HAL open science

Filtrations of the erased-word processes

Stéphane Laurent

▶ To cite this version:

Stéphane Laurent. Filtrations of the erased-word processes. 2015. hal-00999719v3

HAL Id: hal-00999719
https://hal.science/hal-00999719v3

Preprint submitted on 9 Mar 2015

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

09/03/2015

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 of negative integers $-\mathbb{N}$, and consists in picking at random a word W_n with $|n|$ letters at time n and then to obtain the next word W_{n+1} by deleting at random one letter of W_n (thus the final word W_0 is the empty word). More precisely, given a Lebesgue probability space (A, μ), and calling A the *alphabet*, the erased-word process on (A, μ) is the Markov process $(W_n, \eta_n)_{n \leqslant 0}$ whose law is defined as follows: for every $n \leqslant -1$,

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

$$\cdots \quad W_{-3} = bac \quad \xrightarrow{\eta_{-2}=3} \quad W_{-2} = ba \quad \xrightarrow{\eta_{-1}=1} \quad W_{-1} = a \quad \xrightarrow{\eta_0=1} \quad W_0 = \varnothing$$

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

The filtration $\mathcal{F}$ generated by the erased-word process $(W_n, \eta_n)_{n \leqslant 0}$ is defined by $\mathcal{F}_n = \sigma(W_m, \eta_m; m \leqslant n)$. We will sometimes term the η_n as the *erasers*. According to definition given below, the sequence $(\eta_n)_{n \leqslant 0}$ 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 $(\eta_n)_{n\leqslant 0}$ 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 $(\eta'_n)_{n\leqslant 0}$, 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 $\big(\sigma(\eta_n), \mathbb{P}\big)$ and $\big(\sigma(\eta'_n), \mathbb{P}\big)$ for every $n \leqslant 0$. Details about this point can be found in [1] and [5]. Thus, any possible innovation η'_n of the filtration of the erased-word process is uniformly distributed on $|n|+1$ values, similarly to the eraser η_n. For this reason, the filtration $\mathcal{F}$ of the erased-word process is said to be $\big(|n|+1\big)$*-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 $(\eta_n)_{n\leqslant 0}$ 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 $(\eta_n)_{n\leqslant 0}$. Obviously $\mathcal{E} \subset \mathcal{F}$, but $\mathcal{E} \neq \mathcal{F}$ because $\mathcal{E}_n$ is independent of W_n for every $n \leqslant 0$. 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 $(\tilde{\eta}_n)_{n\leqslant 0}$ which generates $\mathcal{F}$ (then called a *generating process of innovations*), and this says that $\mathcal{E}$ and $\mathcal{F}$ are isomorphic. Thus $\mathcal{F}$, which is bigger 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 (this 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 our three demonstration steps), 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 $\big(|n|+1\big)$-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 ([8, 9, 10, 11]), who developed a theory to characterize the existence of a generating process of innovations for poly-adic filtrations.

To give a better idea of the subtlety of Theorem 1.3, we mention that the opposite conclusion holds for a process seemingly close to the erased-word process. Namely, this process is similar to the erased-words process except that at each time, the letter deleted at random is either the first one or the last one with equal probabilities of 1/2. It

generates a dyadic (2-adic) filtration which is not of product type although its tail σ-field is degenerate. This result is not explicitely written in the literature but Heicklen and Hoffman's proof of their main result in [3] implicitly relies on it.

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], [5]), and this definition directly yields that filtrations of product type are 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], [7], [6]). 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 derived from step 2 and from the heritability property of standardness under immersion, with the help of Theorem 1.4.

The standardness property will be more precisely explained in Section 3, at time we will resort to it.

In Section 4 we derive standardness of the multidimensional Pascal filtration from standardness of the erased-word filtration. This filtration arises when one observes the evolution of the numbers of occurences of each letter in the erased-word process.

It is worth mentioning that the filtration of the erased-word process can be viewed as the filtration induced by an ergodic central measure on a Bratteli graph, because there is a recent interest in the study of standardness of such filtrations ([12, 13, 4]). This Bratelli graph arises by taking as probability space of the erased-word process a so-called canonical space in the theory of stochastic process, that is, a space representing the trajectories of the erased-word process, equipped with a probability measure ν such

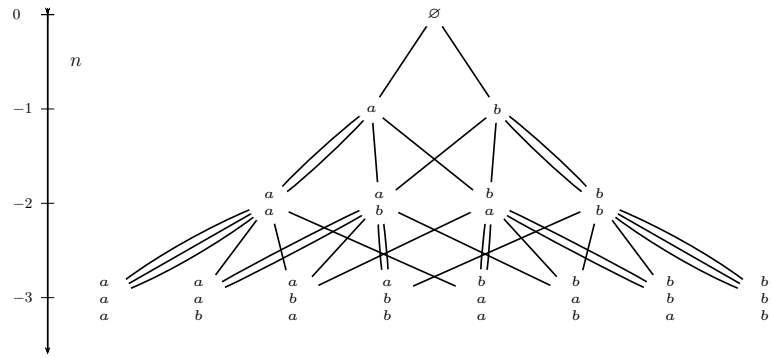

Fig. 2: The Bratteli graph of the erased-word filtration

that picking a trajectory according to ν defines the law of the erased-word process. The Bratteli graph is shown on Figure 2 for a two-letters alphabet $A = \{a, b\}$.

The graph is graded: the vertices at each level $n \leqslant 0$ correspond to the possible states of the random word W_n, in particular it has a unique vertex $\varnothing$ at level 0. The edges connecting a vertex at level $n-1$ to a vertex at level n correspond to the possible values of the eraser η_n. In this way, a trajectory of the erased-word process corresponds to an infinite path in the graph, starting from the root vertex $\varnothing$, and the law of the erased-word process defines a probability measure ν on the set Γ of such paths. This graph is termed as Bratelli because in addition to be graded, each vertex at level $n \leqslant 0$ is connected to at least one vertex at level $n-1$, and each vertex at level $n \leqslant -1$ is connected to at least one vertex at level n. Thus, a trajectory of the erased-words process can be viewed as an infinite path in Γ taken at random according to ν, and the σ-field $\mathcal{F}_n$ is the one generated by the path observed up to level n. We can similarly define the filtration $\mathcal{F}$ for any Bratteli graph and a given probability measure ν on the space Γ of its infinite paths. In our case where ν is the law of the erased-word process, it is an *ergodic central measure* with the terminology of Vershik ([12, 13]). The measure ν is said to be *central* when for every given path observed up to level n, the remaining finite piece of the path from the vertex picked at level n to the root vertex $\varnothing$ at level 0 is taken uniformly on the set of all such finite pieces of path (see [4] for more details). This obviously holds in our case because of the poly-adicity of $\mathcal{F}$. The measure ν is said to be *ergodic* when the tail σ-field $\mathcal{F}_{-\infty}$ is degenerate, and as already said before, this property holds in our case as a consequence of standardness but it is not difficult to prove it directly.

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 $(\tilde{\eta}_n)_{n\leqslant 0}$. We also set $\kappa = \#A$ and we fix a total order on A. Then we denote $A = \{a_1, \ldots, a_\kappa\}$ where a_i is the i-th letter of A.

2.1 Ingredients of the construction

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).

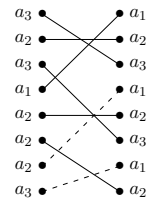

Fig. 3: A canonical coupling

The *canonical word* of length ℓ on A is the word $\tilde{w} \in A^\ell$ in which the letters of A appear in the order and repeat periodically: the i-th letter $\tilde{w}(i)$ of the canonical word $\tilde{w}$ is the r-th letter a_r of A if $i \equiv r \pmod{\kappa}$. For example, the canonical word of length 8 on $A = \{a_1, a_2, a_3\}$, shown at right on Figure 3, is the word $a_1a_2a_3a_1a_2a_3a_1a_2$.

Given a word w and a position in w, that is to say an index i of one letter of w, we denote by $N_i^-(w) = \sum_{k=1}^{i-1} 1_{\{w(k)=w(i)\}}$ the number of occurrences of the i-th letter of w to the left of position i. If $\tilde{w}$ is the canonical word, then $N_i^-(\tilde{w})$ 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 $A = \{a_1, \ldots, a_\kappa\}$ 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 $\{1, \ldots, \ell\}$ of positions in w. Its construction is made in two steps:

- *First step.* Take a position $i \in \{1, \ldots, \ell\}$ in w. If $w(i) = a_r$ then we set $\phi_w(i) = r + \kappa N_i^-(w)$ provided $r + \kappa N_i^-(w) \leqslant \ell$, i.e. when $N_i^-(w)$ is strictly less than the number of occurrences of a_r in the canonical word $\tilde{w}$. Then $w(i) = \tilde{w}\big(\phi_w(i)\big)$ for all such i. This step is illustrated on Figure 3 by the solid lines.

- *Second step.* After performing the first step for every possible i, we assign the remaining positions in w to the remaining positions in $\tilde{w}$ in the increasing way. This step is illustrated on Figure 3 by the dashed lines.

The last ingredient of the construction are the *cascaded permutations*. 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 $|n_0|$ 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 $(W_{n_0}, W_{n_0+1}, \ldots, W_0)$, we construct a sequence of erased words $(W'_{n_0}, W'_{n_0+1}, \ldots, W'_0)$, 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 $(\tilde{w}_{n_0}, W'_{n_0+1}, \ldots, W'_0)$, 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 $n \in \{n_0+1, \ldots, -1\}$ we have a word W'_n of length $|n|$ and a permutation $\phi_{W_{n_0},\eta_{n_0+1},\ldots,\eta_n}$ 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 $\eta'_{n+1} = \phi_{W_{n_0},\eta_{n_0+1},\ldots,\eta_n}(\eta_{n+1})$, and this provides a new permutation $\phi_{W_{n_0},\eta_{n_0+1},\ldots,\eta_{n+1}}$ connecting the letters of W'_{n+1} to the letters of W_{n+1}.

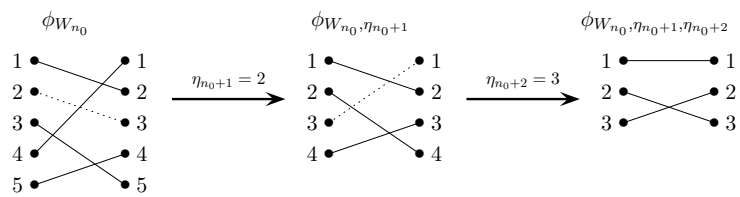

Fig. 4: A cascaded permutation

Figure 4 illustrates the "cascaded" permutations $\phi_{W_{n_0},\eta_{n_0+1},\ldots,\eta_n}$ initiated at time $n_0 = -5$ by the canonical coupling $\phi_{W_{n_0}}$ and sequentially obtained from the erasers η_{n_0+1} and η_{n_0+2}. By this way the random word W'_n is measurable with respect to $\sigma(\eta'_{n_0}+1, \ldots, \eta'_n)$. Lemma 2.1 in the next section shows that the η'_n are innovations of $\mathcal{F}$ and $W_n = W'_n$ with probability as high as desired when the construction starts from an arbitrary small n_0, and this will allow us to construct a generating process of innovations.

2.2 Key lemma

A generating process of innovations of $\mathcal{F}$ will be derived from Lemma 2.1 below. The following construction, already sketched in the previous section, is used in the statement of this lemma. Let f_n be the function from $A^{|n|+1} \times \{1, \ldots, |n|+1\}$ to $A^{|n|}$ defined by setting $f_n(w, e)$ to the word obtained by deleting the e-th letter of w. This function represents the update from W_{n-1} to W_n because of the equality $W_n = f_n(W_{n-1}, \eta_n)$. Now, consider a random vector $(\eta'_{n_0+1}, \ldots, \eta'_0)$ having the same law as $(\eta_{n_0+1}, \ldots, \eta_0)$. Then define a Markov process $\big(Y_n(n_0)\big)_{n_0 \leqslant n \leqslant 0}$ by the initial condition $Y_{n_0}(n_0) = \tilde{w}_{n_0}$ (the canonical word of length $|n_0|$), and by the inductive relation

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

Setting $W'_n = Y_n(n_0)$, the process $\Big((W'_{n_0+1}, \eta'_{n_0+1}), \dots, (W'_0, \eta'_0)\Big)$ has the same distribution as the process $\Big((W_{n_0+1}, \eta_{n_0+1}), \dots, (W_0, \eta_0)\Big)$ conditionally to $W_{n_0} = \tilde{w}_{n_0}$. Lemma 2.1 below shows that $W_n = W'_n$ with probability as high as desired when the construction starts from an arbitrary small n_0 and when we use the innovations η'_n encountered in the previous section when we defined the cascaded permutations.

Lemma 2.1. *Let $n_0 < 0$ be an integer and $\phi_{W_{n_0}}$ be the canonical coupling of W_{n_0}. For every integer $m \in [n_0+1, -1]$ let $\phi_{W_{n_0}, \eta_{n_0+1}, \dots, \eta_m}$ be, as explained in Section 2.1, the cascaded permutation initiated by $\phi_{W_{n_0}}$ and sequentially obtained from $\eta_{n_0}, \dots, \eta_m$, and define $\eta'_{m+1} = \phi_{W_{n_0}, \eta_{n_0+1}, \dots, \eta_m}(\eta_{m+1})$ for every $m \in [n_0, -1]$. Then $(\eta'_n)_{n_0+1 \leqslant n \leqslant 0}$ has the same law as $(\eta_n)_{n_0+1 \leqslant n \leqslant 0}$ and each η'_n is, just as η_n, an innovation of $\mathcal{F}$, that is, η'_n is independent of $\mathcal{F}_{n-1}$ and $\mathcal{F}_n = \mathcal{F}_{n-1} \vee \sigma(\eta'_n)$. Moreover, with the notations above,*

$$\mathbb{P}\Big(W_n \neq Y_n(n_0)\Big) \to 0 \quad \text{as } n_0 \to -\infty$$

for every $n \leqslant 0$, where $\tilde{w}_{n_0}$ is the canonical word of length $|n_0|$.

Proof. It is easy to check that η'_n is an innovation as any other $\mathcal{F}_{n-1}$-measurable random permutation of η_n. The word W_n is a subword of W_{n_0}, and we denote by $Q_{n_0,n} \subset \{1, \dots, |n_0|\}$ the set of positions in the word W_{n_0} forming its subword W_n. Moreover, by construction of the cascaded permutations, $W'_n = Y_n(n_0)$ is a subword of $\tilde{w}_{n_0}$ and the set of positions in $\tilde{w}_{n_0}$ forming W'_n is the image of $Q_{n_0,n}$ by the canonical coupling $\phi_{W_{n_0}}$. We can check the intuitively clear fact that $Q_{n_0,n}$ is independent of W_{n_0} and is uniform on the subsets of $\{1, \dots, |n_0|\}$ having size $|n|$. Indeed, there is a bijective correspondence between the erasers $(\eta_{n_0+1}, \dots, \eta_n)$ and the $(n-n_0)$-uple listing the successive positions in the word W_{n_0} of the letters deleted at times $n_0+1, \dots, n$. The set of all theses positions is exactly the complement of $Q_{n_0,n}$ in $\{1, \dots, |n_0|\}$. Therefore there is a correspondance between $(\eta_{n_0+1}, \dots, \eta_n)$ and $Q_{n_0,n}$, and consequently there is independence between $Q_{n_0,n}$ and W_{n_0}. Moreover, this correspondance between $(\eta_{n_0+1}, \dots, \eta_n)$ and $Q_{n_0,n}$ is surjective and $(n-n_0)!$ to one, wherefrom follows the uniformity of the law of $Q_{n_0,n}$.

Now, to abbreviate notations, set $p = |n_0|$, $q = |n|$, $Q = Q_{n_0,n}$ and $W = W_{n_0}$. Thus we have seen that Q is a random variable independent of W and uniformly distributed on the subsets of $\{1, \dots, p\}$ having size q. With these abbreviated notations, the main statement of the lemma is rephrased by

$$\pi(p,q) := \mathbb{P}\Big(W_{|Q} = \widetilde{w}_{|\phi_W(Q)}\Big) \longrightarrow 1 \quad \text{as } p \to +\infty,$$

where ϕ_w is the canonical coupling of a word w and $\tilde{w}$ is the canonical word of length p, and 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.

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

$$E_1 = \Big\{\max(Q) \leqslant p - p^{3/4} - \kappa\Big\},$$
$$E_2 = \Big\{\forall (i,j) \in Q^2, i = j \text{ or } |i-j| \geqslant 3p^{3/4}\Big\},$$
$$E_3 = \Big\{\forall i \in Q, i - 1 - p^{3/4} \leqslant \kappa N_i^-(W) \leqslant i - 1 + p^{3/4}\Big\},$$

and we are going to show that

$$E_1 \cap E_2 \cap E_3 \subset \left\{W_{|Q} = \widetilde{w}_{|\phi_W(Q)}\right\}$$

if p is sufficiently large, and

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

We firstly show the inclusion. As a first step, we show that $W(i) = \widetilde{w}\Big(\phi_W(i)\Big)$ for every $i \in Q$ on the event $E_1 \cap E_3$. Consider $i \in Q$ and assume that $W(i) = a_r$. On the event E_3,

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

and $\kappa + i - 1 + p^{3/4} \leqslant p$ on the event E_1. Thus, by definition of the canonical coupling, $\phi_W(i) = r + \kappa N_i^-(W)$ on the event $E_1 \cap E_3$ and $W(i) = \widetilde{w}\Big(\phi_W(i)\Big)$ for every $i \in Q$. Moreover,

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

on the event E_3, therefore ϕ_W satisfies the following property on $E_1 \cap E_3$:

$$\forall i \in Q, \quad \Big|i - \phi_W(i)\Big| \leqslant \kappa - 1 + p^{3/4}.$$

Consequently, if we are on $E_1 \cap E_2 \cap E_3$ and if p is sufficiently large so that $3p^{3/4} \geqslant 2\Big(\kappa + p^{3/4}\Big)$, then the restriction of ϕ_W to Q is increasing, and finally $W_{|Q} = \widetilde{w}_{|\phi_W(Q)}$, as desired.

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

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

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

$$\begin{aligned}\mathbb{P}(E_2^c) &\leqslant \frac{6p^{3/4} + 2}{p - 1} + \frac{2(6p^{3/4} + 2)}{p - 2} + \cdots + \frac{(q - 1)(6p^{3/4} + 2)}{p - q + 1} \\ &\leqslant \frac{(q - 1)^2(6p^{3/4} + 2)}{p - q + 1} \xrightarrow[p\to\infty]{} 0.\end{aligned}$$

To find an upper bound of $\mathbb{P}(E_3^c)$, we call I_k the k-th element of Q for every $k \in \{1, \ldots, q\}$. Conditionally to $I_k = i$, the number of occurences $N_{I_k}^-(W)$ 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,

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

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

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

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

$$\mathbb{P}(E_3^c) < \frac{q\kappa}{\sqrt{p}} \xrightarrow[p\to\infty]{} 0,$$

and the proof is over. □ □

2.3 Proof of standardness

We finish to prove that the erased-word process $(W_n, \eta_n)_{n\leqslant 0}$ generates a filtration of product type in the discrete uniform case. This can be quickly proved from Lemma 2.1 with the help of *Vershik's first level criterion* and proposition 2.22 in [5]. Lemma 2.1 says that each random variable W_n satisfies Vershik's first level criterion. Since $(\eta_n)_{n\leqslant 0}$ is a process of innovations of $\mathcal{F}$, and since the σ-fields $\sigma(W_n, \eta_{n+1}, \dots, \eta_0)$ increase to $\mathcal{F}_0$ as $n \to -\infty$, proposition 2.22 in [5] ensures that $\mathcal{F}$ satisfies Vershik's first level criterion, and then $\mathcal{F}$ is of product type by Vershik's theorem (theorem 2.25 in [5])

But we have not stated Vershik's first level criterion in the present paper, and we can give a self-contained proof that $\mathcal{F}$ is of product type by constructing a generating process of innovations. First recall that $W'_n = Y_n(n_0)$ in Lemma 2.1 is measurable with respect to $\sigma(\eta'_{n_0+1}, \dots, \eta'_n)$. Then, given a sequence $(\delta_k)_{k\leqslant 0}$ of real numbers $\delta_k > 0$ satisfying $\delta_k \to 0$ as $k \to -\infty$, recursively applying Lemma 2.1 provides a strictly increasing sequence $(n_k)_{k\leqslant 0}$ of integers with $n_0 = -1$ and an innovation process $(\tilde{\eta}_n)_{n\leqslant 0}$ such that:

(i) $(\tilde{\eta}_{n_k+1}, \dots, \tilde{\eta}_n) = \tau_{W_{n_k},n}(\eta_{n_k+1}, \dots, \eta_n)$ for every $k < 0$ and every integer $n \in [n_k + 1, 0]$, where each $\tau_{w,n}$ is a permutation of $\{1, \dots, |n_k|\} \times \cdots \times \{1, \dots, |n| + 1\}$;

(ii) for every $k \leqslant 0$ 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 $(\tilde{\eta}_n)_{n\leqslant 0}$ generates $\mathcal{F}$. It suffices to construct, for each $n \leqslant 0$ 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 $\mathbb{P}\Big((W_n, \eta_n) \neq (\hat{W}_n, \hat{\eta}_n)\Big) < \delta$. To do so, let k be sufficiently small in order that $\delta_k < \delta$ and $n_k < n$. Then define

$$(\hat{\eta}_{n_k+1}, \dots, \hat{\eta}_n) = \tau^{-1}_{\tilde{W}_{n_k},n}(\tilde{\eta}_{n_k+1}, \dots, \tilde{\eta}_n)$$

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 $\hat{W}_{m+1} = f_{m+1}(\hat{W}_m, \hat{\eta}_{m+1})$ (the functions f_m were introduced before Lemma 2.1). 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.

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 $(W_n, \eta_n)_{n\leqslant 0}$ in the case when $A = [0,1]$ and μ is the Lebesgue measure on A, and denote by $\mathcal{G}$ the filtration it generates. In order to prove that $\mathcal{G}$ is of product type (step 2), the idea consists in approximating this process by an erased-word process on a finite alphabet with equiprobable letters, known to generate a filtration of product type by the previous section (step 1). Then the tools of Vershik's theory of filtrations will allow to conclude.

For every integer $k \geqslant 1$, let $f_k \colon A \to A$ be the function defined by $f_k(x) = 2^{-k}\lfloor 2^k x \rfloor$. Then f_k sends the Lebesgue measure μ to the uniform probability measure on the finite alphabet $A_k := \left\{0, \frac{1}{2^k}, \dots, \frac{2^k-1}{2^k}\right\}$. 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 $\left(f_k(W_n), \eta_n\right)_{n\leqslant 0}$ is an erased-word process generating a filtration of product type by Section 2. Moreover, denoting by $\mathcal{G}^k$ this filtration, the sequence of σ-fields $(\mathcal{G}_0^k)_{k\geqslant 1}$ is increasing and $\vee_{k=1}^{+\infty}\mathcal{G}_0^k = \mathcal{G}_0$

We give two ways to prove that $\mathcal{G}$ is of product type from the fact that each $\mathcal{G}^k$ is of product type. The first one uses *Vershik's first level criterion*, as the proof of the similar result 2.45 in [5] about the split-word processes. Vershik's first level criterion is known to be equivalent to productness (see [5]), hence we know it is satisfied by each filtration $\mathcal{G}^k$. Moreover, every innovation η'_n of $\mathcal{G}^k$ at time n is also an innovation of $\mathcal{G}$, because by lemma 2.4 in [5] it can be written $\eta'_n = \Phi(\eta_n)$ where Φ is a $\mathcal{G}_{n-1}^k$-measurable random bijection from $\{1, \dots, |n|+1\}$ to some finite set of size $|n|+1$. Thus, every random variable in $\cup_k L^1(\mathcal{G}_0^k)$ belongs to the set of random variables in $L^1(\mathcal{G}_0)$ satisfying Vershik's first level criterion with respect to $\mathcal{G}$. But this set is closed in $L^1(\mathcal{G}_0)$ by proposition 2.7 in [5], consequently $\mathcal{G}$ satisfies Vershik's first level criterion.

The second proof we give relies on a more general result stated in our original Proposition 3.1 below. As we have seen, the key point in the previous proof is the fact that every innovation of $\mathcal{G}^k$ is also an innovation of $\mathcal{G}$ and it is very specific to our situation. This fact implies that each $\mathcal{G}^k$ is immersed in $\mathcal{G}$ (see lemma 1.6 in [5]; that means here that the process $\left(f_k(W_n), \eta_n\right)_{n\leqslant 0}$ is Markovian with respect to $\mathcal{G}$), and this is the key point of the second proof.

Proposition 3.1. *Let $\mathcal{F}$ be a filtration. If there exists a sequence of Vershikian filtrations $(\mathcal{F}^k)_{k\geqslant 1}$ such that the sequence of σ-fields $(\mathcal{F}_0^k)_{k\geqslant 1}$ is increasing and satisfies $\vee_{k=1}^{+\infty}\mathcal{F}_0^k = \mathcal{F}_0$, and if each $\mathcal{F}^k$ is immersed in $\mathcal{F}$, then $\mathcal{F}$ is Vershikian.*

Proof. Saying that $\mathcal{F}^k$ is Vershikian means by definition that the final σ-field $\mathcal{F}_0^k$ is Vershikian with respect to the filtration $\mathcal{F}^k$, but thanks to lemma 4.1 in [7], this tantamounts to say that $\mathcal{F}_0^k$ is Vershikian with respect to the filtration $\mathcal{F}$ because of the immersion of $\mathcal{F}^k$ in $\mathcal{F}$. Now, because of $\vee_{k=1}^{+\infty}\mathcal{F}_0^k = \mathcal{F}_0$, lemma 4.2 in [7] (closedness of the set of Vershikian random variables) shows that $\mathcal{F}_0$ is Vershikian with respect to $\mathcal{F}$, that is to say $\mathcal{F}$ is Vershikian. □ □

Thus, we know that $\mathcal{G}$ is standard by Proposition 3.1 and by the equivalence between standardness and the Vershik property. By Theorem 1.4, we conclude that $\mathcal{G}$ is of product type.

Step 2 of the proof of Theorem 1.3 is achieved. Step 3 (the general case) is easily achieved with the help of Theorem 1.4. Consider an arbitrary Lebesgue alphabet (A, μ) and take a measurable function $f\colon [0,1] \to A$ sending the Lebesgue measure to μ. Then the process $\Big(f(W_n), \eta_n\Big)_{n\leqslant 0}$ 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 heritability of standardness under immersion, an immediate consequence of the definition of standardness (see [2] or [5]). Now step 3 is achieved and the proof of Theorem 1.3 is over.

4 Standardness of the multidimensional Pascal filtration

The d-dimensional Pascal filtration is introduced in [4]. It is the filtration generated by the Markov chain $(V_n)_{n\leqslant 0}$ whose distribution depends on a given probability vector $(\theta_1, \dots, \theta_d)$, where $d \geqslant 2$ is a finite integer or $d = \infty$, and is defined as follows:

- *(instantaneous distributions)* the random variable V_n has the multinomial distribution on
$$\mathrm{V}_n^d = \Big\{v \in \mathbb{N}^d \mid v(1) + \cdots + v(d) = |n|\Big\}$$
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 $V_n = \Big(v(1), \dots, v(d)\Big)$, coordinate i is picked at random with probability $\frac{v(i)}{|n|}$ 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 [4] for more detailed explanations.

It has been shown in [4] 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 the introduction and in Section 3. Indeed, taking the erased-word process $(W_n, \eta_n)_{n\leqslant 0}$ 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 $f_n\colon A^{|n|} \to \mathrm{V}_n$ as the one returning the list of the numbers of occurences of each letter of A in a given word of length $|n|$, then the process $\Big(f_n(W_n), \eta_n\Big)_{n\leqslant 0}$ is the

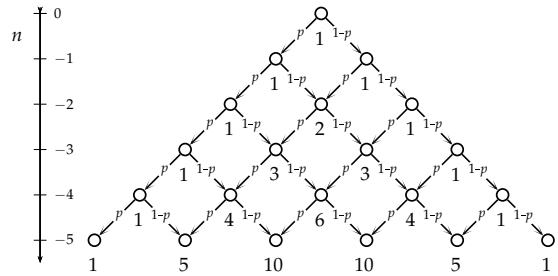

Fig. 5: 2-dimensional Pascal random walk, directed from $n = 0$ to $n = -\infty$

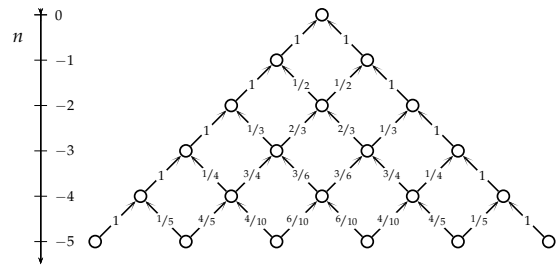

Fig. 6: 2-dimensional Pascal random walk, directed from $n = -\infty$ to $n = 0$

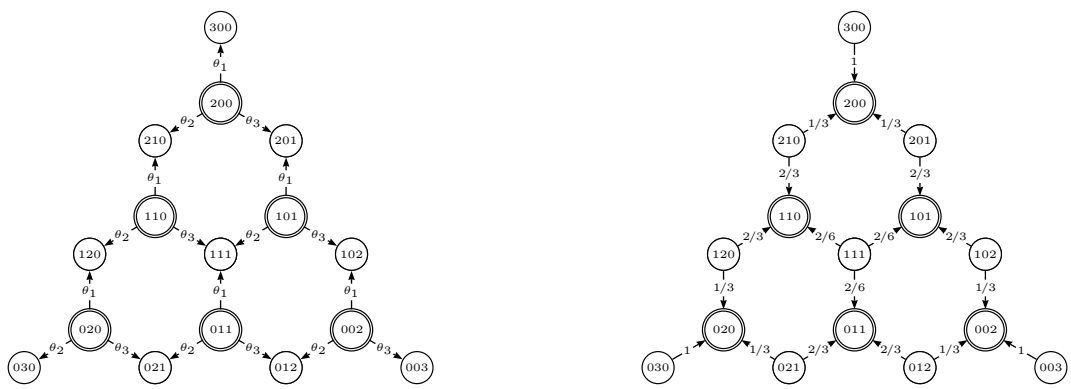

Fig. 7: Step in the 3-dimensional Pascal random walk: from $n = -2$ to $n = -3$ (left), and from $n = -3$ to $n = -2$ (right)

d-dimensional Pascal random walk defined by the probability vector $(\theta_1, \ldots, \theta_d)$, and the filtration $\mathcal{F}$ it generates is immersed in the filtration $\mathcal{G}$ generated by the erased-word process $(W_n, \eta_n)_{n \leqslant 0}$ because $\big(f_n(W_n), \eta_n\big)_{n \leqslant 0}$ is Markovian with respect to $\mathcal{G}$. Then standardness of the d-dimensional Pascal filtration $\mathcal{F}$ results from Theorem 1.3, from the obvious fact that standardness holds for filtrations of product type, and from the heritability property of standardness under immersion.

Acknowledgments. I thank Christophe Leuridan for his helpful comments on the first version of this paper, and the referee for his/her meticulous review which led to substantial improvements of the article.

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] D. Heicklen, C. Hoffman: $[T, T^{-1}]$ *is not standard*. Ergodic Theory and Dynamical Systems 18, 875–878 (1998)

[4] É. Janvresse, S. Laurent, T. de la Rue: *Standardness of the filtration of a monotonic Markov process*. arXiv:1501.02166 (2015). To appear in: Markov Processes and Related Fields.

[5] S. Laurent: *On standardness and I-cosiness*. Séminaire de Probabilités XLIII, Springer Lecture Notes in Mathematics 2006, 127–186 (2010)

[6] 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)

[7] S. Laurent: *On Vershikian and I-cosy random variables and filtrations*. Teoriya Veroyatnostei i ee Primeneniya 55, 104–132 (2010). Also published in: Theory Probab. Appl. 55, 54–76 (2011)

[8] 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)

[9] 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)

[10] A.M. Vershik: *Approximation in measure theory* (in Russian). PhD Dissertation, Leningrad University, 1973. (expanded and updated version: [11])

[11] 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)

[12] A.M. Vershik: *Intrinsic metric on graded graphs, standardness, and invariant measures*. Zapiski Nauchn. Semin. POMI 421, 58–67 (2014)

[13] A.M. Vershik: *The problem of describing central measures on the path spaces of graded graphs*. arXiv:1408.3291 (2013)