



HAL
open science

Filtrations of the erased-word processes

Stéphane Laurent

► **To cite this version:**

| Stéphane Laurent. Filtrations of the erased-word processes. 2014. hal-00999719v2

HAL Id: hal-00999719

<https://hal.science/hal-00999719v2>

Preprint submitted on 15 Jul 2014 (v2), last revised 9 Mar 2015 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Filtrations of the erased-word processes

Stéphane Laurent

15/07/2014

Abstract

We define a class of erased-word processes and prove that the poly-adic filtration generated by such a process is standard. This is shown by firstly constructing a generating process of innovations in the case of a finite alphabet equipped with the uniform probability measure, and then by deriving the general case with the help of the tools of Vershik's theory of filtrations in discrete negative time.

1 The filtration of the erased-word process

An erased-word process is depicted on figure 1. It is a stochastic process indexed by the set negative integers $-\mathbb{N}$, and consists in picking at random a word W_n with $|n| + 1$ letters at time n and then to obtain the next word W_{n+1} by deleting at random one letter of W_n . More precisely, given a Lebesgue probability space (A, μ) , and calling A the *alphabet*, the erased-word process on (A, μ) is the Markov chain $(W_n, \eta_n)_{n \leq 0}$ whose law is defined as follows:

- W_n is a random word on A made of $|n| + 1$ letters i.i.d. according to μ ;
- η_{n+1} is a random variable uniformly distributed on $\{1, 2, \dots, |n| + 1\}$ and independent of the past σ -field $\sigma(W_m, \eta_m; m \leq n)$;
- W_{n+1} is obtained by deleting the η_{n+1} -th letter of W_n .

$$\dots \quad W_{-3} = baca \quad \xrightarrow{\eta_{-2}=2} \quad W_{-2} = bca \quad \xrightarrow{\eta_{-1}=3} \quad W_{-1} = bc \quad \xrightarrow{\eta_0=1} \quad W_0 = c$$

Fig. 1: A trajectory of the erased-word process

The filtration \mathcal{F} generated by the erased-word process $(W_n, \eta_n)_{n \leq 0}$ is defined by $\mathcal{F}_n = \sigma(W_m, \eta_m; m \leq n)$. We will sometimes term the η_n as the *erasers*. According to definition given below, the sequence $(\eta_n)_{n \leq 0}$ made of the erasers is a *process of innovations* of the filtration \mathcal{F} .

Definition 1.1. Let \mathcal{F} be a filtration. A random variable η_n that is independent of \mathcal{F}_{n-1} and such that $\mathcal{F}_n = \mathcal{F}_{n-1} \vee \sigma(\eta_n)$ is called an *innovation* of \mathcal{F} (more precisely, we should say an innovation at time n , but thanks to the subscript in η_n this is not a point worth quibbling about). A sequence $(\eta_n)_{n \leq 0}$ of independent random variables such that each η_n is an innovation of \mathcal{F} at time n , is called a *process of innovations* of \mathcal{F} .

When such a process of innovations exist, it defines the local structure of the filtration: for any other process of innovations $(\eta'_n)_{n \leq 0}$, the two random variables η_n and η'_n possibly generate two different σ -fields $\sigma(\eta_n)$ and $\sigma(\eta'_n)$, but there is a Boolean isomorphism between the measure algebras $(\sigma(\eta_n), \mathbb{P})$ and $(\sigma(\eta'_n), \mathbb{P})$ for every $n \leq 0$. Details about this point can be found in [1] and [4]. Thus, any possible innovation η'_n of the filtration of the erased-word process is uniformly distributed on $|n| + 2$ values, similarly to the eraser η_n . For this reason, the filtration \mathcal{F} of the erased-word process is said to be $(|n| + 2)$ -adic, and it belongs to the class of *poly-adic* filtrations, according to definition below.

Definition 1.2. A filtration \mathcal{F} is *poly-adic* if there exists a process of innovations $(\eta_n)_{n \leq 0}$ of \mathcal{F} such that each η_n is uniformly distributed on a finite set.

The poly-adicity will play an important role in the proof of theorem below, which is the main result of this article.

Theorem 1.3. *For any Lebesgue alphabet (A, μ) , the filtration of the erased-word process is of product type, that is to say, it is generated by a process of innovations.*

Let us comment this theorem. Consider the filtration \mathcal{E} generated by the process of innovations $(\eta_n)_{n \leq 0}$. Obviously $\mathcal{E} \subset \mathcal{F}$, but $\mathcal{E} \neq \mathcal{F}$ because η_n is independent of W_n for every $n \leq 0$. But that does not mean that \mathcal{E} and \mathcal{F} are not isomorphic. Theorem 1.3 asserts that there exists another process of innovations $(\eta'_n)_{n \leq 0}$ which generates \mathcal{F} , and this says that \mathcal{E} and \mathcal{F} are isomorphic. Thus \mathcal{F} , which is in some sense richer than \mathcal{E} , is no more than \mathcal{E} up to isomorphism.

This theorem together with Kolmogorov's zero-one law imply that the tail σ -field $\mathcal{F}_{-\infty} := \cap \mathcal{F}_n$ is degenerate. But it is not difficult to directly prove the degeneracy of $\mathcal{F}_{-\infty}$ with the help of the reverse martingale convergence theorem (the proof would be similar to the one given in [2] for the dyadic split-word process), whereas the proof of theorem 1.3, even in the simpler case when A is finite and μ is uniform (see below), is not easy. The motivation of theorem 1.3 is precisely the surprising fact that it is not a trivial result once we know that $\mathcal{F}_{-\infty}$ is degenerate: it is known that for any type of poly-adicity (such as the $(|n| + 2)$ -adicity), there exist some filtrations whose tail σ -fields are degenerate but for which there does not exist any generating process of innovations. Thus, such a filtration is locally isomorphic to \mathcal{F} and, similarly to \mathcal{F} , has a degenerate σ -field, but is not isomorphic to \mathcal{F} . This surprising fact has been discovered by Vershik ([7, 8, 9, 10]), who developed a theory to characterize the existence of a generating process of innovations for poly-adic filtrations.

We will use the tools of the theory of filtrations developed by Vershik to derive the general case in theorem 1.3 from the particular case when μ is uniform on a finite alphabet A . More precisely, our theorem will be proved in three steps:

1. we will prove theorem 1.3 in the case when μ is the uniform probability measure on a finite alphabet A using a 'bare-hands' approach, that is, we will construct a generating process of innovations in this case;
2. using some tools of Vershik's theory, we will prove theorem 1.3 in the case when μ is the Lebesgue measure on $A = [0, 1]$;

3. again using some tools of Vershik's theory, we will derive the general case of theorem 1.3 from the case when μ is the Lebesgue measure on $A = [0, 1]$.

The main theorem of Vershik's theory that will be used is the equivalence between the existence of a generating process of innovations and *standardness* in the case of poly-adic filtrations:

Theorem 1.4. *A poly-adic filtration is of product type if and only if it is standard.*

Different definitions of standardness are used in the literature. Probabilists usually consider that a standard filtration is by definition a filtration which can be *immersed* in a filtration of product type ([2], [4]), and using this definition it is straightforward to see that any filtration of product type is standard. The deep assertion of Theorem 1.4 is the reciprocal fact. Under the usual assumption that the final σ -field \mathcal{F}_0 of the filtration \mathcal{F} is essentially separable, standardness is known to be characterized by a criterion discovered by Vershik, called *Vershik's standardness criterion* or *the Vershik property* ([2], [6], [5]). A filtration satisfying the Vershik property is also said to be *Vershikian*. We will write an easy proposition (proposition 3.1) about the Vershik property to derive step 2 from step 1 in the proof of theorem 1.3. Then step 3 will be straightforwardly derived from step 2 and from the hereditability of standardness under immersion. The standardness property will be more precisely explained in section 3, at time we will resort to it.

There is a recent interest in the study of *adic filtrations* (see [11] and [3]). The filtration of the erased-word process can be viewed as an adic filtration generated by a central random walk on the Bratteli graph such as the one shown on figure 2 for the alphabet $A = \{a, b\}$. This point of view will not be developed in the present paper. Theorem 1.4 shows that the distributions of the erased-word processes studied in this paper are ergodic central measures inducing a standard filtration. The identification of all ergodic central measures and the question of the standardness of the induced filtration is an open direction of further research.

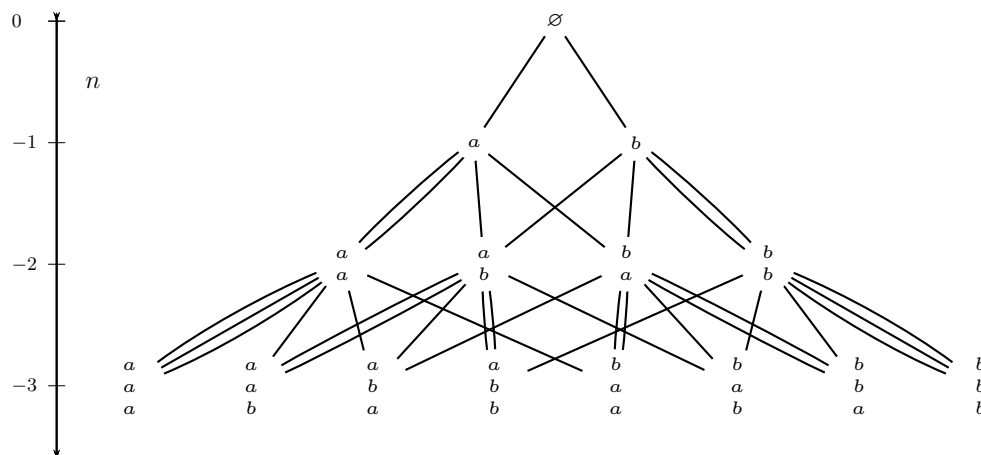


Fig. 2: The Bratteli graph of the erased-word filtration

2 Discrete uniform case

Throughout this section, we assume that A is finite and μ is the uniform probability measure on A . We will prove theorem 1.3 in this case by a more or less explicit construction of a generating process of innovations $(\tilde{\eta}_n)_{n \leq 0}$. We also set $\kappa = \#A$ and we fix a total order on A . Then we denote $A = \{a_1, \dots, a_\kappa\}$ where a_i is the i -th letter of A .

The main ingredient of the construction is the *canonical coupling*. It is very easy to roughly explain what is the canonical coupling with the help of the picture shown on figure 3, but it is a bit tedious to write its rigorous definition. Below, we split the description of the canonical coupling into three paragraphs: we firstly define the *canonical word* (the periodic word at right on figure 3), then we introduce the notation $N_i^-(w)$ for the number of occurrences of the i -th letter of w to the left of position i , and finally we define the canonical coupling of a word (the permutation shown on figure 3, induced by the word at left).

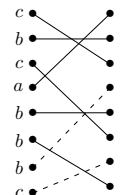


Fig. 3: A canonical coupling

The *canonical word* of length ℓ on A is the word $w \in A^\ell$ in which the letters of A appear in the order and repeat periodically: the i -th letter $w(i)$ of the canonical word w is the r -th letter of A if $i \equiv r \pmod{\kappa}$. For example, the canonical word of length 8 on $A = \{a, b, c\}$, shown at right on figure 3, is the word $abcabcab$.

Given a word w and an index i of one letter of w , we denote by $N_i^-(w) = \sum_{k=1}^{i-1} \mathbb{1}_{w(k)=w(i)}$ the number of occurrences of the i -th letter of w to the left of position i . If w is the canonical word, then $N_i^-(w)$ is the quotient in the Euclidean division of $i - 1$ by $\kappa = \#A$.

The *canonical coupling* ϕ_w of a word w on the finite ordered alphabet $A = \{a_1, \dots, a_\kappa\}$ is the permutation illustrated on Figure 3 and rigorously defined as follows. Let \tilde{w} be the canonical word on A having the same length ℓ as w . The canonical coupling ϕ_w is a permutation of the set $\{1, \dots, \ell\}$ of indices of the letters of w . Its construction is made in two steps:

- *First step.* Take $r \in \{1, \dots, \kappa\}$ and i an index of the letters of w such that the i -th letter $w(i)$ of w is the r -th letter a_r of A . If $N_i^-(w)$ is strictly less than the number of occurrences of a_r in \tilde{w} , then we define $\phi_w(i) = r + N_i^-(w)\kappa$; thus the $\phi_w(i)$ -letter of the canonical word \tilde{w} is the i -th letter of the word w . To sum up, if $w(i)$ is the r -th letter of A , then

$$\left[r + \kappa N_i^-(w) \leq p \right] \implies \left[\phi_w(i) = r + \kappa N_i^-(w) \quad \text{and} \quad w(i) = \tilde{w}(\phi_w(i)) \right]. \quad (2.1)$$

- *Second step.* After performing the first step for every possible r and i , we assign the remaining letters of w to the remaining letters of \tilde{w} in an arbitrary way.

The canonical coupling will help us to construct the generating process of innovations $(\tilde{\eta}_n)_{n \leq 0}$ because of the following reason. Consider the canonical coupling $\phi_{W_{n_0}}$ of W_{n_0} for an arbitrary small n_0 , providing a correspondence between W_{n_0} and the canonical word of length $|n_0| + 1$ denoted by \tilde{w}_{n_0} . Figure 3 is helpful to keep in mind that $\phi_{W_{n_0}}$ represents

one-to-one connections between the letters of W_{n_0} and the letters of \tilde{w}_{n_0} . In parallel to $(W_{n_0}, W_{n_0+1}, \dots, W_0)$, we construct a sequence of erased words $(W'_{n_0}, W'_{n_0+1}, \dots, W'_0)$, starting from $W'_{n_0} = \tilde{w}_{n_0}$ and erasing one letter at each step as follows. At time $n = n_0 + 1$, the word W_{n_0+1} is obtained by deleting the η_{n_0+1} -th letter of W_{n_0} , and we delete the corresponding $\phi_{W_{n_0}}(\eta_{n_0+1})$ -th letter of the canonical word $\tilde{w}_{n_0} = W'_{n_0}$, thereby obtaining a subword W'_{n_0+1} of W'_{n_0} having the same length as W_{n_0+1} . Thus $\eta'_{n_0+1} := \phi_{W_{n_0}}(\eta_{n_0+1})$ is the first eraser in the parallel erased-word sequence $(\tilde{w}_{n_0}, W'_{n_0+1}, \dots, W'_0)$, and its realization fully determines the realization of the random word W'_{n_0+1} . Moreover, by deleting the connection between η_{n_0+1} and η'_{n_0+1} in the canonical coupling $\phi_{W_{n_0}}$, we obtain a new permutation $\phi_{W_{n_0}, \eta_{n_0+1}}$ representing one-to-one connections between the letters of W_{n_0+1} and the letters of W'_{n_0+1} . Then we continue so on (this is illustrated on Figure 4):

- At each time $n \in \{n_0 + 1, \dots, -1\}$ we have a word W'_n of length $|n| + 1$ and a permutation $\phi_{W_{n_0}, \eta_{n_0+1}, \dots, \eta_n}$ representing one-to-one connections between the letters of W'_n and the letters of W_n .
- At time $n + 1$ we have a word W'_{n+1} obtained by deleting the η'_{n+1} -th letter of W'_n , where $\eta'_{n+1} = \phi_{W_{n_0}, \eta_{n_0+1}, \dots, \eta_n}(\eta_{n+1})$, and this provides a new permutation $\phi_{W_{n_0}, \eta_{n_0+1}, \dots, \eta_{n+1}}$ connecting the letters of W'_{n+1} to the letters of W_{n+1} .

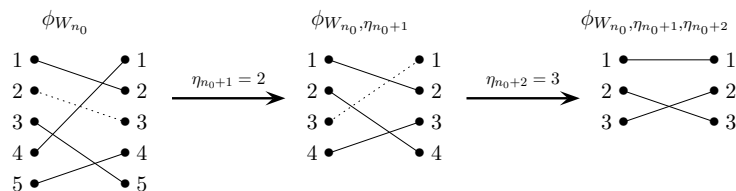


Fig. 4: A cascaded permutation

Figure 4 illustrates the “cascaded” permutations $\phi_{W_{n_0}, \eta_{n_0+1}, \dots, \eta_n}$ initiated at time $n_0 = -4$ by the canonical coupling $\phi_{W_{n_0}}$ and sequentially obtained from the erasers η_{n_0+1} and η_{n_0+2} .

The word W_n is a subword of W_{n_0} , and there is a bijective correspondence between the erasers $(\eta_{n_0+1}, \dots, \eta_n)$ and the random set $Q_{n_0, n} \subset \{1, \dots, |n_0| + 1\}$ of the indices of the letters of the word W_{n_0} forming its subword W_n . This is similar for the parallel sequence $(W'_{n_0}, W'_{n_0+1}, \dots, W'_n)$, and $Q'_{n_0, n}$ is the image of $Q_{n_0, n}$ by the canonical coupling $\phi_{W_{n_0}}$. Lemma 2.1 below, applied with $W = W_{n_0}$ and $Q = Q_{n_0, n}$, says that $W_n = W'_n$ with probability as high as desired when the construction starts from an arbitrary small n_0 . We will see in addition that the η'_n are innovations of \mathcal{F} and this will allow us to derive a generating process of innovations of \mathcal{F} . In Lemma 2.1, we use the notation $w_{|J}$ to denote the subword of a word w obtained by keeping only those of its letters whose indices belong to the subset J .

Lemma 2.1. *Let p and q be integer numbers satisfying $p > q > 0$. Let W be a random variable having the uniform law on the p -words on A and let Q be a random variable*

independent of W having the uniform law on the subsets of $\{1, \dots, p\}$ having size q . We denote by $\pi(p, q)$ the probability

$$\pi(p, q) = \mathbb{P}\left(W|_Q = \tilde{w}|_{\phi_W(Q)}\right),$$

where ϕ_w is the canonical coupling of a word w and \tilde{w} is the canonical word of length p . Then $\pi(p, q) \rightarrow 1$ when $p \rightarrow \infty$.

Proof. Recall the notation $\kappa = \#A$. To show that $\pi(p, q) \rightarrow 1$ when $p \rightarrow \infty$, we introduce the three events

$$\begin{aligned} E_1 &= \left\{ \max(Q) \leq p - p^{3/4} - \kappa \right\}, \\ E_2 &= \left\{ \forall i, j \in Q, i = j \text{ or } |i - j| \geq 3p^{3/4} \right\}, \\ E_3 &= \left\{ \forall i \in Q, i - 1 - p^{3/4} \leq \kappa N_i^-(w) \leq i - 1 + p^{3/4} \right\}, \end{aligned}$$

and we are going to show that

$$E_1 \cap E_2 \cap E_3 \subset \left\{ W|_Q = \tilde{w}|_{\phi_W(Q)} \right\}$$

if p is sufficiently large, and

$$\mathbb{P}(E_1^c \cup E_2^c \cup E_3^c) \xrightarrow{p \rightarrow \infty} 0.$$

We firstly show the inclusion. As a first step, we show that $W|_Q$ and $\tilde{w}|_{\phi_W(Q)}$ have the same letters on $E_1 \cap E_3$. Consider $i \in Q$. On the event E_3 , if $W(i)$ is the r -th letter of A then

$$r + \kappa N_i^-(W) \leq \kappa + \kappa N_i^-(W) \leq \kappa + i + p^{3/4},$$

and $\kappa + i + p^{3/4} \leq p$ on the event E_1 . Thus, by the fundamental equation (2.1) of the canonical coupling, $\phi_W(i) = r + \kappa N_i^-(W)$ on the event $E_1 \cap E_3$ and the two words $W|_Q$ and $\tilde{w}|_{\phi_W(Q)}$ have the same letters on the event $E_1 \cap E_3$. Moreover,

$$r + i - 1 - p^{3/4} \leq r + \kappa N_i^-(W) \leq r + i + p^{3/4},$$

on the event E_3 , and then we get $i - p^{3/4} \leq \phi_W(i) \leq \kappa + i + p^{3/4}$ on the event $E_1 \cap E_3$. Consequently ϕ_W satisfies the following property on $E_1 \cap E_3$:

$$\forall i \in Q, \quad \left| i - \phi_W(i) \right| \leq \kappa + p^{3/4}.$$

Now, if we are on $E_1 \cap E_2 \cap E_3$ and if p is sufficiently large so that $3p^{3/4} > 2(\kappa + p^{3/4})$, then the restriction of ϕ_W to Q is increasing because of the previous property about ϕ_W holding on $E_1 \cap E_3$ and because of the definition of E_2 . Finally, as we have seen that the two words $W|_Q$ and $\tilde{w}|_{\phi_W(Q)}$ have the same letters on the event $E_1 \cap E_3$, then they are equal on $E_1 \cap E_2 \cap E_3$, and we get

$$E_1 \cap E_2 \cap E_3 \subset \left\{ W|_Q = \tilde{w}|_{\phi_W(Q)} \right\},$$

as desired.

It remains to show that $\mathbb{P}(E_1^c \cup E_2^c \cup E_3^c) \xrightarrow{p \rightarrow \infty} 0$. The following majoration of $\mathbb{P}(E_1^c)$ is easily obtained:

$$\mathbb{P}(E_1^c) \leq q \frac{p^{3/4} + \kappa + 1}{p} \xrightarrow{p \rightarrow \infty} 0.$$

The following majoration of $\mathbb{P}(E_2^c)$ is obtained by sampling the elements of Q without replacement:

$$\begin{aligned} \mathbb{P}(E_2^c) &\leq \frac{6p^{3/4} + 2}{p-1} + \frac{2(6p^{3/4} + 2)}{p-2} + \dots + \frac{(q-1)(6p^{3/4} + 2)}{p-q+1} \\ &\leq \frac{(q-1)^2(6p^{3/4} + 2)}{p-q+1} \xrightarrow{p \rightarrow \infty} 0. \end{aligned}$$

To find a majoration of $\mathbb{P}(E_3^c)$, we call I_k the k -th element of Q for every $k \in \{1, \dots, q\}$. Conditionally to $I_k = i$, the number of occurrences $N_{I_k}^-(W)$ has the binomial distribution with size $i-1$ and probability of success $1/\kappa$, because of the independence between W and Q . Therefore, using Bienaymé-Chebyshev's inequality,

$$\begin{aligned} \mathbb{P}\left(\left|N_{I_k}^-(W) - \frac{I_k - 1}{\kappa}\right| > \frac{p^{3/4}}{\kappa} \mid I_k = i\right) &\leq \left(\frac{\kappa}{p^{3/4}}\right)^2 (i-1) \frac{1}{\kappa} \left(1 - \frac{1}{\kappa}\right) \\ &< \frac{p\kappa}{p^{3/2}} = \frac{\kappa}{\sqrt{p}}, \end{aligned}$$

and this being true for every $i \in \{1, \dots, p\}$,

$$\mathbb{P}\left(\left|N_{I_k}^-(W) - \frac{I_k - 1}{\kappa}\right| > \frac{p^{3/4}}{\kappa}\right) < \frac{\kappa}{\sqrt{p}}.$$

By summing this equality over all $k \in \{1, \dots, q\}$,

$$\mathbb{P}(E_3^c) < \frac{q\kappa}{\sqrt{p}} \xrightarrow{p \rightarrow \infty} 0.$$

□

We now formalize the discussion above lemma 2.1. Denote by f_n the function such that $W_n = f_n(W_{n-1}, \eta_n)$, that is to say $f_n(w, e)$ is simply defined as the word obtained by deleting the e -th letter of w , and consider a random vector $(\eta'_{n_0+1}, \dots, \eta'_0)$ having the same law as $(\eta_{n_0+1}, \dots, \eta_0)$. Then define a Markov process $(Y_n(n_0, w_{n_0}))_{n_0 \leq n \leq 0}$ by the initial condition $Y_{n_0}(n_0, w_{n_0}) = w_{n_0}$ for some given word w_{n_0} in the state space of W_{n_0} , and by the inductive relation

$$Y_{n+1}(n_0, w_{n_0}) := f_{n+1}(Y_n(n_0, w_{n_0}), \eta'_{n+1}).$$

Setting $W'_n = Y_n(n_0, w_{n_0})$, the process $((W'_{n_0+1}, \eta'_{n_0+1}), \dots, (W'_0, \eta'_0))$ has the same distribution as the process $((W_{n_0+1}, \eta_{n_0+1}), \dots, (W_0, \eta_0))$ conditionally to $W_{n_0} = w_{n_0}$. The case discussed above lemma 2.1 is the context of the following lemma, where N plays the role of n_0 .

Lemma 2.2. *Let $N < 0$ be an integer and ϕ_{W_N} be the canonical coupling of W_N . For every integer $m \in [N + 1, -1]$ let $\phi_{W_N, \eta_{N+1}, \dots, \eta_m}$ be, as explained above lemma 2.1, the cascaded permutation initiated by ϕ_{W_N} and sequentially obtained from η_N, \dots, η_m , and define $\eta'_{m+1} = \phi_{W_N, \eta_{N+1}, \dots, \eta_m}(\eta_{m+1})$ for every $m \in [N, -1]$. Then $(\eta'_n)_{N+1 \leq n \leq 0}$ has the same law as $(\eta_n)_{N+1 \leq n \leq 0}$ and each η'_n is, just as η_n , an innovation of \mathcal{F} , that is, η'_n is independent of \mathcal{F}_{n-1} for every n and $\mathcal{F}_n = \mathcal{F}_{n-1} \vee \sigma(\eta'_n)$. Moreover, with the notations above,*

$$\mathbb{P}(W_m \neq Y_m(N, \tilde{w}_N)) \rightarrow 0 \quad \text{as } N \rightarrow -\infty$$

for every $m \in [N + 1, 0]$, where \tilde{w}_N is the canonical word of length $|N| + 1$.

Proof. It is easy to check that η'_n is an innovation as any \mathcal{F}_{n-1} -measurable random permutation of η_n . The rest of the lemma is nothing but lemma 2.1. \square

Proposition 2.3. *The erased-words process $(W_n, \eta_n)_{n \leq 0}$ generates a filtration of product type.*

Proof. First note that $Y_m(N, \tilde{w}_N)$ in Lemma 2.2 is measurable with respect to $\sigma(\eta'_{N+1}, \dots, \eta'_m)$. Then, given a sequence $(\delta_k)_{k \leq 0}$ of real numbers $\delta_k > 0$ satisfying $\delta_k \rightarrow 0$ as $k \rightarrow -\infty$, applying Lemma 2.2 recursively provides a strictly increasing sequence $(N_k)_{k \leq 0}$ of integers with $N_0 = 0$ and an innovation process $(\tilde{\eta}_n)_{n \leq 0}$ such that, for every $k < 0$,

- (i) $(\tilde{\eta}_{N_k+1}, \dots, \tilde{\eta}_n) = \tau_{W_{N_k}}(\eta_{N_k+1}, \dots, \eta_n)$ for every integer $n \in [N_k + 1, 0]$, where each τ_w is a permutation preserving the law of $(\eta_{N_k+1}, \dots, \eta_n)$;
- (ii) there is a random word \tilde{W}_{N_k} measurable with respect to $\sigma(\tilde{\eta}_{N_k-1+1}, \dots, \tilde{\eta}_{N_k})$ and satisfying $\mathbb{P}(W_{N_k} \neq \tilde{W}_{N_k}) < \delta_k$.

Now we check that $(\tilde{\eta}_n)_{n \leq 0}$ generates \mathcal{G} . It suffices to construct, for each $n \leq 0$ and every $\delta > 0$, a pair of random variables $(\hat{W}_n, \hat{\eta}_n)$ that is measurable with respect to $\sigma(\dots, \tilde{\eta}_{n-1}, \tilde{\eta}_n)$ and that satisfies $\mathbb{P}((W_n, \eta_n) \neq (\hat{W}_n, \hat{\eta}_n)) < \delta$. To do so, let k be sufficiently small in order that $\delta_k < \delta$ and $N_k < n$. Then define

$$(\hat{\eta}_{N_k+1}, \dots, \hat{\eta}_n) = \tau_{\tilde{W}_{N_k}}^{-1}(\tilde{\eta}_{N_k+1}, \dots, \tilde{\eta}_n)$$

and define \hat{W}_m for $m \in [N_k, n]$ by initially setting $\hat{W}_{N_k} = \tilde{W}_{N_k}$ and recursively setting $\hat{W}_{m+1} = f_{m+1}(\hat{W}_m, \hat{\eta}_{m+1})$. Now, $(\hat{\eta}_{N_k+1}, \dots, \hat{\eta}_n) = (\eta_{N_k+1}, \dots, \eta_n)$ on the event $\{W_{N_k} = \tilde{W}_{N_k}\}$, hence $\hat{W}_n = W_n$ on this event too. \square

Alternatively, one could derive proposition 2.3 from lemma 2.2 with the help of *Vershik's first level criterion* and proposition 2.22 in [4].

3 Vershikian tools and the general case

Here we finish the proof of theorem 1.3 by following step 2 and step 3 announced in the introduction.

Consider the erased-word process $(W_n, \eta_n)_{n \leq 0}$ in the case when $A = [0, 1]$ and μ is the Lebesgue measure on A , and denote by \mathcal{G} the filtration it generates. Take a sequence $(f_k)_{k \geq 1}$ of functions $f_k: A \rightarrow A_k := \{0, \frac{1}{k}, \dots, \frac{k-1}{k}\}$, each sending μ to the uniform probability measure on A_k , and such that $f_k(U) \rightarrow_{k \rightarrow +\infty} U$ almost surely for any random variable $U \sim \mu$. Applying f_k to each letter of a word w on A gives a word on A_k denoted by $f_k(w)$. Then the process $(f_k(W_n), \eta_n)_{n \leq 0}$ is the erased-word process on the alphabet A_k equipped with the uniform probability measure. Denote its filtration by \mathcal{G}^k . We know by proposition 2.3 that each filtration \mathcal{G}^k is of product type. Using this result, the tools of the theory of filtrations will allow us to prove that \mathcal{G} is of product type. We will use theorem 1.4 and the equivalence between standardness and the Vershik property, as well as our original proposition 3.1 stated below. The key point is the fact that each \mathcal{G}^k is immersed in \mathcal{G} , which means here that the process $(f_k(W_n), \eta_n)_{n \leq 0}$ is Markovian with respect to \mathcal{G} (see Lemma 2.6 in [6], we do not provide in the present paper the general definition of the immersion property), and this allows us to use the following proposition.

Proposition 3.1. *Let \mathcal{F} be a filtration. If there exists a sequence of Vershikian filtrations $(\mathcal{F}^k)_{k \geq 1}$ such that $\bigvee_{k=1}^{+\infty} \mathcal{F}_0^k = \mathcal{F}_0$ and such that each \mathcal{F}^k is immersed in \mathcal{F} , then \mathcal{F} is Vershikian.*

Proof. This is an easy corollary of Lemma 4.1 and Lemma 4.2 in [6]. □

Thus, we know that \mathcal{G} is of product type by proposition 3.1 and theorem 1.4, and step 2 of the proof of theorem 1.3 is achieved. Now consider an arbitrary Lebesgue alphabet (A, μ) and take a measurable function $f: [0, 1] \rightarrow A$ sending the Lebesgue measure on A . Then the process $(f(W_n), \eta_n)_{n \leq 0}$ is the erased-word process on (A, μ) , and the filtration it generates is immersed in \mathcal{G} . We conclude that this filtration is of product type by using theorem 1.4 and the hereditability of standardness under immersion, an immediate consequence of the definition of standardness (see [2] or [4]). Now step 3 is achieved and the proof of theorem 1.3 is over.

Alternatively, instead of using proposition 3.1 involving Vershik's standardness criterion, step 2 could have been proved by using *Vershik's first level criterion*, similarly to the proof of result 2.45 in [4] about the split-word processes. This proof would be more elementary but less concise.

4 Standardness of the multidimensional Pascal filtration

The d -dimensional Pascal filtration is introduced in [3]. It is the filtration generated by the Markov chain $(V_n)_{n \leq 0}$ whose distribution is defined from a given probability vector $(\theta_1, \dots, \theta_d)$, where $d \in \mathbb{N} \cup \{+\infty\}$, as follows:

- (*instantaneous distributions*) the random variable V_n has the multinomial distribution on

$$\mathbb{V}_n^d = \left\{ (i_1, \dots, i_d) \in \mathbb{N}^d \mid i_1 + \dots + i_d = |n| \right\}$$

with success probability vector $(\theta_1, \dots, \theta_d)$;

- (*Markov transitions*) the transition laws from n to $n + 1$ are

$$\mathcal{L}(V_{n+1} \mid V_n = v) = \sum_{i=1}^d \frac{v(i)}{|n|} \delta_{v-e_i}, \tag{4.1}$$

where e_i is the vector in \mathbb{R}^d whose i -th term is 1 and all the other ones are 0. In other words, given $V_n = (v_n(1), \dots, v_n(d))$, coordinate i is picked at random with probability $\frac{v_n(i)}{|n|}$ and V_{n+1} is obtained by subtracting 1 to this coordinate.

The case when $d = 2$ is illustrated on Figure 5 and Figure 6. (with p playing the role of θ_1), and the case when $d = 3$ is illustrated on Figure 7. We refer to [3] for more detailed explanations.

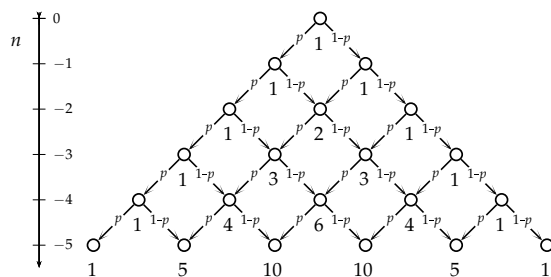


Fig. 5: 2-dimensional Pascal random walk, directed from $n = 0$ to $n = -\infty$

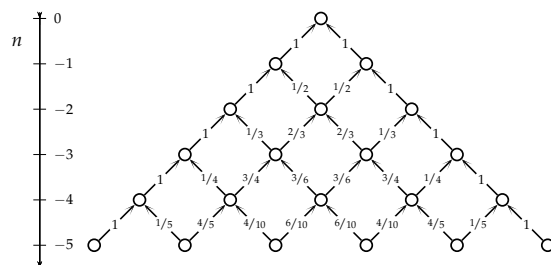


Fig. 6: 2-dimensional Pascal random walk, directed from $n = -\infty$ to $n = 0$

It has been shown in [3] that the filtration generated by the d -dimensional Pascal random walk is standard for any d and any $(\theta_1, \dots, \theta_d)$. This result is straightforwardly derived from our Theorem 1.3 and from the heritability property of standardness under immersion, already mentioned in section 3. Indeed, taking the erased-word process $(W_n, \eta_n)_{n \leq 0}$ on an alphabet A with d letters and equipped with the probability μ whose masses are given by the probability vector $(\theta_1, \dots, \theta_d)$, and defining the function $f_n: A^{|n|+1} \rightarrow \mathbb{V}_n$ as the one

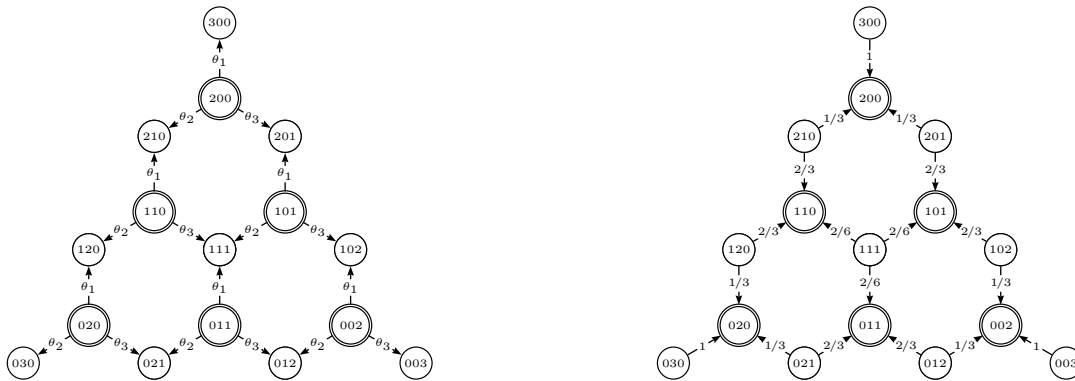


Fig. 7: Step in the 3-dimensional Pascal random walk: from $n = -2$ to $n = -3$ (left), and from $n = -3$ to $n = -2$ (right)

returning the list of the numbers of occurrences of each letter of A in a given word of length $|n| + 1$, then the process $(f_n(W_n), \eta_n)_{n \leq 0}$ is the d -dimensional Pascal random walk defined by the probability vector $(\theta_1, \dots, \theta_d)$, and the filtration \mathcal{F} it generates is immersed in the filtration \mathcal{G} generated by the erased-word process $(W_n, \eta_n)_{n \leq 0}$ because $(f_n(W_n), \eta_n)_{n \leq 0}$ is Markovian with respect to \mathcal{G} . Then standardness of the d -dimensional Pascal filtration \mathcal{F} results from theorem 1.3 and from the hereditability property of standardness under immersion.

Acknowledgments. I thank Christophe Leuridan for his relevant comments on the first version of this paper.

References

- [1] M. Barlow, M. Émery, F. Knight, S. Song, M. Yor: *Autour d'un théorème de Tsirelson sur des filtrations browniennes et non-browniennes*. Séminaire de probabilités XXXII, Springer Lectures Notes in Math. 1686, 264–305, 1998.
- [2] M. Émery, W. Schachermayer: *On Vershik's standardness criterion and Tsirelson's notion of cosiness*. Séminaire de Probabilités XXXV, Springer Lectures Notes in Math. 1755, 265–305 (2001)
- [3] E. Janvresse, S. Laurent, T. de la Rue: *Standardness of the filtration of a monotonic Markov process*. Preprint 2014 (in preparation).
- [4] S. Laurent: *On standardness and I-cosiness*. Séminaire de Probabilités XLIII, Springer Lecture Notes in Mathematics 2006, 127–186 (2010)
- [5] S. Laurent: *Vershik's Intermediate Level Standardness Criterion and the Scale of an Automorphism*. Séminaire de Probabilités XLV, Springer Lecture Notes in Mathematics 2078, 123–139 (2013).
- [6] S. Laurent: *On Vershikian and I-cosy random variables and filtrations*. Teoriya Veroyatnosti i ee Primeneniya 55, 104–132 (2010). Also published in: Theory Probab. Appl. 55, 54–76 (2011)

-
- [7] A.M. Vershik: *Theorem on lacunary isomorphisms of monotonic sequences of partitions*. Funktsional'nyi Analiz i Ego Prilozheniya **2**:3, 17–21, 1968. English translation: Functional Analysis and Its Applications **2**:3, 200–203, 1968.
- [8] A.M. Vershik: *Decreasing sequences of measurable partitions, and their applications*. Dokl. Akad. Nauk SSSR 193, 748–751 (1970). English translation: Soviet Math. Dokl. 11, 1007–1011 (1970)
- [9] A.M. Vershik: *Approximation in measure theory* (in Russian). PhD Dissertation, Leningrad University, 1973. (expanded and updated version: [10])
- [10] A.M. Vershik: *The theory of decreasing sequences of measurable partitions (in Russian)*. Algebra i Analiz, 6:4, 1–68 (1994). English translation: St. Petersburg Mathematical Journal, 6:4, 705–761 (1995)
- [11] A.M. Vershik: *Intrinsic metric on graded graphs, standardness, and invariant measures*. Zapiski Nauchn. Semin. POMI 421, 58–67 (2014). arXiv:1312.7239. <http://arxiv.org/abs/1312.7239>