Probab. Theory Relat. Fields (2012) 153:269–292
DOI 10.1007/s00440-011-0346-7

The filtration of the split-words process

Gaël Ceillier

Received: 5 July 2010 / Revised: 6 January 2011 / Published online: 13 September 2011

Abstract Smorodinsky and Laurent have initiated the study of the filtrations of split-word processes, in the framework of discrete negative time. For these filtrations, we show that Laurent's sufficient condition for non standardness is also necessary, thus yielding a practical standardness criterion. In turn, this criterion enables us to exhibit a non standard filtration which becomes standard when time is accelerated by omitting infinitely many instants of time.

Mathematics Subject Classification (2000) 60G15 · 60A99

1 Introduction

We shall be interested in filtrations, in the setting of discrete, negative time: given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, a filtration is an increasing family $\mathcal{F} = (\mathcal{F}_n)_{n\leqslant 0}$ of sub-σ-fields of $\mathcal{A}$; observe that the time n ranges over all negative integers. (Equivalently, one could consider decreasing families of σ-fields indexed by positive integers, known as reverse filtrations; but we find it more convenient to let time run forward, with $n+1$ posterior to n, at the mild cost of dealing with negative instants.) As discovered by Vershik [13], in this framework very subtle phenomena occur in the vicinity of time $-\infty$.

For a simple example, suppose that the σ-field $\cap_n \mathcal{F}_n$ is degenerate and that, for each n, $\mathcal{F}_n$ is generated by $\mathcal{F}_{n-1}$ and by some Bernoulli random variable U_n which is independent of $\mathcal{F}_{n-1}$ and uniformly distributed on the 2-set $\{0, 1\}$. Under these

G. Ceillier (✉)
Institut Fourier, UMR CNRS 5582, UFR de Mathématiques,
Université Grenoble I, BP 74, 100 Rue des Maths,
38402 St Martin d'Hères cedex, France
e-mail: gael.ceillier@gmail.com

hypotheses, it may happen that $\mathcal{F}$ contains more information than the natural filtration of the Bernoulli process $U = (U_n)_{n\leqslant 0}$ (this is similar to weak solutions in SDEs); but something more surprising is also possible: that $\mathcal{F}$ is not generated by any Bernoulli process whatsoever. Such filtrations have been called non standard by Vershik, who has given in [13] a necessary and sufficient criterion for standardness, and several examples of non standard filtrations. The rigorous definition of a standard filtration will be recalled later, in Sect. 4.

All filtrations considered in this study have an additional property: for each n, $\mathcal{F}_n$ is generated by $\mathcal{F}_{n-1}$ and by some random variable U_n which is independent from $\mathcal{F}_{n-1}$ and uniformly distributed on some finite set with r_n elements. Such a filtration is called (r_n)-adic. For these filtrations, as shown by Vershik [13], standardness turns out to be tantamount to a simpler, much more intuitive property: an (r_n)-adic filtration $\mathcal{F}$ is standard if and only if $\mathcal{F}$ is of product type, that is, $\mathcal{F}$ is the natural filtration of some process $V = (V_n)_{n\leqslant 0}$ where the V_n are independent random variables (In this case, it is easy to see that the process V can be chosen with the same law as U.). So, at first reading, 'standard' can be replaced with 'of product type' in this introduction.

When time is accelerated by extracting a subsequence, that is, when $(\mathcal{F}_n)_{n\in\mathbb{Z}^-}$ is replaced with $(\mathcal{F}_n)_{n\in Q}$ where Q is some infinite subset of the time-axis $\mathbb{Z}^-$, a standard filtration always remains standard, but a non standard one may become standard (or not). Examples of this phenomenon were first studied by Vershik in the framework of ergodic theory, and then, in a probabilistic setting, by Laurent [5]. Lacunary isomorphism theorem [2] states that, from any filtration $(\mathcal{F}_n)_{n\in\mathbb{Z}^-}$ (such that $\mathcal{F}_0$ is essentially separable), there exists $Q \subset \mathbb{Z}^-$ such that $(\mathcal{F}_n)_{n\in Q}$ is standard.

By varying the parameters in an example initially due to Vershik [13] and later modified (in the dyadic case) by Smorodinsky [9], Laurent [5] has described a family of filtrations, the split-word filtrations; he has shown some of them (the fastest ones) to be standard, and some other ones (the slowest ones) to be non standard; but an intermediate class was left undecided. We continue his study, and show that all these intermediate filtrations are in fact standard. This yields an easily verifiable necessary and sufficient condition for a split-word filtration to be standard.

As the family of split-word filtrations is stable by extracting subsequences, this criterion makes it simple to observe on these examples the transition from nonstandardness to standardness when time is accelerated. We find that this transition is, in some sense, sharp: in Example 2, we exhibit a non standard filtration $\mathcal{F}$ such that, for every infinite subset Q of $\mathbb{Z}^-$ with infinite complementary, the corresponding extracted filtration $(\mathcal{F}_n)_{n\in Q}$ is standard. This $\mathcal{F}$ is as close to being standard as possible, for, if $Q \subset \mathbb{Z}^-$ is cofinite and if $\mathcal{G}$ is any filtration, the extracted filtration $(\mathcal{G}_n)_{n\in Q}$ clearly has the same asymptotic properties (standardness, product type, etc.) as $\mathcal{G}$. To our knowledge, in the earlier literature, the best result in this direction was the existence of a non standard filtration $\mathcal{F}$ such that $(\mathcal{F}_{2n})_{n\leqslant 0}$ is standard (examples are given by Vershik [13], Gorbulsky [3] and Tsirelson in an unpublished paper).

In this paper we study the filtrations of split-word processes. These processes are inspired by examples given by Vershik [13, example 2,3,4] and have been introduced and studied in terms of probability theory by Smorodinsky [9] in the dyadic case and by Laurent [5] in the general case.

The distribution of a split-word process depends on an alphabet A of size $N \geqslant 2$ and on a sequence of positive integers $(\ell_n)_{n\leqslant 0}$ such that $\ell_0 = 1$ and, for every $n \leqslant 0$, the ratio $r_n = \ell_{n-1}/\ell_n$ is an integer $r_n \geqslant 2$. The sequence $(X_n)_{n\leqslant 0}$ of split words is indexed by the nonpositive integers. For every $n \leqslant 0$, the law of X_n is uniform on the set of words of length ℓ_n on A. Moreover, if one splits the word X_{n-1}, whose length is ℓ_{n-1}, into r_n subwords of length ℓ_n, then the word X_n is chosen uniformly among these subwords, independently of everything up to time $n-1$. More precisely, denote by V_n the location of the subword X_n in X_{n-1}. Then V_n is uniform in $\{1, 2, \dots, r_n\}$. The split-word process is $(X_n, V_n)_{n\leqslant 0}$.

Call $\mathcal{F}^{X,V} = (\mathcal{F}^{X,V}_n)_{n\leqslant 0}$ the natural filtration of $(X_n, V_n)_{n\leqslant 0}$. Clearly, every subsequence $(\mathcal{F}^{X,V}_n)_{n\in Q}$ with $Q \subset \mathbb{Z}^-$ is the natural filtration of a split-word process with lengths $(\ell_n/\ell_m)_{n\in Q}$ on the alphabet A^{ℓ_m}, where $m = \max Q$.

The filtration $\mathcal{F}^{X,V}$ is (r_n)-adic since for every $n \leqslant 0$,

$$\mathcal{F}^{X,V}_n = \mathcal{F}^{X,V}_{n-1} \vee \sigma(V_n) \text{ with } V_n \text{ independent of } \mathcal{F}^{X,V}_{n-1}.$$

Moreover, the tail σ-field $\mathcal{F}^{X,V}_{-\infty}$ is trivial, thanks to proposition 6.2.1 in [5]. Yet, the inclusion $\mathcal{F}^{V}_n \subset \mathcal{F}^{X,V}_n$ is clearly strict since X_0 is independent of $(V_n)_{n\leqslant 0}$. However, the filtration $\mathcal{F}^{X,V}$ may still be a product type filtration (generated by some other independent sequence).

1.1 Results

Surprisingly, the nature of the filtration $\mathcal{F}^{X,V}$ depends on the sequence $(\ell_n)_{n\leqslant 0}$:

Theorem 1 *The filtration $\mathcal{F}^{X,V}$ is standard (or equivalently, is of product type) if and only if, the series $\sum_n \dfrac{\ln(r_n)}{\ell_n}$ diverges.*

The ‘if’ part of the theorem, which is new, will be proved in Sect. 3, and the ‘only if’ part in Sect. 4.

Note that the convergence of the series $\sum_n \dfrac{\ln(r_n)}{\ell_n}$ is equivalent to the condition Δ of Laurent [5], who established the ‘only if’ part of Theorem 1.

The condition that we call $\neg\Delta$ (the divergence of the series $\sum_n \dfrac{\ln(r_n)}{\ell_n}$), which implies standardness, improves on Laurent’s sufficient condition of standardness:

(∇_N) There exists $\alpha < 1$ such that $r_n^\alpha \gg N^{\ell_n}$ as $n \to -\infty$.

Laurent notices that conditions ∇_N are weaker than the condition

$$\frac{\ln(r_n)}{\ell_n} \to \infty \quad \text{as} \quad n \to -\infty, \qquad (\nabla)$$

which does not depend on N. Laurent indicates that conditions ∇ and Δ were previously introduced by Vershik [13, example 1], in the context of decreasing sequences of measurable partitions.

Since conditions ∇ and Δ do not exhaust all possible situations, both Vershik and Laurent asked what happens "between ∇ and Δ". In Theorem 1, we solve Laurent's question: condition Δ is in fact necessary and sufficient for the filtration of the split-word process to be non standard.

After the completion of this paper, Anatoly Vershik drew our attention to the paper [4], where Heicklen obtains a result equivalent to Theorem 1. The (r_n)-adic filtrations studied by Heicklen were introduced by Vershik [11] as follows. (We mention that both Vershik and Heicklen index sequences by the set of nonnegative integers whereas we index them by the set of nonpositive integers, and that this is the only difference between their presentation and ours, given below.)

Let A be a finite alphabet and $(G_n)_{n\leqslant 0}$ the decreasing sequence of groups defined by

$$G_n = \sum_{k=n+1}^{0} \mathbb{Z}/r_k\mathbb{Z}.$$

Let G denote the union over $n \leqslant 0$ of the groups G_n. For every $n \leqslant 0$, the group G_n acts on A^G (on the left) by canonical shifts. Namely, for every $g \in G_n$ and $f \in A^G$, one defines $g \cdot f \in A^G$ by $(g \cdot f)(x) = f(xg)$ for every $x \in G$. Let $\mathrm{Orb}_{G_n}(f)$ denote the orbit of a given $f \in A^G$ under the action of G_n, that is, $\mathrm{Orb}_{G_n}(f) = \{g \cdot f\ ;\ g \in G_n\}$.

Let $F = (F(x))_{x\in G} \in A^G$ be a random function whose coordinates $F(x)$ are independent and uniformly distributed in A. The filtration $(\mathcal{F}_n)_{n\leqslant 0}$ studied by Vershik and Heicklen is the natural filtration of $(\mathrm{O}_n)_{n\leqslant 0}$, where

$$\mathrm{O}_n = \mathrm{Orb}_{G_n}(F).$$

For each $n \leqslant 0$, O_{n-1} is the union of r_n orbits under the action of G_n, one orbit for each element of G_{n-1}/G_n. Almost surely, these orbits are all different since the shifted functions $g \cdot F$ are different. Futhermore, conditionally on $(\mathrm{O}_{n-1}, \mathrm{O}_{n-2}, ...)$, the random variable O_n is uniformly distributed on these r_n orbits. This shows that $(\mathcal{F}_n)_{n\leqslant 0}$ is an (r_n)-adic filtration.

One can show that the tail σ-field $\mathcal{F}_{-\infty}$ is trivial and that the filtration of the split-words process is immersed in $(\mathcal{F}_n)_{n\leqslant 0}$. Informally, the word X_n at time n is given by the values at e (the identity of the group G) of the elements of O_n. One gets X_n from X_{n-1} by splitting the orbits under G_{n-1} into r_n orbits under G_n and by choosing one of these orbits uniformly randomly.

Thus, the standardness of $(\mathcal{F}_n)_{n\leqslant 0}$ when condition Δ fails implies that the natural filtration of the split-words process is standard and, therefore, that it is of product type. Heicklen's proof relies on Vershik's standardness criterion and uses the language of ergodic theory. Although Heicklen's result and our Theorem 1 are logically equivalent, we believe that our proof is interesting because it relies on a constructive, direct and probabilistic method.

This result has interesting applications in ergodic theory as we now explain. Recall that entropy is a well known invariant associated to an automorphism of a probability space (that is, a bimeasurable application preserving the measure). Vershik [12] defined a much more elaborate invariant, named the scale. But computing this invariant is a very difficult task, even in simple cases. However, Laurent showed that our result provides the exact scale of a dyadic transformation [7]: more precisely, the scale of this transformation is the set of sequences $(r_n)_{n\leqslant 0}$ fulfilling condition Δ.

Using Vershik's theory, one can deduce from Theorem 1 that the filtration of the split-word process on any separable alphabet (endowed with an arbitrary measure) is standard under condition $\neg\Delta$, see [6] for a proof.

1.2 Examples

Condition $\neg\Delta$ forces the length ℓ_n to grow very quickly as n goes to $-\infty$. If a given sequence $(\ell_n)_{n\leqslant 0}$ is $\neg\Delta$, then every sequence $(\ell'_n)_{n\leqslant 0}$ such that $(\ln r'_n)/\ell'_n \geqslant (\ln r_n)/\ell_n$ is $\neg\Delta$ as well. No similar property holds for the sequence $(\ell_n)_{n\leqslant 0}$ only, nor for the sequence $(r_n)_{n\leqslant 0}$ only. Indeed, the first example of this section provides a sequence $(r_n)_n$ which is $\neg\Delta$ and such that $(r_n^2)_{n\leqslant 0}$ is Δ.

Example 1 (A standard split-word filtration) *Let* $\ell_0 = 1$ *and* $\ell_{n-1} = 2^{\ell_n}$ *for every* $n \leqslant 0$. *That is to say*

$$\ell_n = 2^{2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}} \quad \textit{where the fig. } 2 \textit{ appears } |n| \textit{ times.}$$

Then for every $n < 0$,

$$r_n = \ell_{n-1}/\ell_n = 2^{\ell_n - \ell_{n+1}}.$$

Therefore

$$\log_2(r_n)/\ell_n = \frac{\ell_n - \ell_{n+1}}{\ell_n} \to 1,$$

which proves that (r_n) *is* $\neg\Delta$.

Theorem 1 has another interesting consequence which we now explain. Recall that, by the lacunary isomorphism theorem [2], from any filtration $(\mathcal{F}_n)_{n\leqslant 0}$ (such that $\mathcal{F}_0$ is essentially separable), one can extract a filtration $(\mathcal{F}_n)_{n\in Q}$ which is standard. In [13], Vershik provides an example where $(\mathcal{F}_n)_{n\leqslant 0}$ is non standard whereas $(\mathcal{F}_{2n})_{n\leqslant 0}$ is standard. In [3], Gorbulsky also gives such an example. Theorem 1 provides an example of a non standard filtration (example 2 below) in which the transition from the non standard case to the standard case is very sharp: $(\mathcal{F}_n)_{n\in Q}$ is standard for every infinite subset Q of $\mathbb{Z}^-$ with infinite complementary.

Example 2 (A non standard filtration close to standardness) *Set* $\ell_0 = 1$ *and* $\ell_{n-1} = 4^{\sqrt{\ell_n}}$ *for every* $n \leqslant 0$. *That is to say*

$$\ell_n = 4^{2^{2^{\cdot^{\cdot^{\cdot^{2}}}}}} \quad \textit{where the fig. } 2 \textit{ appears } |n| - 1 \textit{ times.}$$

Then the filtration $(\mathcal{F}_n^{(X,V)})_{n\leqslant 0}$ is not of product type. Yet, if ϕ is a strictly increasing application from $-\mathbb{N}$ into $-\mathbb{N}$ such that $\phi(n)-n \to -\infty$ as $n \to -\infty$, then the filtration $(\mathcal{F}_{\phi(n)}^{(X,V)})_{n\leqslant 0}$ is of product type.

Proof On the one hand

$$\frac{\log_2 r_n}{\ell_n} \leqslant \frac{\log_2 \ell_{n-1}}{\ell_n} = \frac{2}{\sqrt{\ell_n}},$$

which is the general term of a convergent series, thus $(\ell_n)_{n\leqslant 0}$ is Δ. On the other hand the filtration $(\mathcal{F}_{\phi(n)}^{(X,V)})_{n\leqslant 0}$ is the filtration of a split-word process of length process $(\ell'_n)_{n\leqslant 0} = (\ell_{\phi(n)})_{n\leqslant 0}$. The ratios between successive lengths are, for $n \leqslant 0$,

$$r'_n = \ell'_{n-1}/\ell'_n = \ell_{\phi(n-1)}/\ell_{\phi(n)}.$$

If $\phi(n)-n \to -\infty$ when $n \to -\infty$, then $\phi(n-1) \leqslant \phi(n)-2$ infinitely often. For these n,

$$r'_n \geqslant \frac{\ell_{\phi(n)-2}}{\ell_{\phi(n)}} = \frac{4^{\sqrt{\ell_{\phi(n)-1}}}}{\ell_{\phi(n)}} = \frac{4^{2^{\sqrt{\ell_{\phi(n)}}}}}{\ell_{\phi(n)}}$$

hence

$$\frac{\log_2 r'_n}{\ell'_n} \geqslant 2\frac{2^{\sqrt{\ell_{\phi(n)}}}}{\ell_{\phi(n)}} - \frac{\log_2 \ell_{\phi(n)}}{\ell_{\phi(n)}}.$$

This shows that a subsequence of $(\log_2(r'_n)/\ell'_n)_n$ converges to infinity, hence that $(\ell'_n)_n$ is $\neg\Delta$. $\square$

2 Laurent's method and tools

In this section, we introduce the tools used by Laurent to prove that under condition ∇, the filtration of the split-word process is of product type. Laurent used a canonical coupling to build explicitly a sequence of innovations $(V'_n)_{n\leqslant 0}$ which generates the process $(X_n, V_n)_{n\leqslant 0}$.

Definition 1 If $(\mathcal{F}_n)_{n\leqslant 0}$ is a filtration, and $(U_n)_{n\leqslant 0}$ is a sequence of random variables such that for every $n \leqslant 0$,

$$\mathcal{F}_n = \mathcal{F}_{n-1} \vee \sigma(U_n) \quad \text{with } U_n \text{ independent of } \mathcal{F}_{n-1},$$

one says that $(U_n)_{n\leqslant 0}$ is a sequence of innovations for $(\mathcal{F}_n)_{n\leqslant 0}$.

This method is strengthened in Sect. 3, where we consider a partial canonical coupling to improve on condition ∇.

We remind the reader that sequences are indexed by the nonpositive integers.

2.1 Change of innovations

We start with a complete definition of split-word processes.

Definition 2 (*Split-word process*) Let $(r_n)_{n\leqslant 0}$ denote a sequence of integers such that $r_n \geqslant 2$ for every $n \leqslant 0$. Set $\ell_0 = 1$ and, for every $n \leqslant 0$, $\ell_{n-1} = r_n \ell_n$. Let A denote a finite set, called the alphabet, with cardinal $N \geqslant 2$.

A split-word process is any process $(X_n, V_n)_{n\leqslant 0}$ such that, for every $n \leqslant 0$,

- X_n is uniformly distributed on A^{ℓ_n},
- V_n is uniformly distributed on $\{1, \ldots, r_n\}$ and independent of the σ-algebra $\mathcal{F}_{n-1}^{(X,V)} = \sigma(X_m, V_m\,;\; m \leqslant n-1)$,
- if the word X_{n-1} (with length $l_{n-1} = l_n r_n$) is partitioned into r_n subwords of length l_n, X_n is the V_nth among those r_n subwords.

The sequence $(X_n)_{n\leqslant 0}$ is a inhomogeneous Markov process indexed by the negative integers and generated by the innovations $(V_n)_{n\leqslant 0}$. The existence of such a process $(X_n, V_n)_{n\leqslant 0}$ is guaranteed by Kolmogorov's theorem.

To prove that, under some conditions, the filtration of the split-word process is of product type, one has to switch from one set of innovations to another. Lemma 2.1 provides a general method to build new innovations.

Lemma 2.1 (Change of innovations) *For every $n \leqslant -1$, let $\{\varphi_w^n\}_w$ denote a family of permutations of $\{1, \ldots, r_{n+1}\}$, indexed by the elements w of A^{ℓ_n}, and let*

$$V'_{n+1} = \varphi^n_{X_n}(V_{n+1}).$$

Then $(V'_n)_{n\leqslant 0}$ is a sequence of generating innovations for $(X_n)_{n\leqslant 0}$. This means that, for every negative integer $n \leqslant -1$, the following properties hold:

- *The random variable V'_{n+1} is uniformly distributed on $\{1, \ldots, r_{n+1}\}$.*
- *The random variable V'_{n+1} is independent of $\mathcal{F}_n^{X,V}$ and therefore also of $\mathcal{F}_n^{X,V'}$.*
- *The random variable X_{n+1} is a measurable function of X_n and V'_{n+1}.*

Proof of lemma 2.1 For every negative integer n and every v such that $1 \leqslant v \leqslant r_{n+1}$, a simple computation proves that

$$\mathbb{P}[V'_{n+1} = v | \mathcal{F}_n^{X,V}] = \mathbb{P}[V_{n+1} = (\varphi^n_{X_n})^{-1}(v) | \mathcal{F}_n^{X,V}] = 1/r_{n+1}.$$

This shows the first two properties. The third property follows from the fact that X_{n+1} is the kth subword of X_n, where $k = (\varphi^n_{X_n})^{-1}(V'_n)$. □

2.2 Canonical word and coupling

To build the innovations which generate the process $(X_n)_{n\leqslant 0}$, one can use, and improve on, Laurent's construction under the stronger condition ∇. This uses the notions of canonical word and canonical coupling, which we recall below.

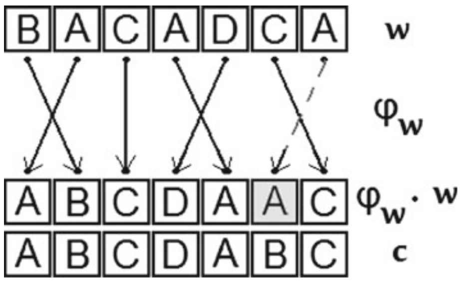

Fig. 1 Example of canonical coupling

Definition 3 (*Canonical alphabets and canonical words*) For every integer $M \geqslant 2$, the canonical alphabet on M letters is $A_M = \{1, \dots, M\}$. Canonical words on A_M are the words whose ith letter is congruent to i modulo M. Hence, the letters of A_M appear in order and are repeated periodically.

Canonical words will usually be denoted by the letter c. For example the canonical word of length 11 on A_3 is 12312312312.

Notation 1 (*General notations*) *To simplify the definition of the canonical coupling, one identifies any ordered alphabet B of size $M \geqslant 2$ with A_M according to the rank of each letter in the alphabet B.*
The ith letter of a word w is denoted by $w(i)$. Let $w = (w(i))_{1 \leqslant i \leqslant r}$ denote a word of length r. For every $1 \leqslant i \leqslant r$, $H(w, i)$ denotes the number of instances of the letter $w(i)$ among the $(i - 1)$ first letters of w:

$$H(w, i) = \sum_{1 \leqslant j < i} \mathbf{1}_{\{w(j) = w(i)\}}.$$

Definition 4 (*Canonical coupling*) Let w denote a word of length r on an ordered alphabet B of size $M \geqslant 2$. The canonical coupling associated to w is the permutation φ_w of $\{1, \dots, r\}$ defined as follows: for every $i \leqslant r$,

$$\varphi_w(i) = w(i) + H(w, i)M \quad \text{if} \quad w(i) + H(w, i)M \leqslant r.$$

After this process has been applied to every i, one chooses $\varphi_w(j)$ for the integers j such that $w(j) + H(w, j)M > r$, in an increasing way and in order to make φ_w a bijection. (So $\varphi_w(j)$ is the smallest k which does not belong yet to the range of φ_w.)

Later on, we apply the notions of canonical word and canonical coupling to some alphabets A^ℓ with $\ell \geqslant 1$ (Fig. 1).

By construction φ_w is one of the permutations φ such that $\varphi \cdot w := w \circ \varphi^{-1}$ is as close as possible to a canonical word. Lemma 2.2 makes this statement more precise.

Lemma 2.2 (Comparison of w and $c \circ \varphi_w$) Let c be the canonical word of length $r \geqslant 1$ on an ordered alphabet B of size $M \geqslant 2$. Then for every $1 \leqslant i \leqslant r$ and $w \in B^r$,

$$c(\varphi_w(i)) = w(i) \qquad \text{if and only if} \qquad w(i) + H(w, i)M \leqslant r.$$

Proof of lemma 2.2 By definition of c, the number of instances of j in c is

$$K_j = \max\{k \geqslant 0 : j + (k-1)M \leqslant r\} = \lfloor \frac{r-j}{M} \rfloor + 1 = \lceil \frac{r+1-j}{M} \rceil.$$

If $w(i) + H(w,i)M \leqslant r$, then $\varphi_w(i) = w(i) + H(w,i)M$, thus $c(\varphi_w(i)) = w(i)$ by definition of c.

Otherwise, $w(i) + H(w,i)M > r$, hence the number $H(w,i)$ of instances of $w(i)$ among the $i-1$ first letters of w is at least the number $K_{w(i)}$ of instances of the letter $w(i)$ in c. According to the first case, φ_w sends the ranks of the $K_{w(i)}$ first instances of $w(i)$ in w on the ranks of the instances of $w(i)$ in c. Since φ_w is bijective, the rank of every instance of the letter $w(i)$ in c has an antecedent by φ_w that is less than i. Therefore the letter in c with rank $\varphi_w(i)$ cannot be $w(i)$, that is, $c(\varphi_w(i)) \neq w(i)$. □

Lemma 2.2 implies Lemma 2.3 below, which is a slight improvement on Laurent's Lemma 6.3.2 of [5].

Lemma 2.3 *Let X denote a uniform random word of length r on an ordered alphabet B of size $M \geqslant 2$, V a uniform random variable on $\{1, \dots, r\}$, independent of X, and c the canonical word of length r on B. Then,*

$$\mathbb{P}[X(V) \neq c(\varphi_X(V))] \leqslant M/r + 2(M/r)^{1/3}.$$

Proof of lemma 2.3 Since $X(V) \leqslant M$ and $\{X(V) = c(\varphi_X(V))\} = \{X(V) + H(X,V)M \leqslant r\}$,

$$\{X(V) = c(\varphi_X(V))\} \supset \{H(X,V)M \leqslant r - M\},$$

hence, for every positive s,

$$\{X(V) = c(\varphi_X(V))\} \supset \{H(X,V)M \leqslant V - 1 + Ms\} \cap \{V - 1 + Ms \leqslant r - M\}.$$

Taking complements, one gets

$$\{X(V) \neq c(\varphi_X(V))\} \subset \{H(X,V)M > V - 1 + Ms\} \cup \{V - 1 + Ms > r - M\},$$

which yields

$$\mathbb{P}[X(V) \neq c(\varphi_X(V))] \leqslant \mathbb{P}[H(X,V)M > V - 1 + Ms] + \mathbb{P}[V - 1 + Ms > r - M].$$

Since $r + 1 - V$ and V are both uniform on $\{1, \dots, r\}$,

$$\mathbb{P}[V - 1 + Ms > r - M] = \mathbb{P}[V < M(s+1)] \leqslant M(s+1)/r.$$

On the other hand,

$$\mathbb{P}[H(X,V)M > V - 1 + Ms] \leqslant \mathbb{P}\left[\left|H(X,V) - (V-1)/M\right| > s\right].$$

Since V and X are independent, conditionally on V the distribution of $H(X, V)$ is binomial $\mathrm{Bin}(V-1, 1/M)$. The conditional expectation of $H(X, V)$ is $(V-1)/M$ and the conditional variance of $H(X, V)$ is

$$(V-1)(M-1)/M^2 \leqslant V/M,$$

hence the Bienaymé-Chebychev inequality yields

$$\mathbb{P}[H(X, V)M > V - 1 + Ms] \leqslant \mathbb{E}[V]/(Ms^2) = (r+1)/(2Ms^2) \leqslant r/(Ms^2).$$

Finally,

$$\mathbb{P}[X(V) \neq c(\varphi_X(V))] \leqslant M(s+1)/r + r/(Ms^2).$$

This upper bound for $s = (r/M)^{2/3}$ implies the statement of the lemma. □

2.3 Adapting Laurent's proof

Let us state a slight improvement on Laurent's result [5, proposition 6.3.3]:

Theorem 2 *If $r_n \gg N^{\ell_n}$ as $n \to -\infty$, then the natural filtration of $(X_n, V_n)_{n \leqslant 0}$ is of product type.*

Let us introduce notations for the proof. From now on, a word of length ℓ_n on the alphabet A will often be seen as a word of length r_{n+1} on the alphabet $A^{\ell_{n+1}}$.
Proof of theorem 2.

- Choice among sub-words: for every x in $A^{\ell_{n-1}}$ and v in $\{1, \ldots, r_n\}$, denote by $f_n(x, v) = x(v)$ the vth letter of x seen as a word of length r_n on the alphabet A^{ℓ_n}. So $X_n = f_n(X_{n-1}, V_n)$.
- New innovations: let c_{n-1} be the canonical word of length r_n on the alphabet A^{ℓ_n} (for some fixed order on A^{ℓ_n}). Let $\varphi_{X_{n-1}}$ be the canonical coupling associated to X_{n-1} seen as a word of length r_n on A^{ℓ_n}. Set $V'_n = \varphi_{X_{n-1}}(V_n)$.
- Construction of a sequence $(X'_n)_{n \leqslant 0}$ approximating $(X_n)_{n \leqslant 0}$: set

$$X'_n = f_n(c_{n-1}, V'_n) = f_n(c_{n-1}, \varphi_{X_{n-1}}(V_n)).$$

The key point is to show that

$$\mathbb{P}[X_n = X'_n] \to 1 \quad \text{as} \quad n \to -\infty$$

by bounding above

$$\mathbb{P}[X_{n-1}(V_n) \neq c_{n-1}(\varphi_{X_{n-1}}(V_n))].$$

Applying Lemma 2.3 to X_{n-1} seen as a word of length r_n on the alphabet A^{ℓ_n}, one gets that,

$$\mathbb{P}[X_n \neq X'_n] \leqslant N^{\ell_n}/r_n + 2(N^{\ell_n}/r_n)^{1/3}.$$

Hence $\mathbb{P}[X_n \neq X'_n] \to 0$ since $N^{\ell_n} \ll r_n$.

Define an application $f'_n : A^{\ell_{n-1}} \times \{1, \ldots, r_n\} \to A^{\ell_n}$ by $f'_n(x, v') = f_n(x, \varphi_x^{-1}(v'))$. Then $f'_n(X_{n-1}, V'_n) = X_n$ and $f'_n(\cdot, V'_n)\circ\cdots\circ f'_{m+1}(\cdot, V'_{m+1})(X_m) = X_n$ for $m \leqslant n \leqslant 0$. Therefore, under the assumption ∇, one has,

$$\begin{aligned}\mathbb{P}\left[X_n \neq f'_n(\ \cdot\ , V'_n) \circ \cdots \circ f'_{m+1}(\ \cdot\ , V'_{m+1})(X'_m)\right] &\leqslant \mathbb{P}[X_m \neq X'_m] \\ &\to 0 \quad \text{as} \quad m \to -\infty.\end{aligned}$$

This implies the convergence in probability:

$$\begin{aligned}X_n &= \lim_{m\to-\infty} f'_n(\ \cdot\ , V'_n) \circ f'_{n-1}(\ \cdot\ , V'_{n-1}) \circ \cdots \circ f'_{m+1}(\ \cdot\ , V'_{m+1})(X'_m) \\ &= \lim_{m\to-\infty} f'_n(\ \cdot\ , V'_n) \circ \cdots \circ f'_{m+1}(\ \cdot\ , V'_{m+1}) \circ f_m(\ \cdot\ , V'_m)(c_{m-1})\end{aligned}$$

and proves that the innovations $(V'_k)_{k\leqslant n}$ determine the words $(X_k)_{k\leqslant n}$ and the innovations $(V_k = \varphi^{-1}_{X_{k-1}}(V'_k))_{k\leqslant n}$. Thus, the filtration $(\mathcal{F}_n^{(X,V)})_{n\leqslant 0}$ is generated by the innovations $(V'_n)_{n\leqslant 0}$. □

3 Improving on condition ∇

3.1 Statement of the main result

As said before, there is a gap between the conditions ∇ and Δ under which the problem has been solved by Laurent. Our next theorem bridges the gap between the two conditions.

Theorem 3 *If the series* $\sum_n \ln(r_n)/\ell_n$ *diverges (condition* $\neg\Delta$*), then the filtration* $(\mathcal{F}_n^{(X,V)})_{n\leqslant 0}$ *is of product type.*

Remark 1 *Laurent* [5] *states condition* Δ *as the convergence of the series*

$$\sum_n \frac{\ln(r_n!)}{\ell_{n-1}}.$$

The inequalities $\frac{1}{2}r\ln(r) \leqslant \ln(r!) \leqslant r\ln(r)$*, valid for every* $r \geqslant 2$*, ensure that the series*

$$\sum_n \frac{\ln(r_n)}{\ell_n} \quad \textit{and} \quad \sum_n \frac{\ln(r_n!)}{\ell_{n-1}}$$

both converge or both diverge. Hence Laurent's condition Δ *and the condition* Δ *which we stated in our introduction and used since, are indeed equivalent.*

Condition $\neg\Delta$ is easy to express, but less handy to prove things. The equivalent wording below, which is closer to Theorem 2, is more convenient.

Proposition 3.1 (Rewording of condition $\neg\Delta$) *Condition* Δ *fails if and only if there exists a sequence* $(\alpha_n)_{n\leqslant 0}$ *of nonnegative real numbers and an increasing application* $\phi : -\mathbb{N} \to -\mathbb{N}$*, such that the following properties hold:*

1. *For every* $n \leqslant 0$, $r_{\phi(n)} \geqslant N^{2\alpha_n \ell_{\phi(n)}}$.
2. *The series* $\sum_n \alpha_n$ *diverges.*

Furthermore, when they exist, the sequence $(\alpha_n)_{n\leqslant 0}$ *and the application* ϕ *can be chosen in such a way that the additional properties below hold:*

3. *When* $n \to -\infty$, $r_{\phi(n)} \gg N^{2\alpha_n \ell_{\phi(n)}}$ *(in particular,* $r_{\phi(n)} \to +\infty$*).*
4. *The series* $\sum_n \alpha_{2n}$ *diverges.*
5. *For every* $n \leqslant 0$, $0 < \alpha_n \leqslant 1$.
6. *For every* $n \leqslant -2$*, the ratio* $\alpha_n \ell_{\phi(n)}/\ell_{\phi(n+1)}$ *is an integer.*

The proof of this result can be found in Sect. 3.4.

3.2 Construction of the new innovations

The construction of the new innovations uses a *partial canonical coupling*. This tool sharpens the canonical coupling that was introduced in Sect. 2.2 and which has to be kept in mind.

Under the hypotheses of Theorem 3, the ratios $(r_n)_{n\leqslant 0}$ are no longer big enough for the innovations associated to the canonical coupling to approach the entire word X_n in only one step. Therefore several steps are necessary to get a good information on the word.

Definition 5 (*Partial canonical coupling*) Let w be a word of length ℓr on the alphabet A and $\lambda \in \{1, \ldots, \ell\}$ an integer. Denote by $\tilde{w}$ the word extracted from w by splitting w into r sub-words of length ℓ and keeping only the first λ letters of each sub-word. In other words, if $w = (w_1, \ldots, w_{\ell r})$, then

$$\tilde{w} = (w_1, \ldots, w_\lambda, w_{\ell+1}, \ldots, w_{\ell+\lambda}, \ldots, w_{(r-1)\ell+1}, \ldots, w_{(r-1)\ell+\lambda})$$

is the word constituted of the letters w_i such that $i = j \bmod \ell$ with $1 \leqslant j \leqslant \lambda$.

Let $\varphi_{\tilde{w}}$ be the canonical coupling of $\tilde{w}$ towards the canonical word $\tilde{c}$ of length r on the alphabet A^λ. Since $\varphi_{\tilde{w}}$ belongs to $\mathfrak{S}_r$ (the symmetric group on r letters), one can apply it to the r sub-words of w of length ℓ. The permutation $\varphi_w^{\lambda/\ell} = \varphi_{\tilde{w}}$ is called the partial canonical coupling of rate λ/ℓ associated to w (Fig. 2).

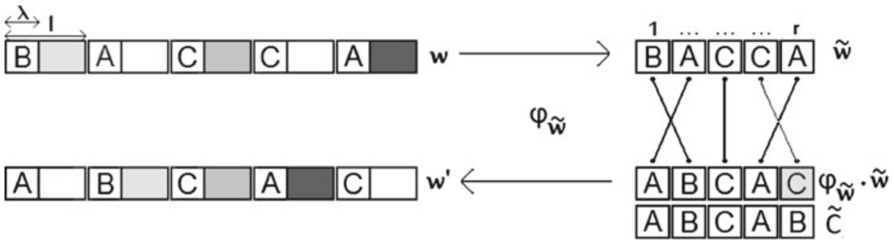

Fig. 2 Example of a partial canonical coupling. In the proof, this coupling will be considered with $\ell = \ell_{\phi(2n)}$, $\lambda = \alpha_{2n}$, $r = r_{\phi(2n)}$, $w = X_{\phi(2n)-1}$ and $\tilde{w} = \tilde{X}_{\phi(2n)-1}$.

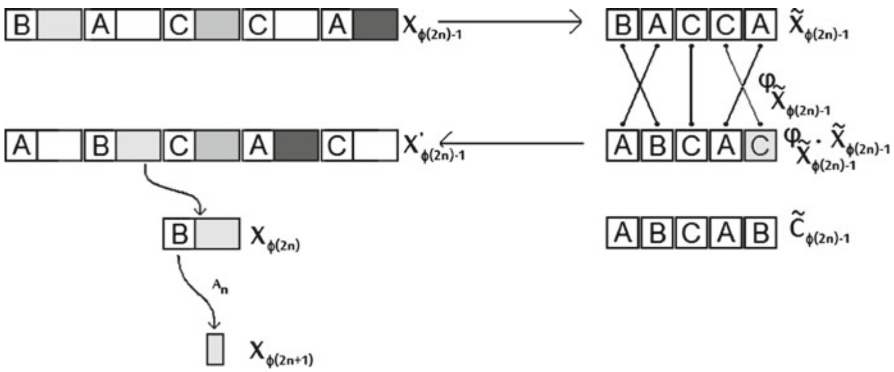

Fig. 3 Description of the method.

Proof of theorem 3. Assume that $\neg\Delta$ holds. Fix ϕ and $(\alpha_n)_{n\leqslant 0}$ which fulfill conditions 1,2,3,4,5 and 6 of proposition 3.1. We now construct new innovations $(V'_n)_{n\leqslant 0}$ which generate the same filtration than the process X.

Let us define new innovations $(V'_k)_{k\leqslant 0}$ as follows. For every $k \leqslant 0$, define $V'_k = \varphi^{\alpha_{2n}}_{X_{k-1}}(V_k) = \varphi_{\tilde{X}_{k-1}}(V_k)$ if there exists an integer n (necessarily unique) such that $k = \phi(2n)$, and $V'_k = V_k$ otherwise.

Lemma 2.3 will be used to show that with probability close to 1, the first $\alpha_{2n}\ell_{\phi(2n)}$ letters of the words $X_{\phi(2n)}$ and those of $f_{\phi(2n)}(C_{\phi(2n)-1}, V'_{\phi(2n)})$ coincide (Fig. 3).

3.3 Proof of the main result

This proof is split into three steps.

First step: $X_{\phi(2n+1)}$ comes from the beginning of $X_{\phi(2n)}$ infinitely often

One focuses on the events

$$A_n = \{X_{\phi(2n+1)} \text{ comes from the first } \alpha_{2n}\ell_{\phi(2n)} \text{ letters of } X_{\phi(2n)}\}.$$

One computes $\mathbb{P}[A_n]$ by counting: the number of possible choices for the innovations V_k for $\phi(2n)+1 \leqslant k \leqslant \phi(2n+1)$ is

$$\prod_{k=\phi(2n)+1}^{\phi(2n+1)} r_k = \ell_{\phi(2n)}/\ell_{\phi(2n+1)}.$$

The number of cases such that A_n occurs is the number of sub-words of length $\ell_{\phi(2n+1)}$ entirely included in the first $\alpha_{2n}\ell_{\phi(2n)}$ letters of $X_{\phi(2n)}$: this number is $\alpha_{2n}\ell_{\phi(2n)}/\ell_{\phi(2n+1)}$ thanks to the additional hypothesis that $\alpha_{2n}\ell_{\phi(2n)}/\ell_{\phi(2n+1)}$ is an integer. Therefore $\mathbb{P}[A_n] = \alpha_{2n}$ and the series $\sum_n \mathbb{P}[A_n]$ diverges.

Moreover A_n is a (deterministic) function of V_k for $\phi(2n)+1 \leqslant k \leqslant \phi(2n+1)$, hence the events A_n are independent and the Borel-Cantelli lemma ensures that almost surely, A_n occurs for infinitely many n.

Note that $A_n \in \mathcal{F}^{V'}_{\phi(2n+1)}$ for every n since $V'_k = V_k$ for every time k which is not one of the integers $\phi(2n)$.

Second step: Use of lemma 2.3

Our purpose is to prove Lemma 3.2.

Lemma 3.2 *For every $n \leqslant 0$, set $I_{\phi(2n)} = \{1, \ldots, \alpha_{2n}\ell_{\phi(2n)}\}$ and fix a word $C_{\phi(2n)-1}$ of length $\ell_{\phi(2n)-1} = r_{\phi(2n)}\ell_{\phi(2n)}$ on A such that $\tilde{C}_{\phi(2n)-1}$ is the canonical word of length $r_{\phi(2n)}$ on the alphabet $A^{\alpha_{2n}\ell_{\phi(2n)}}$. The probability for $X_{\phi(2n)}(I_{\phi(2n)})$ (the first $\alpha_{2n}\ell_{\phi(2n)}$ letters of $X_{\phi(2n)}$) to be the $V'_{\phi(2n)}$th sub-word of $\tilde{C}_{\phi(2n)-1}$ converges to 1 as n tends towards $-\infty$. That is to say,*

$$\mathbb{P}\Big[X_{\phi(2n)}(I_{\phi(2n)}) = f_{\phi(2n)}(C_{\phi(2n)-1}, V'_{\phi(2n)})(I_{\phi(2n)})\Big] \to 1, \quad \textit{as } n \to -\infty.$$

Proof of lemma 3.2 Note that $X_{\phi(2n)}(J_{\phi(2n)})$ is the $V_{\phi(2n)}$th letter of $\tilde{X}_{\phi(2n)-1}$ seen as a word of length $r_{\phi(2n)}$ on the alphabet $A^{\alpha_{2n}\ell_{\phi(2n)}}$ (where $\tilde{X}_{\phi(2n)-1}$ is built from $X_{\phi(2n)-1}$ according to Definition 5 and the caption of figure 3.2). Hence

$$\begin{aligned}\mathbb{P}\Big[X_{\phi(2n)}(I_{\phi(2n)}) &= f_{\phi(2n)}(C_{\phi(2n)-1}, V'_{\phi(2n)})(I_{\phi(2n)})\Big] \\ &= \mathbb{P}\Big[\tilde{X}_{\phi(2n)-1}(V_{\phi(2n)}) = \tilde{C}_{\phi(2n)-1}(\varphi_{\tilde{X}_{\phi(2n)-1}}(V_{\phi(2n)}))\Big]\end{aligned}$$

Lemma 2.3 applied to $\tilde{X}_{\phi(2n)-1}$ seen as a word of length $r_{\phi(2n)}$ on the alphabet $A^{\alpha_{2n}\ell_{\phi(2n)}}$ provides

$$\begin{aligned}&\mathbb{P}\Big[\tilde{X}_{\phi(2n)-1}(V_{\phi(2n)}) \neq f_{\phi(2n)}(C_{\phi(2n)-1}, V'_{\phi(2n)})(I_{\phi(2n)})\Big] \\ &\leqslant N^{\alpha_{2n}\ell_{\phi(2n)}}/r_{\phi(2n)} + 2(N^{\alpha_{2n}\ell_{\phi(2n)}}/r_{\phi(2n)})^{1/3}.\end{aligned}$$

From proposition 3.1, each term converges to 0, hence lemma 3.2 holds.

Third step: Use of the innovations to recover $(X_n)_{n\leqslant 0}$

Our aim is to show that, for every $m \leqslant 0$, X_m is a function of the innovations $(V_k)_{k\leqslant m}$. Consider once again the events

$$A_n = \{X_{\phi(2n+1)} \text{ comes from the first } \alpha_{2n}\ell_{\phi(2n)} \text{ letters of } X_{\phi(2n)}\}.$$

If the event A_n occurs and if

$$X_{\phi(2n)}(I_{\phi(2n)}) = f_{\phi(2n)}(C_{\phi(2n)-1}, V'_{\phi(2n)})(I_{\phi(2n)}),$$

then

$$X_{\phi(2n+1)} = f_{\phi(2n+1)}(\cdot, V'_{\phi(2n+1)}) \circ \cdots \circ f_{\phi(2n)}(\cdot, V'_{\phi(2n)})(C_{\phi(2n)-1}).$$

Moreover, since A_n depends only on $(V'_{\phi(2n)+1}, \ldots, V'_{\phi(2n+1)})$, it is independent of $\{X_{\phi(2n)}(I_{\phi(2n)}) = f_{\phi(2n)}(C_{\phi(2n)-1}, V'_{\phi(2n)})(I_{\phi(2n)})\}$. Thus

$$\begin{aligned}&\mathbb{P}[X_{\phi(2n+1)} = f_{\phi(2n+1)}(\cdot, V'_{\phi(2n+1)}) \circ \cdots \circ f_{\phi(2n)}(\cdot, V'_{\phi(2n)})(C_{\phi(2n)-1}) \mid A_n]\\ &\quad \geq \mathbb{P}[X_{\phi(2n)}(I_{\phi(2n)}) = f_{\phi(2n)}(C_{\phi(2n)-1}, V'_{\phi(2n)})(I_{\phi(2n)})],\end{aligned}$$

which tends to 1 as n goes to $-\infty$ by lemma 3.2.

Some formulas below will be easier to read thanks to the introduction of the function g_n which associates X_{n+1} to (X_n, V'_{n+1}). Namely, for every integer $n \leqslant 0$, every word x in A^{ℓ_n} and every integer $1 \leqslant v \leqslant r_n$, define

$$\begin{aligned} g_n(x, v) &= f_n(x, \varphi_{\tilde{x}}(v)) \quad \text{if } n \text{ is one of the integers } \phi(2k),\\ &= f_n(x, v) \quad \text{otherwise.}\end{aligned}$$

Let $X'_{\phi(2n)}$ be the word of length $\ell_{\phi(2n)}$ whose first $\alpha_{2n}\ell_{\phi(2n)}$ letters are those of the word $f_{\phi(2n)}(C_{\phi(2n)-1}, V'_{\phi(2n)})$ and the others are set to 1. Then, by lemma 3.2,

$$\mathbb{P}\left[X_{\phi(2n)}(I_{\phi(2n)}) = X'_{\phi(2n)}(I_{\phi(2n)})\right] \to 1 \quad \text{as } n \to -\infty.$$

For $n < m \leqslant 0$, call $X'_{m,n}$ the offspring of $X'_{\phi(2n)}$ at time m, that is,

$$X'_{m,n} = g_{m-1}(\cdot, V'_m) \circ g_{m-2}(\cdot, V'_{m-1}) \circ \cdots \circ g_{\phi(2n)+1}(\cdot, V'_{\phi(2n)})(X'_{\phi(2n)-1}).$$

Then,

$$\mathbb{P}\left[X_m \neq X'_{m,n}|A_n\right] \leqslant \mathbb{P}\left[X_{\phi(2n)}(I_{\phi(2n)}) \neq X'_{\phi(2n)}(I_{\phi(2n)})|A_n\right],$$

hence $\mathbb{P}\left[X_m \neq X'_{m,n}|A_n\right] \to 0$, as n goes to $-\infty$.

Lemma 3.3 below, which will be proved at the end of Sect. 3.4, enables us to use Borel-Cantelli's lemma twice.

Lemma 3.3 *Let $(a_n)_{n\geqslant 0}$ and $(b_n)_{n\geqslant 0}$ 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 application $\theta : \mathbb{N} \to \mathbb{N}$ such that the series $\sum_n a_{\theta(n)}$ converges and the series $\sum_n b_{\theta(n)}$ diverges.*

Continuation of the proof of theorem 3
Since the series $\sum_n \mathbb{P}(A_n)$ diverges and, for every fixed $m \leqslant 0$,

$$\mathbb{P}[X_m \neq X'_{m,n} | A_n] \to 0, \text{ when } n \to -\infty,$$

our last lemma applied to the sequences $(\mathbb{P}[A_n])_{n\leqslant m}$ and $(\mathbb{P}[X_m \neq X'_{m,n}; A_n])_{n\leqslant m}$ provides a deterministic increasing application $\theta : -\mathbb{N} \to -\mathbb{N}$ such that

$$\sum_{n\leqslant m} \mathbb{P}\left[X_m \neq X'_{m,\theta(n)}; A_{\theta(n)}\right] < \infty \quad \text{and} \quad \sum_{n\leqslant m} \mathbb{P}\left[A_{\theta(n)}\right] = \infty.$$

By Borel-Cantelli's lemma, the events $\{X_m \neq X'_{m,\theta(n)}\} \cap A_{\theta(n)}$ occur only for a finite number of times n, whereas the independent events $A_{\theta(n)}$ occur infinitely often. Thus for every word x in A^{l_m}, almost surely,

$$\{X_m = x\} = \limsup_{n\to-\infty} A_{\theta(n)} \cap \{X'_{m,\theta(n)} = x\}.$$

This proves that $\{X_m = x\}$ belongs to $\mathcal{F}_m^{V'}$, hence X_m is a function of the innovations $(V'_m)_{m\leqslant 0}$. □

3.4 Proof of some auxiliary facts

Proof of proposition 3.1 Assume that $(\alpha_n)_{n\leqslant 0}$ and ϕ exist such that 1 and 2 hold. Then, for every $n \leqslant 0$, $r_{\phi(n)} \geqslant N^{2\alpha_n \ell_{\phi(n)}}$ hence $\log_N r_{\phi(n)} \geqslant 2\alpha_n \ell_{\phi(n)}$. Therefore

$$\sum_n \frac{\log_N r_n}{\ell_n} \geqslant \sum_n \frac{\log_N r_{\phi(n)}}{\ell_{\phi(n)}} \geqslant 2 \sum_n \alpha_n,$$

and the last series diverges hence condition $\neg\Delta$ holds.

Conversely, assume that condition $\neg\Delta$ holds. Let $\beta_n = \frac{1}{4}\log_N(r_n)/\ell_n$. Then, $r_n = N^{4\beta_n \ell_n}$ hence $r_n/N^{2\beta_n\ell_n} = N^{2\beta_n\ell_n}$.

Since $\sum_{n\leqslant 0} \beta_n$ diverges and $\sum_{n\leqslant 0} |n|2^n$ converges, there exists an increasing application ϕ such that the series $\sum_n \beta_{\phi(n)}$ diverges and such that $\beta_{\phi(n)} \geqslant 2^{\phi(n)}|\phi(n)|$ for every $n \leqslant 0$.

Replacing, if necessary, ϕ by φ given by $\varphi(n) = \phi(n-1)$, one can ensure that the series $\sum_n \beta_{\phi(2n)}$ diverges as well.

Since $r_n \geqslant 2$ for every n, $\ell_{\phi(n)} \geqslant 2^{-\phi(n)}$, hence $\beta_{\phi(n)} \geqslant 2^{\phi(n)}|\phi(n)|$ implies that $\beta_{\phi(n)}\ell_{\phi(n)} \geqslant |\phi(n)| \geqslant |n|$. Hence,

$$r_{\phi(n)}/N^{2\beta_{\phi(n)}\ell_{\phi(n)}} = N^{2\beta_{\phi(n)}\ell_{\phi(n)}} \geqslant N^{2|n|} \geqslant 1,$$

and this sequence converges to $+\infty$. Furthermore, defining $\alpha_n = \min(\beta_{\phi(n)}, 1)$, one sees that $(\alpha_n)_n$ fulfills conditions 1-2-4-5 (condition 3 being a consequence of condition 4).

We now show how to build from $(\alpha_n)_n$ a sequence $(\alpha'_n)_n$ such that condition 6 holds as well.

Recall that from the construction above, $r_n = N^{4\beta_n\ell_n}$ and $\beta_{\phi(n)}\ell_{\phi(n)} \geqslant |\phi(n)|$, hence $r_{\phi(n)} \geqslant N^{4|\phi(n)|} \geqslant 2^{-4\phi(n)}$. Also $\alpha_n \geqslant 2^{\phi(n)}|\phi(n)|$.

These inequalities implies that the ratio $\varrho_n = \alpha_n\ell_{\phi(n)}/\ell_{\phi(n+1)}$ is such that

$$\begin{aligned}
\varrho_n &= \alpha_n \left(\prod_{k=\phi(n)+1}^{\phi(n+1)-1} r_k\right) r_{\phi(n+1)} \\
&\geqslant 2^{\phi(n)}|\phi(n)|\, 2^{\phi(n+1)-\phi(n)-1}2^{-4\phi(n+1)} \\
&= 2^{-3\phi(n+1)-1}|\phi(n)|.
\end{aligned}$$

Since $\phi(n+1) \leqslant -1$ and $|\phi(n)| \geqslant 2$ for every $n \leqslant -2$, this shows that $\varrho_n \geqslant 8$ for every $n \leqslant -2$.

Thus, $8\varrho_n/9 \leqslant \lfloor \varrho_n \rfloor \leqslant \varrho_n$ and the sequence $(\alpha'_n)_n$ defined by

$$\alpha'_n = \frac{\lfloor \alpha_n\ell_{\phi(n)}/\ell_{\phi(n+1)} \rfloor}{\ell_{\phi(n)}/\ell_{\phi(n+1)}},$$

is such that $\frac{8}{9}\alpha_n \leqslant \alpha'_n \leqslant \alpha_n$ for every $n \leqslant -2$. Therefore it fulfills the conditions already satisfied by the sequence $(\alpha_n)_{n\leqslant 0}$ and the additional condition that $\alpha'_n\ell_{\phi(n)}/\ell_{\phi(n+1)}$ is an integer for every $n \leqslant 0$. □

Proof of lemma 3.3 Call B any finite upper bound of the sequence $(b_n)_{n\in\mathbb{N}}$. Since $a_n \ll b_n$, for any positive integer k, there exists an integer N_k such that for every $n \geqslant N_k, a_n \leqslant b_n 2^{-k}$. We now define a sequence of disjoint intervals of integers $J_k = \{i_k, \ldots, j_k\}$, as follows. Set $j_{-1} = -1$ and let $k \geqslant 0$.

Once j_{k-1} is defined, let $i_k = \max\{j_{k-1}+1, N_k\}$. Since the series $\sum_n b_n$ diverges and $0 \leqslant b_n \leqslant B$ for every $n \in \mathbb{N}$, one can choose an integer $j_k \geqslant i_k$ such that $B \leqslant \sum_{n=i_k}^{j_k} b_n < 2B$ and let $I_k = \{i_k, \ldots, j_k\}$. Note that $\sum_{n\in J_k} b_n \geqslant B$ and $\sum_{n\in J_k} a_n \leqslant 2B/2^k$.

Calling Q the set $\bigcup_{k\in\mathbb{N}} I_k$, one gets

$$\sum_{n\in Q} a_n = \sum_{k\in\mathbb{N}} \sum_{n\in J_k} a_n \leqslant 2B \sum_{k\in\mathbb{N}} 1/2^k = 4B,$$

and

$$\sum_{n\in Q} b_n = \sum_{k\in\mathbb{N}} \sum_{n\in J_k} b_n \geqslant \sum_{k\in\mathbb{N}} B = \infty,$$

which completes the proof of the lemma. □

4 Proof of non standardness under Δ

The non standardness of the split-word process was established by Laurent in his thesis [5] and a similar result was obtained by Vershik in [13] in the context of decreasing sequences of measurable partitions. In this section we give a simplified presentation of Laurent's proof.

The proof of this result involves a subtle notion on filtrations, which is standardness. The notion of standardness was first introduced by Vershik for decreasing measurable partitions. This notion has been adapted to continuous time filtrations by Tsirelson [10], it has been formulated by Dubins, Feldman, Smorodinsky and Tsirelson [1] for continuous time filtrations and by Emery and Schachermayer [2] for discrete time filtrations.

Many necessary and sufficient conditions for standardness have been established, for instance Vershik's self-joining criterion and various notions of cosiness. All these criterions are based on coupling methods. Checking them in specific cases is often a technical task. Yet, these criterions are the key tool to solve some difficult problems. For example, Tsirelson defines and uses a notion of cosiness to prove that the filtration of Walsh's Brownian motion is not Brownian since it is non standard.

By definition, a filtration $\mathcal{F} = (\mathcal{F}_n)_{n\leqslant 0}$ indexed by $n \leqslant 0$ is *standard* if, modulo an enlargement of the probability space, one can immerse $\mathcal{F}$ in a filtration generated by an i.i.d. process. Recall that a filtration $\mathcal{F} = (\mathcal{F}_n)_{n\leqslant 0}$ is *immersed* in a filtration $\mathcal{G} = (\mathcal{G}_n)_{n\leqslant 0}$ if, for every $n \leqslant 0$, $\mathcal{F}_n \subset \mathcal{G}_n$ and $\mathcal{F}_n$ and $\mathcal{G}_{n-1}$ are independent conditionally on $\mathcal{F}_{n-1}$. Roughly speaking, this means that $\mathcal{G}_{n-1}$ gives no further information on $\mathcal{F}_n$ than $\mathcal{F}_{n-1}$ does. Equivalently, $\mathcal{F}$ is immersed in $\mathcal{G}$ if every $\mathcal{F}$-martingale is a $\mathcal{G}$-martingale.

Laurent negates the so-called I-cosiness property to prove that under Δ, the filtration is non standard and therefore non of product type.

We follow the method that Smorodinsky [9] used to prove the non-existence of a "generating parametrization" in the case where $r_n = 2$ for every $n \leqslant 0$. This method still works in the general case and provides the non standard behaviour of the filtration.

The purpose of this section is to show that if the sequence $(r_n)_{n\leqslant 0}$ is Δ, then the filtration $\mathcal{F}^{(X,V)}$ is non standard. Note that Δ holds for every bounded sequence $(r_n)_n$.

4.1 Preliminary notions

We first recall the notion of I-cosiness, due to Émery and Schachermayer [2].

Definition 6 (*Immersion and co-immersion*) Let $\mathcal{F} = (\mathcal{F}_n)_{n\leqslant 0}$ and $\mathcal{G} = (\mathcal{G}_n)_{n\leqslant 0}$ denote two filtrations defined on the same probability space.

- $\mathcal{F}$ is immersed in $\mathcal{G}$ if every martingale in $\mathcal{F}$ is a martingale in $\mathcal{G}$.
- $\mathcal{F}$ and $\mathcal{G}$ are co-immersed if $\mathcal{F}$ and $\mathcal{G}$ are both immersed in $\mathcal{F} \vee \mathcal{G}$.

Definition 7 (*I-cosiness*) Let $\mathcal{F}$ be a filtration on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. One says that $\mathcal{F}$ satisfies the I-cosiness criterion if for every random variable Y measurable for $\mathcal{F}_0$ with values in a finite set and for every real $\delta > 0$, there exists a probability space $(\overline{\Omega}, \overline{\mathcal{A}}, \overline{\mathbb{P}})$ and two filtrations $\mathcal{F}'$ and $\mathcal{F}''$ on $(\overline{\Omega}, \overline{\mathcal{A}}, \overline{\mathbb{P}})$, such that the following properties hold.

- The filtrations $\mathcal{F}'$ and $\mathcal{F}''$ are both isomorphic to $\mathcal{F}$.
- The filtrations $\mathcal{F}'$ and $\mathcal{F}''$ are co-immersed.
- There exists an integer n_0 such that $\mathcal{F}'_{n_0}$ and $\mathcal{F}''_{n_0}$ are independent.
- The copies Y' and Y'' of Y by the isomorphisms of the first condition, verify $\overline{\mathbb{P}}[Y' \neq Y''] < \delta$.

The proof of the non standardness of the filtration $\mathcal{F}^{(X,V)}$ uses the easy part of the equivalence between I-cosiness and standardness.

Theorem 4 (Corollary 5 [2]) *A filtration is standard if and only if it satisfies the I-cosiness criterion and is essentially separable.*

To prove the non standardness of $\mathcal{F}^{(X,V)}$, it is therefore sufficient to show that $\mathcal{F}^{(X,V)}$ does not satisfies the I-cosiness criterion. The tools of the proof are introduced just below.

Let n_0 be a negative integer which will be fixed later.

Definition 8 (Definition of Aut_n) The intervals of integers $\{1, \ldots, \ell_k\}$, $\{\ell_k + 1, \ldots, 2\ell_k\}, \ldots, \{\ell_n - \ell_k + 1, \ldots, \ell_n\}$ are called blocks of $\{1, \ldots, \ell_n\}$ of length ℓ_k. Every permutation of the ℓ_n/ℓ_{n_0} blocks of length ℓ_{n_0} which induces for every $k \in \{n, \ldots, n_0\}$ a bijection between the blocks of length ℓ_k is called an automorphism of $\{1, \ldots, \ell_n\}$ adapted to $\{r_{n+1}, ..., r_{n_0}\}$. One denotes by $\mathrm{Aut_n}$ the set of those permutations.

One can enumerate the automorphisms adapted to $(r_n)_{n\leqslant 0}$, by induction, as follows. By definition, an automorphism a in Aut_{n-1} is built from a permutation σ of $\{1, \ldots, r_n\}$ and from r_n automorphisms $(a_k)_{1\leqslant k\leqslant r_n}$ in Aut_n. One gets a from σ and $(a_k)_{1\leqslant k\leqslant r_n}$ by setting, for every $1 \leqslant j \leqslant r_n$ and every $1 \leqslant k \leqslant \ell_n$,

$$a((j-1)\ell_n + k) = (\sigma(j) - 1)\ell_n + a_j(k).$$

Therefore $\#(\mathrm{Aut}_{n-1}) = \#(\mathfrak{S}_{r_n})(\#(\mathrm{Aut}_n))^{r_n} = r_n!(\#(\mathrm{Aut}_n))^{r_n}$. By induction

$$\#(\mathrm{Aut}_n) = \prod_{k=n+1}^{n_0} (r_k!)^{r_{n+1}\ldots r_{k-1}} = \prod_{k=n+1}^{n_0} (r_k!)^{\ell_n/\ell_{k-1}}.$$

Note that

$$\#(\mathrm{Aut}_n) = \exp(\ell_n S_n) \text{ where } S_n = \sum_{k=n+1}^{n_0} \frac{\ln r_k!}{\ell_{k-1}}.$$

We now define a semi-metrics based on the Hamming distance.

Definition 9 (*Semi-metrics e_n on A^{ℓ_n}*) Recall that the Hamming distance on A^{ℓ_n} is defined by

$$d_n^H(x, x') = \#\{k \in \{1, \dots, \ell_n\} \;:\; x(k) \neq x'(k)\}.$$

One defines an action of the group Aut_n on A^{ℓ_n}, seen as $(A^{\ell_{n_0}})^{\ell_n/\ell_{n_0}}$, by

$$a \cdot x = x \circ a^{-1},$$

and a semi-metrics e_n on A^{ℓ_n}, by

$$e_n(x, x') = \frac{1}{\ell_n} \min\{d_n^H(a \cdot x, x')\,;\; a \in \mathrm{Aut}_n\}.$$

The quantity $e_n(x, x')$ is the smallest proportion of letters which are different between the words x' and $x \circ a^{-1}$ as a goes through Aut_n. Our next result is a recursion relation, useful to compute e_n.

Lemma 4.1 *For every $n \leqslant 0$, $x = (w_1, \dots, w_{r_n})$ and $x' = (w'_1, \dots, w'_{r_n})$ where w_i and w'_j belong to A^{ℓ_n},*

$$e_{n-1}(x, x') = \min_{\sigma \in \mathfrak{S}_{r_{n-1}}} \left(\frac{1}{r_n} \sum_{j=1}^{r_n} e_n(w_j, w'_{\sigma(j)}) \right).$$

Proof of lemma 4.1 Using the decomposition of every automorphism a of Aut_{n-1} into a permutation σ of $\{1, \dots, r_n\}$ and r_n automorphisms a_k, $1 \leqslant k \leqslant r_n$ in Aut_n, and the additivity of the restricted Hamming distance, one sees that $\ell_{n-1} e_{n-1}(x, x')$ is the minimum over these σ and a_k of the sums

$$\sum_{j=1}^{r_n} d_n^H(a_j \cdot w_j, w'_{\sigma(j)}),$$

hence

$$\ell_{n-1} e_{n-1}(x, x') = \min_{\sigma \in \mathfrak{S}_{r_n}} \left(\ell_n \sum_{j=1}^{r_n} e_n(w_j, w'_{\sigma(j)}) \right).$$

$\square$

To prove that the filtration $\mathcal{F}^{(X,V)}$ is non standard under the assumption Δ, by denying the I-cosiness criterion, one considers X' and X'' two copies of $(X_n)_{n\leqslant 0}$ such that:

- The associated filtrations $\mathcal{F}' = \mathcal{F}^{(X')}$ and $\mathcal{F}'' = \mathcal{F}^{(X'')}$ are co-immersed;
- There exists an integer $m < n_0$ such that $(X'_k)_{k\leqslant m}$ and $(X''_k)_{k\leqslant m}$ are independent.

Our proof of the non standardness of $\mathcal{F}^{(X,V)}$ includes three steps:

- We prove the inequality $\mathbb{P}[X'_{n_0} \neq X''_{n_0}] \geqslant E[e_n(X'_n, X''_n)]$.
- We bound below $E[e_n(X'_n, X''_n)]$ when X'_n and X''_n are independent.
- We negate the I-cosiness criterion.

4.2 Proof of the inequality $\mathbb{P}[X'_{n_0} \neq X''_{n_0}] \geqslant E[e_n(X'_n, X''_n)]$ for $n \leqslant n_0$

For $n \leqslant n_0$ one denotes by

$$M_n = \mathbb{P}[X'_{n_0} \neq X''_{n_0} | \mathcal{F}'_n \vee \mathcal{F}''_n], \qquad L_n = e_n(X'_n, X''_n).$$

By construction, $(M_n)_{n\leqslant n_0}$ is a martingale. Loosely speaking, this martingale measures the influence of the past before time n on the word at time n_0. The key step is to prove that $(M_n)_{n\leqslant n_0}$ is bounded below by $(L_n)_{n\leqslant n_0}$.

Let us prove that $(L_n)_{n\leqslant n_0}$ is a sub-martingale in the filtration $\mathcal{F}' \vee \mathcal{F}''$. Let $n \leqslant n_0$. Using the co-immersion of the filtrations $\mathcal{F}'$ and $\mathcal{F}''$, the conditional law of X'_n given $\mathcal{F}'_{n-1} \vee \mathcal{F}''_{n-1}$ is uniform on the r_n sub-words of X'_{n-1} of length ℓ_n:

$$\mathcal{L}(X'_n | \mathcal{F}'_{n-1} \vee \mathcal{F}''_{n-1}) = \frac{1}{r_n} \sum_{i=1}^{r_n} \delta_{f_n(X'_{n-1}, i)},$$

where $f_n(x, v)$ denotes the vth sub-word (of length ℓ_n) of x, for $x \in A^{\ell_{n-1}}$ and $1 \leqslant v \leqslant r_n$ (as in the proof of Theorem 2). Furthermore,

$$\mathcal{L}(X''_n | \mathcal{F}'_{n-1} \vee \mathcal{F}''_{n-1}) = \frac{1}{r_n} \sum_{j=1}^{r_n} \delta_{f_n(X''_{n-1}, j)}.$$

Therefore, the conditional law of (X'_n, X''_n) given $\mathcal{F}'_{n-1} \vee \mathcal{F}''_{n-1}$ can be written as follows

$$\mathcal{L}((X'_n, X''_n) | \mathcal{F}'_{n-1} \vee \mathcal{F}''_{n-1}) = \frac{1}{r_n} \sum_{1\leqslant i,j\leqslant r_n} C_{i,j}\ \delta_{\left(f_n(X'_{n-1}, i), f_n(X''_{n-1}, j)\right)},$$

where $(C_{i,j})_{1\leqslant i,j\leqslant r_n}$ is a bistochastic matrix measurable for $\mathcal{F}'_{n-1} \vee \mathcal{F}''_{n-1}$. In particular,

$$E[L_n | \mathcal{F}'_{n-1} \vee \mathcal{F}''_{n-1}] = \frac{1}{r_n} \sum_{1\leqslant i,j\leqslant r_n} C_{i,j}\ e_n\left(f_n(X'_{n-1}, i), f_n(X''_{n-1}, j)\right).$$

This quantity is the image of the matrix $(C_{i,j})_{1\leqslant i,j\leqslant r_n}$ by a linear form. Since bistochastic matrices belong to the convex hull of the permutation matrices, one gets

$$E[L_n|\mathcal{F}'_{n-1} \vee \mathcal{F}''_{n-1}] \geqslant \frac{1}{r_n} \inf_{\sigma\in\mathfrak{S}_{r_n}} \sum_{i=1}^{r_n} e_n\left(f_n(X'_{n-1}, i), f_n(X''_{n-1}, \sigma(i))\right),$$

hence

$$E[L_n|\mathcal{F}'_{n-1} \vee \mathcal{F}''_{n-1}] \geqslant e_{n-1}(X'_{n-1}, X''_{n-1}) = L_{n-1},$$

thanks to the recursion relation verified by the semi-metrics e_n. This shows that $(L_n)_{n\leqslant n_0}$ is a sub-martingale in the filtration $\mathcal{F}' \vee \mathcal{F}''$.

Since $M_{n_0} = \mathbf{1}_{\{X'_{n_0}\neq X''_{n_0}\}} \geqslant L_{n_0}$, one gets, for every $n \leqslant n_0$

$$M_n = E[M_{n_0}|\mathcal{F}'_n \vee \mathcal{F}''_n] \geqslant E[L_{n_0}|\mathcal{F}'_n \vee \mathcal{F}''_n] \geqslant L_n,$$

yielding the inequality $\mathbb{P}[X'_{n_0} \neq X''_{n_0}] \geqslant E[e_n(X'_n, X''_n)]$ by taking the expectations.

This inequality is going to be used to prove that the probability $\mathbb{P}[X'_{n_0} \neq X''_{n_0}]$ can not be made as small as one wishes if the processes X' and X'' are independent until a given time n.

4.3 Bounding below $E[e_n(X'_n, X''_n)]$ when X'_n and X''_n are independent

The purpose of this subsection is to prove the next inequality:

Lemma 4.2 *Let $(X'_k)_{k\leqslant 0}$ and $(X''_k)_{k\leqslant 0}$ be two copies of the split-word process that are independent until time n. Then, for every $\alpha > 0$ and $n \leqslant n_0$,*

$$\mathbb{P}\left[e_n(X'_n, X''_n) \leqslant 1 - N^{-1} - \alpha\right] \leqslant \exp\left(\ell_n(S_n - 2\alpha^2)\right), \text{ where } S_n = \sum_{k=n+1}^{n_0} \frac{\ln(r_k!)}{\ell_{k-1}}$$

The proof is based on Hoeffding's large deviations inequality, see [8] for a proof of this inequality.

Lemma 4.3 (*Hoeffding*) *Let q in $]0, 1[$ and $(\epsilon_k)_{1\leqslant k\leqslant n}$ denote n independent Bernoulli random variables of parameter q, and Z_n their average. Then, for every real $\alpha > 0$,*

$$\mathbb{P}[Z_n \leqslant q - \alpha] \leqslant \exp(-2n\alpha^2).$$

Proof of lemma 4.2 One applies lemma 4.3 to the variables

$$\epsilon_i = \mathbf{1}_{\{X'_n(i)\neq X''_n(i)\}}, \qquad 1 \leqslant i \leqslant \ell_n.$$

Since X'_n and X''_n are independent and uniform on A^{ℓ_n}, the ϵ_i are independent and Bernoulli with parameter $q_N = 1 - 1/N$. Thus, denoting by d_n^H the Hamming distance on A^{ℓ_n},

$$\mathbb{P}\left[\frac{1}{\ell_n} d_n^H(X'_n, X''_n) \leqslant q_N - \alpha\right] \leqslant \exp(-2\ell_n \alpha^2).$$

For every a in Aut_n, $a \cdot X''$ has the same law as X'' and is independent of X'. Since, by definition $e_n(X'_n, X''_n) = \min\{d_n^H(X'_n, a \cdot X''_n)\,;\ a \in \mathrm{Aut}_n\}$, one gets

$$\left[e_n(X'_n, X''_n) \leqslant q_N - \alpha\right] = \left[\exists a \in \mathrm{Aut}_n,\ \frac{1}{\ell_n} d_n^H(X'_n, a \cdot X''_n) \leqslant q_N - \alpha\right],$$

hence

$$\begin{aligned}\mathbb{P}\left[e_n(X'_n, X''_n) \leqslant q_N - \alpha\right] &\leqslant \sum_{a \in \mathrm{Aut}_n} \mathbb{P}\left[\frac{1}{\ell_n} d_n^H(X'_n, a \cdot X''_n) \leqslant q_N - \alpha\right] \\ &= \sharp(\mathrm{Aut}_n)\, \mathbb{P}\left[\frac{1}{\ell_n} d_n^H(X'_n, X''_n) \leqslant q_N - \alpha\right] \\ &\leqslant \exp(\ell_n S_n) \exp(-2\ell_n \alpha^2),\end{aligned}$$

since $\sharp(\mathrm{Aut}_n) = \exp(\ell_n S_n)$. □

Choice of the integer n_0 Under assumption Δ, by remark 1 after Theorem 3, one can choose $n_0 \leqslant 0$ such that

$$S_{-\infty} < 2\,(1 - 1/N)^2, \text{ where } S_{-\infty} = \sum_{k=-\infty}^{n_0} \frac{\ln(r_k!)}{\ell_{k-1}}.$$

Once n_0 is fixed, choose a real α such that $\sqrt{S_{-\infty}/2} < \alpha < 1 - 1/N$.

Since $S_n - 2\alpha^2 \leqslant S_{-\infty} - 2\alpha^2 < 0$ and $\ell_n \geqslant \ell_{n_0}$ for every $n \leqslant n_0$, lemma 4.2 yields

$$\begin{aligned}\mathbb{P}\left[e_n(X'_n, X''_n) \leqslant 1 - 1/N - \alpha\right] &\leqslant \exp\Big(\ell_n (S_n - 2\alpha^2)\Big) \\ &\leqslant \exp\Big(\ell_{n_0} (S_{-\infty} - 2\alpha^2)\Big) = \beta,\end{aligned}$$

with $\beta < 1$. Hence, $E[e_n(X'_n, X''_n)] \geqslant (1 - \beta)\,(1 - 1/N - \alpha)$, which is positive.

4.4 Negation of the I-cosiness criterion

If (X', V') and (X'', V'') are two copies of the split-word processus whose filtrations are co-immersed and independent until time $n \leqslant n_0$,

$$\mathbb{P}[X'_{n_0} \neq X''_{n_0}] \geqslant E[e_n(X'_n, X''_n)] \geqslant (1 - \beta)(1 - 1/N - \alpha),$$

and this lower bound is positive, contradicting the I-cosiness criterion. Thus the filtration of the split-word process is non standard. □

References

1. Dubins, L., Feldman, J., Smorodinsky, M., Tsirelson, B.: Decreasing sequences of σ-fields and a measure change for Brownian motion. Ann.Probab. **24**(2), 882–904 (1996)
2. Émery, M., Schachermayer, W.: On Vershik's standardness criterion and Tsirelson's notion of cosiness. Séminaire de Probabilités, XXXV **1755**, 265–305 (2001)
3. Gorbulsky, A.: About one property of entropy of decreasing sequence of measurable partitions. Zapiski Nauchnykh Seminarov POMI **256**, 19–24 (1999)
4. Heicklen, D.: Bernoullis are standard when entropy is not an obstruction. Israel J. Math. **107**, 141–155 (1998)
5. Laurent, S.: Phd thesis : Filtrations à temps discret négatif. Université Louis Pasteur, Institut de Recherche Mathématique Avancée, Strasbourg (2004)
6. Laurent, S.: On standardness and I-cosiness. Technical Report TR09017, IAP Network, (2009)
7. Laurent, S.: On Vershikian and I-cosy random variables and filtrations. Technical Report TR09018, IAP Network, (2009)
8. Shiryaev, A.N.: Probability, volume 95 of Graduate Texts in Mathematics. Springer-Verlag, (1996)
9. Smorodinsky, M.: Processes with no standard extension. Israel J. Math. **107**, 327–331 (1998)
10. Tsirelson, B.: Triple points: from non-Brownian filtrations to harmonic measures. Geom. Funct. Anal. **7**(6), 1096–1142 (1997)
11. Vershik, A.M.: Decreasing sequences of measurable partitions, and their applications. Soviet Math. Doklady **11**, 1007–1011 (1970)
12. Vershik, A.M.: Four definitions of the scale of an automorphism. Funct. Anal. Appl. **7**(3), 169–181 (1973)
13. Vershik, A.M.: Theory of decreasing sequences of measurable partitions. Saint Petersburg Math. J. **6**, 705–761 (1995)