-2q}{1-q} \times \frac{1}{2} \times \frac{1-q}{2} \\
&\qquad = \frac{1-2q}{4}.
\end{aligned}
$$

This shows that the random variables $U_n = (X_n, \varepsilon_n, \kappa_n)$ constitute a governing sequence for the process Z''.

Let us show the inclusion $\mathcal{F}_n^{Z''} \subset \mathcal{F}_n^U$ for any $n \in \mathbb{Z}$, that is to say that the sequence $(U_k)_{k \leq n}$ is sufficient to recover the scenery G_n seen from S_n. The variables $(X_k)_{k \leq n}$ determine the increments $(S_n - S_k)_{k \leq n}$ and for every $s \in \mathbb{Z}$, $S_n - S_k = s$ for an infinite number of times $k \leq n$. Among those times, there is an infinite number of times such that $\varepsilon_k = 1$. The value of κ_k at the last time $k \leq n$, such that $S_n - S_k = s$ and $\varepsilon_k = 1$, is equal to $G_n(s)$. Therefore $\mathcal{F}^G \subset \mathcal{F}^U$, and since $\mathcal{F}^X \subset \mathcal{F}^U$, one gets $\mathcal{F}^{Z''} \subset \mathcal{F}^U$. Finally, Lemma 1.1 yields that $\mathcal{F}^{Z''}$ is immersed in $\mathcal{F}^U$, and therefore the natural filtration of the process $(Z''_n)_{n \in \mathbb{Z}}$ is standard.

5. Proofs of auxiliary facts.

5.1. *Inequalities involving $\alpha_n, \delta_n, \gamma_n$ and η_n.* *To prove that $\delta_n \leq \gamma_n$ for every* $n \geq 0$, consider x and y in $A^{\triangleleft}$ and $z \in A^n$. Then,

$$
\begin{aligned}
\|p(\cdot|xz) - p(\cdot|yz)\| &= \sum_{a \in A} [p(a|xz) - p(a|yz)]_+ \\
&= \sum_{a \in A} p(a|xz) \left(1 - \frac{p(a|yz)}{p(a|xz)}\right)_+ \\
&\leq \sum_{a \in A} p(a|xz) \gamma_n = \gamma_n.
\end{aligned}
$$

Taking the supremum over x, y and z, one gets $\delta_n \le \gamma_n$.

To prove that $\eta_n \le \delta_n$ *for every* $n \ge 0$, consider for every $z \in A^n$, the law Q_z of $X^{\lhd}_{-n-1}$ conditionally on $X_{-n:-1} = z$. Then,

$$p(\cdot|z) = \int_{A^{\lhd}} p(\cdot|yz) Q_z(\mathrm{d}y).$$

Thus, for every x in $A^{\lhd}$ and z in A^n,

$$\begin{aligned}\|p(\cdot|z) - p(\cdot|xz)\| &= \left\| \int_{y \in A^{\lhd}} \big(p(\cdot|yz) - p(\cdot|xz)\big) Q_z(\mathrm{d}y) \right\| \\ &\le \int_{y \in A^{\lhd}} \|p(\cdot|yz) - p(\cdot|xz)\| Q_z(\mathrm{d}y) \\ &\le \sup_{y \in A^{\lhd}} \|p(\cdot|yz) - p(\cdot|xz)\| \le \delta_n.\end{aligned}$$

For every z in A^n, $\|p(\cdot|z) - p(\cdot|X^{\lhd}_{-n-1}z)\| \le \delta_n$ almost surely. Taking the expectation and the average over z, one gets $\eta_n \le \delta_n$.

To prove that $\delta_n \le \alpha_n$ *for every* $n \ge 0$, consider $z \in A^n$ and $y, y' \in A^{\lhd}$. Then,

$$\begin{aligned}\|p(\cdot|yz) - p(\cdot|y'z)\| &= \sum_{a \in A} |p(a|yz) - p(a|y'z)|_+ \\ &= \sum_{a \in A} \big(p(a|yz) - \min(p(a|yz), p(a|y'z))\big) \\ &\le 1 - \inf_{z \in A^n} \sum_{a \in A} \inf\{p(a|yz) : y \in A^{\lhd}\} \\ &= \alpha_n.\end{aligned}$$

This ends the proof.

5.2. *Proof of Lemma* 1.1. Assume that X is a process valued in a measurable space $(E, \mathfrak{E})$ and that U is a governing and generating process of X. Let $n \in \mathbb{Z}$. Since U governs X, there exists a measurable function ψ_n such that $X_{n+1} = \psi_n(U_{n+1}, X^{\lhd}_n)$ [axiom (ii)]. Let $B \in \mathfrak{E}$. We try to estimate

$$\rho = \mathbb{P}[X_{n+1} \in B | \mathcal{F}^{X,U}_n] = \mathbb{P}[\psi_n(U_{n+1}, X^{\lhd}_n) \in B | \mathcal{F}^{X,U}_n].$$

Since U governs X, U_{n+1} and $\mathcal{F}^{X,U}_n$ are independent [axiom (i)]; hence ρ is a function of $X^{\lhd}_n$ only, that is,

$$\rho = \mathbb{P}[\psi(U_{n+1}, X^{\lhd}_n) \in B | \mathcal{F}^X_n] = \mathbb{P}[X_{n+1} \in B | \mathcal{F}^X_n].$$

Hence $\mathcal{F}^X_{n+1}$ is independent of $\mathcal{F}^U_n$ conditionally on $\mathcal{F}^X_n$. This shows that $\mathcal{F}^X$ is immersed in $\mathcal{F}^U$.

5.3. *Proof of Corollary* 1.2. Assume that X is a chain with memory of variable length, and let Y be an independent copy of X. As $p(\cdot|X^{\triangleleft}_{-n-1}Y_{-n:-1}) = p(\cdot|Y_{-n:-1})$. on the event $\{\ell(Y^{\triangleleft}_{-1}) \leq n\}$,

$$\| p(\cdot|X^{\triangleleft}_{-n-1}Y_{-n:-1}) - p(\cdot|Y_{-n:-1})\| \leq \mathbf{1}_{\{\ell(Y^{\triangleleft}_{-1}) \geq n+1\}}.$$

Taking expectations, one gets $\eta_n \leq \mathbb{P}[\ell(Y^{\triangleleft}_{-1}) \geq n+1]$, hence

$$\sum_{n\geq 0} \eta_n \leq \mathbb{E}[\ell(Y^{\triangleleft}_{-1})] < +\infty.$$

This ends the proof.

5.4. *Proof that $\mathcal{H}(\gamma)$ provides a positive lower bound for $p(a|x)$.* We show that $\mathcal{H}(\gamma)$ implies the existence of $c > 0$ such that $p(a|x) \geq c$ for every x in $A^{\triangleleft}$ and a such that $\mathbb{P}[X_0 = a] > 0$.

Assume that $\mathcal{H}(\gamma)$, that is,

$$\sum_{k=0}^{+\infty} \prod_{n=0}^{k} (1-\gamma_n) = +\infty.$$

Therefore, $1-\gamma_0 > 0$. By definition of γ_0, for every $a \in A$, $x, y \in A^{\triangleleft}$,

$$p(a|x) \geq (1-\gamma_0) p(a|y).$$

Integrating this inequality with respect to the law of $X^{\triangleleft}$, one gets

$$p(a|x) \geq (1-\gamma_0)\mathbb{P}[X_0 = a].$$

Since A is finite, this ends the proof.

REFERENCES

[1] BRESSAUD, X., MAASS, A., MARTINEZ, S. and SAN MARTIN, J. (2006). Stationary processes whose filtrations are standard. *Ann. Probab.* **34** 1589–1600. MR2257656

[2] COMETS, F., FERNÁNDEZ, R. and FERRARI, P. A. (2002). Processes with long memory: Regenerative construction and perfect simulation. *Ann. Appl. Probab.* **12** 921–943. MR1925446

[3] ÉMERY, M. and SCHACHERMAYER, W. (2001). On Vershik's standardness criterion and Tsirelson's notion of cosiness. In *Séminaire de Probabilités, XXXV. Lecture Notes in Math.* **1755** 265–305. Springer, Berlin. MR1837293

[4] GALVES, A. and LÖCHERBACH, E. (2008). Stochastic chains with memory of variable length. *TICSP Series* **38** 117–133.

[5] HEICKLEN, D. and HOFFMAN, C. (1998). T, T^{-1} is not standard. *Ergodic Theory Dynam. Systems* **18** 875–878. MR1645322

[6] KALLIANPUR, G. and WIENER, N. (1956). Non-linear prediction. Technical Report No. I, ONR, Cu-2-56-Nonr-266, (39)-CIRMIP, Proj. NR-047-015.

[7] LAURENT, S. (2004). Filtrations à temps discret négatif. Ph.D. thesis, Institut de Recherche Mathématique Avancée, Univ. Louis Pasteur, Strasbourg, France.

[8] MASANI, P. (1966). Wiener's contributions to generalized harmonic analysis, prediction theory and filter theory. *Bull. Amer. Math. Soc.* **72** 73–125. MR0187018

[9] SCHACHERMAYER, W. (1999). On certain probabilities equivalent to Wiener measure, d'après Dubins, Feldman, Smorodinsky and Tsirelson. In *Séminaire de Probabilités, XXXIII. Lecture Notes in Math.* **1709** 221–239. Springer, Berlin. MR1767997

[10] SMORODINSKY, M. (1998). Processes with no standard extension. *Israel J. Math.* **107** 327–331. MR1658583

[11] VERSHIK, A. M. (1995). Theory of decreasing sequences of measurable partitions. *Saint Petersburg Mathematical Journal* **6** 705–761.

ATER
IUT ROBERT SCHUMANN
5 QUAI DE TURCKHEIM
67000 STRASBOURG
FRANCE
E-MAIL: gael.ceillier@gmail.com