



**HAL**  
open science

## Standardness of monotonic Markov filtrations

Elise Janvresse, Stéphane Laurent, Thierry de La Rue

► **To cite this version:**

Elise Janvresse, Stéphane Laurent, Thierry de La Rue. Standardness of monotonic Markov filtrations. Markov Processes And Related Fields, 2016, 22 (4), pp.697-736. hal-01101631

**HAL Id: hal-01101631**

**<https://hal.science/hal-01101631>**

Submitted on 9 Jan 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.

# Standardness of monotonic Markov filtrations

Élise Janvresse<sup>\*1</sup>, Stéphane Laurent<sup>†2</sup>, and Thierry de la Rue<sup>‡1</sup>

<sup>1</sup>Laboratoire de Mathématiques Raphaël Salem, Normandie Université,  
Université de Rouen, CNRS

<sup>2</sup>Independent researcher

## Abstract

We derive a practical standardness criterion for the filtration generated by a monotonic Markov process. This criterion is applied to show standardness of some adic filtrations.

**Keywords:** Standardness of filtrations; monotonic Markov processes; Adic filtrations; Vershik's intrinsic metrics.

**MSC classification:** 60G05, 60J05, 60B99, 05C63.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Standardness . . . . .	2
1.2	Generating parameterization criterion . . . . .	3
<b>2</b>	<b>Standardness for the filtration of a Markov process</b>	<b>3</b>
2.1	Markov updating functions and the Propp-Wilson coupling . . . . .	4
2.2	Iterated Kantorovich pseudometrics and Vershik's criterion . . . . .	5
<b>3</b>	<b>Monotonic Markov processes</b>	<b>6</b>
3.1	Monotonic Markov processes and their representation . . . . .	6
3.2	Standardness criterion for monotonic Markov processes . . . . .	6
3.2.1	Tool 1: Convergence of $\mathcal{L}(X   \mathcal{F}_n)$ . . . . .	7
3.2.2	Tool 2: Ordered couplings and linear metrics . . . . .	8
3.3	Proof of Theorem 3.6 . . . . .	10
<b>4</b>	<b>Multidimensional monotonic Markov processes</b>	<b>11</b>
4.1	Monotonicity for multidimensional Markov processes . . . . .	11
4.2	Standardness for monotonic multidimensional Markov processes . . . . .	12
4.3	Computation of iterated Kantorovich metrics . . . . .	16
<b>5</b>	<b>Standardness of adic filtrations</b>	<b>17</b>
5.1	Adic filtrations and other filtrations on Bratteli graphs . . . . .	18
5.2	Vershik's intrinsic metrics . . . . .	20
<b>6</b>	<b>Pascal filtration</b>	<b>21</b>
6.1	Standardness . . . . .	21
6.2	Intrinsic metrics on the Pascal graph . . . . .	22

---

\*elise.janvresse@univ-rouen.fr

†laurent\_step@yahoo.fr

‡thierry.de-la-rue@univ-rouen.fr

<b>7 Euler filtration</b>	<b>22</b>
7.1 Standardness . . . . .	23
7.2 Intrinsic metrics on the Euler graph . . . . .	24
<b>8 Multidimensional Pascal filtration</b>	<b>24</b>
8.1 First proof of standardness, using monotonicity of multidimensional Markov processes . . . . .	26
8.2 Second proof of standardness, computing intrinsic metrics . . . . .	27
8.3 Third proof of standardness, constructing a generating parameterization . . . . .	27

## 1 Introduction

The theory of filtrations in discrete negative time was originally developed by Vershik in the 70's. It mainly deals with the identification of standard filtrations. Standardness is an invariant property of filtrations  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  in discrete negative time, whose definition is recalled below (Definition 1.1). It only concerns the case when the  $\sigma$ -field  $\mathcal{F}_0$  is essentially separable, and in this situation one can always find a Markov process<sup>1</sup>  $(X_n)_{n \leq 0}$  that generates the filtration  $\mathcal{F}$  by taking for  $X_n$  any random variable generating the  $\sigma$ -field  $\mathcal{F}_n$  for every  $n \leq 0$ .

In Section 2, we provide two standardness criteria for a filtration given as generated by a Markov process. The first one, Lemma 2.1, is a somewhat elementary criterion involving a construction we call the *Propp-Wilson coupling* (Section 2.1). The second one, Lemma 2.5, is borrowed from [17]. It is a particular form of Vershik's standardness criterion which is known to be equivalent to standardness (see [10]).

The main result of this paper is stated and proved in Section 3 (Theorem 3.6): It provides a very convenient standardness criterion for filtrations which are given as generated by a *monotonic* Markov process  $(X_n)_{n \leq 0}$  (see Definition 3.3). It is generalized in Section 4 (Theorem 4.5) to multidimensional Markov processes.

There is a revival interest in standardness due to the recent works of Vershik [28, 29, 30] which connect the theory of filtrations to the problem of identifying ergodic central measures on Bratteli graphs, which is itself closely connected to other problems of mathematics. As we explain in Section 5, an ergodic central measure on (the path space of) a Bratteli graph generates a filtration we call an *adic filtration*, and the recent discoveries by Vershik mainly deal with standardness of adic filtrations. Using our standardness criterion for the filtration of a monotonic Markov process, we show standardness for some adic filtrations arising from the Pascal graph and the Euler graph in the subsequent sections 6, 7 and 8. As a by-product, our results also provide a new proof of ergodicity of some adic transformations on these graphs. We also discuss the case of non-central measures.

### 1.1 Standardness

A filtration  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  is said to be *immersed* in a filtration  $\mathcal{G} = (\mathcal{G}_n)_{n \leq 0}$  if  $\mathcal{F} \subset \mathcal{G}$  and for each  $n < 0$ , the  $\sigma$ -field  $\mathcal{F}_{n+1}$  is conditionally independent of  $\mathcal{G}_n$  given  $\mathcal{F}_n$ . When  $\mathcal{F}$  is the filtration generated by a Markov process  $(X_n)_{n \leq 0}$ , then saying that  $\mathcal{F}$  is immersed in some filtration  $\mathcal{G}$  tantamounts to say that  $\mathcal{F} \subset \mathcal{G}$  and that  $(X_n)_{n \leq 0}$  has the Markov property with respect to the bigger filtration  $\mathcal{G}$ , that is,

$$\mathcal{L}(X_{n+1} | \mathcal{G}_n) = \mathcal{L}(X_{n+1} | \mathcal{F}_n) = \mathcal{L}(X_{n+1} | X_n)$$

for every  $n < 0$ .

A filtration is said to be of *product type* if it is generated by a sequence of independent random variables.

**Definition 1.1.** A filtration  $\mathcal{F}$  is said to be *standard* when it is immersed in a filtration of product type, possibly up to isomorphism (in which case we say that  $\mathcal{F}$  is *immersible* in a filtration of product type).

---

<sup>1</sup>By *Markov process* we mean any stochastic process  $(X_n)_{n \leq 0}$  satisfying the Markov property, but no stationarity and no homogeneity in time are required.

When  $(X_n)_{n \leq 0}$  is any stochastic process generating the filtration  $\mathcal{F}$ , then a filtration isomorphic to  $\mathcal{F}$  is a filtration generated by a *copy* of  $(X_n)_{n \leq 0}$ , that is to say a stochastic process  $(X'_n)_{n \leq 0}$  defined on any probability space and having the same law as  $(X_n)_{n \leq 0}$ .

By Kolmogorov's 0-1 law, a necessary condition for standardness is that the filtration  $\mathcal{F}$  be *Kolmogorovian*, that is to say that the tail  $\sigma$ -algebra  $\mathcal{F}_{-\infty}$  be degenerate<sup>2</sup>.

## 1.2 Generating parameterization criterion

We prove in this section that a filtration having a generating parameterization is standard, after introducing the required definitions. Constructing a generating parameterization is a frequent way to establish standardness in practice.

**Definition 1.2.** Let  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  be a filtration. A *parameterization* of  $\mathcal{F}$  is a sequence of (independent) random variables  $\mathbf{U} = (U_n)_{n \leq 0}$  such that for each  $n \leq 0$ , the random variable  $U_n$  is independent of  $\mathcal{F}_{n-1} \vee \sigma(U_m; m \leq n-1)$ , and satisfies  $\mathcal{F}_n \subset \mathcal{F}_{n-1} \vee \sigma(U_n)$ . We say that the parameterization  $\mathbf{U}$  is *generating* if  $\mathcal{F} \subset \mathcal{U}$ , where  $\mathcal{U}$  is the filtration generated by  $\mathbf{U}$ .

It is shown in [13] that, up to isomorphism, every filtration  $\mathcal{F}$  having an essentially separable  $\sigma$ -field  $\mathcal{F}_0$  has a parameterization  $(U_n)_{n \leq 0}$  where each  $U_n$  has a uniform distribution on  $[0, 1]$ .

The following lemma is shown in [14]. It is the key point to show that a filtration having a generating parameterization is standard (Lemma 1.5).

**Lemma 1.3.** *Let  $\mathcal{F}$  be a filtration having a parameterization  $\mathbf{U} = (U_n)_{n \leq 0}$ , and let  $\mathcal{U} = (\mathcal{U}_n)_{n \leq 0}$  be the filtration generated by  $\mathbf{U}$ . Then  $\mathcal{F}$  and  $\mathcal{U}$  are both immersed in the filtration  $\mathcal{F} \vee \mathcal{U}$ .*

**Definition 1.4.** The filtration  $\mathcal{F} \vee \mathcal{U}$  in the above lemma is called the *extension of  $\mathcal{F}$  with the parameterization  $\mathbf{U}$* , and is also said to be a *parametric extension* of  $\mathcal{F}$ .

**Lemma 1.5.** *If  $\mathbf{U}$  is a generating parameterization of the filtration  $\mathcal{F}$ , then  $\mathcal{F}$  as well as  $\mathcal{F} \vee \mathcal{U}$  are standard.*

*Proof.* Obviously  $\mathcal{F} \vee \mathcal{U}$  is standard because  $\mathcal{U}$  is standard (even of product type), and  $\mathcal{F} \vee \mathcal{U} = \mathcal{U}$  under the generating assumption. Then the filtration  $\mathcal{F}$  is standard as well, because by Lemma 1.3 it is immersed in the filtration  $\mathcal{U}$ .  $\square$

Whether any standard filtration admits a generating parameterization is an open question of the theory of filtrations.

## 2 Standardness for the filtration of a Markov process

From now on, we consider a Markov process  $(X_n)_{n \leq 0}$  where, for each  $n$ ,  $X_n$  takes its values in a standard Borel space  $A_n$ , and whose transition probabilities are given by the sequence of kernels  $(P_n)_{n \leq 0}$ : For each  $n \leq 0$  and each measurable subset  $E \subset A_n$ ,

$$\mathbb{P}(X_n \in E \mid X_{n-1}) = P_n(X_{n-1}, E) \quad a.s.$$

We denote by  $\mathcal{F}$  the filtration generated by  $(X_n)_{n \leq 0}$ . In this section, we provide two practical criteria to establish standardness of  $\mathcal{F}$ : the Propp-Wilson coupling in Section 2.1 (Lemma 2.1) and a simplified form of Vershik's standardness criterion in Section 2.2 (Lemma 2.5, borrowed from [17]). Recall that any filtration having an essentially separable final  $\sigma$ -field  $\mathcal{F}_0$  can always be generated by a Markov process  $(X_n)_{n \leq 0}$ . But practicality of the standardness criteria we present in this section lies on the choice of the generating Markov process.

---

<sup>2</sup>The introduction of the word *Kolmogorovian* firstly occurred in [13] and [14] and was motivated by the so-called Kolmogorov's 0-1 law in the case of a product type filtration. By the correspondance between  $(-\mathbb{N})$ -indexed filtrations and  $\mathbb{N}$ -indexed decreasing sequences of measurables partitions, one could also say *ergodic*, because this property is equivalent to ergodicity of the equivalence relation defined by the tail partition.

The Propp-Wilson coupling is a practical criterion to construct a generating parameterization of  $\mathcal{F}$ . It will be used to prove our standardness criterion for monotonic Markov processes (Theorem 3.6) which is the main result of this article. The simplified form of Vershik's standardness criterion we provide in Lemma 2.5 will not be used to prove Theorem 3.6, but the iterated Kantorovich pseudometrics  $\rho_n$  introduced to state this criterion will play an important role in the proof of Theorem 3.6, and they will also appear in Section 5 as the *intrinsic metrics* in the particular context of adic filtrations. Lemma 2.5 itself will only be used in section 8.

The general statement of Vershik's standardness criterion concerns an arbitrary filtration  $\mathcal{F}$  and it is known to be equivalent to standardness as long as the final  $\sigma$ -field  $\mathcal{F}_0$  is essentially separable. Its statement is simplified in Lemma 2.5, mainly because it is specifically stated for the case when  $\mathcal{F}$  is the filtration of the Markov process  $(X_n)_{n \leq 0}$ , together with an identifiability assumption on the Markov kernels  $P_n$ .

## 2.1 Markov updating functions and the Propp-Wilson coupling

For the filtration  $\mathcal{F}$  generated by the Markov process  $(X_n)_{n \leq 0}$ , it is possible to have, up to isomorphism, a parameterization  $(U_n)_{n \leq 0}$  of  $\mathcal{F}$  with the additional property

$$\sigma(X_n) \subset \sigma(X_{n-1}, U_n) \quad \text{for each } n \leq 0.$$

This fact is shown in [14] but we will consider it from another point of view here. The above inclusion means that  $X_n = f_n(X_{n-1}, U_n)$  for some measurable function  $f_n$ . Such a function is appropriate when it is an *updating function* of the Markov kernel  $P_n$ , that is to say a measurable function  $f_n: (x, u) \mapsto f_n(x, u) \in A_n$  such that  $f_n(x, \cdot)$  sends the distribution law of  $U_n$  to  $P_n(x, \cdot)$  for each  $x \in A_{n-1}$ .

Such updating functions, associated to random variables  $U_n$  which are uniformly distributed in  $[0, 1]$ , always exist. Indeed, there is no loss of generality to assume that each  $X_n$  takes its values in  $\mathbb{R}$ . Then, the most common choice of  $f_n$  is the *quantile updating function*, defined as the inverse of the right-continuous cumulative distribution function of the conditional law  $\mathcal{L}(X_n | X_{n-1} = x) = P_n(x, \cdot)$ :

$$\text{For } 0 < u < 1, \quad f_n(x, u) = \inf \{t \in \mathbb{R} : \mathbb{P}(X_n \leq t | X_{n-1} = x) \geq u\}. \quad (2.1)$$

Once the updating functions  $f_n$  are given, it is not difficult to get, up to isomorphism, a parameterization  $(U_n)_{n \leq 0}$  for which  $X_n = f_n(X_{n-1}, U_n)$ , using the Kolmogorov extension theorem. We then say that  $(X_n)_{n \leq 0}$  is *parameterized by*  $(f_n, U_n)_{n \leq 0}$  and that  $(f_n, U_n)_{n \leq 0}$  is a *parametric representation* of  $(X_n)_{n \leq 0}$ .

Given a parametric representation  $(f_n, U_n)_{n \leq 0}$  of  $(X_n)_{n \leq 0}$ , the *Propp-Wilson coupling* is a practical tool to check whether  $(U_n)_{n \leq 0}$  is a generating parameterization of the filtration  $\mathcal{F}$  generated by  $(X_n)_{n \leq 0}$ . Given  $n_0 \leq -1$  and a point  $x_{n_0}$  in  $A_{n_0}$ , there is a natural way to construct, on the same probability space, a Markov process  $(Y_n(n_0, x_{n_0}))_{n_0 \leq n \leq 0}$  with initial condition  $Y_{n_0}(n_0, x_{n_0}) = x_{n_0}$  and having the same transition kernels as  $(X_n)_{n_0 \leq n \leq 0}$ : It suffices to set the initial condition  $Y_{n_0}(n_0, x_{n_0}) = x_{n_0}$  and to use the inductive relation

$$\forall n_0 \leq n < 0, \quad Y_{n+1}(n_0, x_{n_0}) := f_{n+1}(Y_n(n_0, x_{n_0}), U_{n+1}).$$

We call this construction the *Propp-Wilson coupling* because it is a well-known construction used in Propp and Wilson's coupling-from-the-past algorithm [22]. The word "coupling" refers to the fact that the random variables  $Y_n$  are constructed on the same probability space as the Markov process  $(X_n)_{n \leq 0}$ . The following lemma shows how to use the Propp-Wilson coupling to prove the generating property of  $(U_n)_{n \leq 0}$ .

**Lemma 2.1.** *Assume that, for every  $n \leq 0$ , the state space  $A_n$  of  $X_n$  is Polish under some distance  $d_n$  and that  $\mathbb{E}[d_n(X_n, Y_n(m, x_m))] \rightarrow 0$  as  $m \rightarrow -\infty$  for some sequence  $(x_m)$  (possibly depending on  $n$ ) such that  $x_m \in A_m$ . Then  $(U_n)_{n \leq 0}$  is a generating parameterization of the filtration  $\mathcal{F}$  generated by  $(X_n)_{n \leq 0}$ . In particular,  $\mathcal{F}$  is standard.*

*Proof.* The assumption implies that every  $X_n$  is measurable with respect to  $\sigma(\dots, U_{n-1}, U_n)$  because  $Y_n(m, x_m)$  is  $\sigma(U_{m+1}, \dots, U_n)$ -measurable. Then it is easy to check that  $(U_n)_{n \leq 0}$  is a generating parameterization of  $\mathcal{F}$ .  $\square$

## 2.2 Iterated Kantorovich pseudometrics and Vershik's criterion

Vershik's standardness criterion will only be necessary to prove the second multidimensional version of Theorem 3.6 (Theorem 4.8). However the iterated Kantorovich pseudometrics lying at the heart of Vershik's standardness will be used in the proof of Theorem 3.6.

A *coupling* of two probability measures  $\mu$  and  $\nu$  is a pair  $(X_\mu, X_\nu)$  of two random variables defined on the same probability space with respective distribution  $\mu$  and  $\nu$ . When  $\mu$  and  $\nu$  are defined on the same separable metric space  $(E, \rho)$ , the *Kantorovich distance* between  $\mu$  and  $\nu$  is defined by

$$\rho'(\mu, \nu) := \inf \mathbb{E}[\rho(X_\mu, X_\nu)], \quad (2.2)$$

where the infimum is taken over all couplings  $(X_\mu, X_\nu)$  of  $\mu$  and  $\nu$ .

If  $(E, \rho)$  is compact, the weak topology on the set of probability measures on  $E$  is itself compact and metrized by the Kantorovich metric  $\rho'$ . If  $\rho$  is only a pseudometric on  $E$ , one can define  $\rho'$  in the same way, but we only get a pseudometric on the set of probability measures.

The iterated Kantorovich pseudometrics  $\rho_n$  defined below arise from the translations of Vershik's ideas [27] into the context of our Markov process  $(X_n)_{n \leq 0}$ . Let  $n_0 \leq 0$  be an integer and assume that we are given a compact pseudometric  $\rho_{n_0}$  on the state space  $A_{n_0}$  of  $X_{n_0}$ . Then for every  $n \leq n_0$  we recursively define a compact pseudometric  $\rho_n$  on the state space  $A_n$  of  $X_n$  by setting

$$\rho_n(x_n, x'_n) := (\rho_{n+1})'(P_n(x_n, \cdot), P_n(x'_n, \cdot))$$

where  $(\rho_{n+1})'$  is the Kantorovich pseudometric derived from  $\rho_{n+1}$  as explained above.

**Definition 2.2.** With the above notations, we say that the random variable  $X_{n_0}$  satisfies the *V' property* if  $\mathbb{E}[\rho_n(X'_n, X''_n)] \rightarrow 0$  where  $X'_n$  and  $X''_n$  are two independent copies of  $X_n$ .

Note that the *V' property* of  $X_{n_0}$  is not only a property of the random variable  $X_{n_0}$  alone, since its statement relies on the Markov process  $(X_n)_{n \leq 0}$ . Actually the *V' property* of  $X_{n_0}$  is a rephrasing of the *Vershik property* (not stated in the present paper) of  $X_{n_0}$  with respect to the filtration  $\mathcal{F}$  generated by  $(X_n)_{n \leq 0}$ , in the present context when  $(X_n)_{n \leq 0}$  is a Markov process. The equivalence between these two properties is shown in [17], but in the present paper we do not introduce the general Vershik property. The definition also relies on the choice of the initial compact pseudometric  $\rho_{n_0}$ , but it is shown in [14] and [17] that the Vershik property of  $X_{n_0}$  (with respect to  $\mathcal{F}$ ) and actually is a property about the  $\sigma$ -field  $\sigma(X_{n_0})$  generated by  $X_{n_0}$  and thus it does not depend on  $\rho_{n_0}$ . Admitting this equivalence between the *V' property* and the Vershik property, and using proposition 6.2 in [14], we get the following proposition.

**Proposition 2.3.** *The filtration generated by the Markov process  $(X_n)_{n \leq 0}$  is standard if and only if  $X_n$  satisfies the *V' property* for every  $n \leq 0$ .*

As shown in [17], there is a considerable simplification of Proposition 2.3 under the identifiability condition defined below. This is rephrased in Lemma 2.5.

**Definition 2.4.** A Markov kernel  $P$  is *identifiable* when  $x \mapsto P(x, \cdot)$  is one-to-one. A Markov process  $(X_n)_{n \leq 0}$  is *identifiable* if for every  $n \leq 0$  its transition distributions  $\mathcal{L}(X_n | X_{n-1} = x)$  are given by an identifiable Markov kernel  $P_n$ .

If  $\rho_{n_0}$  is a metric and the Markov process is identifiable, then it is easy to prove by induction that  $\rho_n$  is a metric for all  $n \leq n_0$ , using the fact that  $(\rho_{n+1})'$  is itself a metric. Lemma 2.5 below, borrowed from [17], provides a friendly statement of Vershik's standardness criterion for the filtration of an identifiable Markov process.

**Lemma 2.5.** Let  $(X_n)_{n \leq 0}$  be an identifiable Markov process with  $X_0$  taking its values in a compact metric space  $(A_0, \rho_0)$ . Then the filtration generated by  $(X_n)_{n \leq 0}$  is standard if and only if  $X_0$  satisfies the  $V'$  property.

### 3 Monotonic Markov processes

Theorem 3.6 in Section 3.2 provides a simple standardness criterion for the filtration of a monotonic Markov process. After defining this kind of Markov processes, we introduce a series of tools before proving the theorem. An example is provided in this section (the Poissonian Markov chain), and examples of adic filtrations will be provided in Section 5.

#### 3.1 Monotonic Markov processes and their representation

**Definition 3.1.** Let  $\mu$  and  $\nu$  be two probability measures on the same ordered set, we say that the coupling  $(X_\mu, X_\nu)$  of  $\mu$  and  $\nu$  is an *ordered coupling* if  $\mathbb{P}(X_\mu \leq X_\nu) = 1$  or  $\mathbb{P}(X_\nu \leq X_\mu) = 1$ .

**Definition 3.2.** Let  $\mu$  and  $\nu$  be two probability measures on an ordered set. We say that  $\mu$  is *stochastically dominated* by  $\nu$ , and note  $\mu \stackrel{\text{st}}{\leq} \nu$ , if there exists an ordered coupling  $(X_\mu, X_\nu)$  such that  $X_\mu \leq X_\nu$  a.s.

**Definition 3.3.** • When  $A$  and  $B$  are ordered, a Markov kernel  $P$  from  $A$  to  $B$  is *increasing* if  $x \leq x' \implies P(x, \cdot) \stackrel{\text{st}}{\leq} P(x', \cdot)$ .

- Let  $(X_n)_{n \leq 0}$  be a Markov process such that each  $X_n$  takes its values in an ordered set. We say that  $(X_n)_{n \leq 0}$  is *monotonic* if the Markov kernel  $P_n(x, \cdot) := \mathcal{L}(X_n | X_{n-1} = x)$  is increasing for each  $n$ .

**Example 3.4 (Poissonian Markov chain).** Given a decreasing sequence  $(\lambda_n)_{n \leq 0}$  of positive real numbers, define the law of a Markov process  $(X_n)_{n \leq 0}$  by:

- (*Instantaneous laws*) each  $X_n$  has the Poisson distribution with mean  $\lambda_n$ ;
- (*Markovian transition*) given  $X_n = k$ , the random variable  $X_{n+1}$  has the binomial distribution on  $\{0, \dots, k\}$  with success probability parameter  $\lambda_{n+1}/\lambda_n$ .

It is easy to check that the binomial distribution  $\mathcal{L}(X_{n+1} | X_n = k)$  is stochastically increasing in  $k$ , hence  $(X_n)_{n \leq 0}$  is a monotonic Markov process. Note that it is identifiable (Definition 2.4).

The notion of updating function for a Markov kernel has been introduced in Section 2.1. Below we define the notion of increasing updating function, in the context of monotonic Markov kernels.

**Definition 3.5.** • Let  $P$  be an (increasing) Markov kernel from  $A$  to  $B$  and  $f$  be an updating function of  $P$ . We say that  $f$  is an *increasing updating function* if  $f(x, u) \leq f(x', u)$  for almost all  $u$  and for every  $x, x' \in A$  satisfying  $x \leq x'$ .

- We say that a parameterization  $(f_n, U_n)_{n \leq 0}$  (defined in Section 2.1) of a (monotonic) Markov process is an *increasing representation* if every  $f_n$  is an increasing updating function, that is, the equality  $f_n(x, U_n) \leq f_n(x', U_n)$  almost surely holds whenever  $x \leq x'$ .

For a real-valued monotonic Markov process, it is easy to check that the quantile updating functions defined by (2.1) provide an increasing representation.

#### 3.2 Standardness criterion for monotonic Markov processes

The achievement of the present section is the following Theorem which provides a practical criterion to check standardness of a filtration generated by a monotonic Markov process.

**Theorem 3.6.** Let  $(X_n)_{n \leq 0}$  be an  $\mathbb{R}$ -valued monotonic Markov process, and  $\mathcal{F}$  the filtration it generates.

1) The following conditions are equivalent.

- (a)  $\mathcal{F}$  is standard.
- (b)  $\mathcal{F}$  admits a generating parameterization.
- (c) Every increasing representation provides a generating parameterization.
- (d) For every  $n \leq 0$ , the conditional law  $\mathcal{L}(X_n | \mathcal{F}_{-\infty})$  is almost surely equal to  $\mathcal{L}(X_n)$ .
- (e)  $\mathcal{F}$  is Kolmogorovian.

2) Assuming that the Markov process is identifiable (Definition 2.4), then the five conditions above are equivalent to the almost-sure equality between the conditional law  $\mathcal{L}(X_0 | \mathcal{F}_{-\infty})$  and  $\mathcal{L}(X_0)$ .

Before giving the proof of the theorem, we isolate the main tools that we will use.

### 3.2.1 Tool 1: Convergence of $\mathcal{L}(X | \mathcal{F}_n)$

Lemma 3.8 is somehow a rephrasing of Lévy's reversed martingale convergence theorem. It says in particular that condition (d) of Theorem 3.6 is the same as the convergence  $\mathcal{L}(X_n | \mathcal{F}_m) \xrightarrow{m \rightarrow -\infty} \mathcal{L}(X_n)$ . We state a preliminary lemma which will also be used in Section 4.2.

Given, on some probability space, a  $\sigma$ -field  $\mathcal{B}$  and a random variable  $X$  taking its values in a Polish space  $A$ , the conditional law  $\mathcal{L}(X | \mathcal{B})$  is a random variable when the narrow topology is considered on the space of probability measures on  $A$ , and this topology coincides with the topology of weak convergence when  $A$  is compact (see [1]).

**Lemma 3.7.** *Let  $A$  be a compact metric space and  $(\Gamma_k)_{k \geq 0}$  a sequence of random variables taking values in the space of probability measures on  $A$  equipped with the topology of weak convergence. Then the sequence  $(\Gamma_k)_{k \geq 0}$  almost surely converges to a random probability measure  $\Gamma_\infty$  if and only if, for every continuous function  $f: A \rightarrow \mathbb{R}$ ,  $\Gamma_k(f)$  almost surely converges to  $\Gamma_\infty(f)$ .*

*Proof.* The "only if" part is obvious. Conversely, if for each continuous function  $f: A \rightarrow \mathbb{R}$ ,  $\Gamma_k(f)$  almost surely converges to  $\Gamma_\infty(f)$ , then the full set of convergence can be taken independently of  $f$  by using the separability of the space of continuous functions on  $A$ . This shows the almost sure weak convergence  $\Gamma_k \rightarrow \Gamma_\infty$  (see [1] or [6] for details).  $\square$

Recall that  $\rho'$  denotes the Kantorovich metric (defined by (2.2)) induced by  $\rho$ .

**Lemma 3.8.** *Let  $\mathcal{F}$  be a filtration and  $X$  an  $\mathcal{F}_0$ -measurable random variable taking its values in a compact metric space  $(A, \rho)$ . Then one always has the almost sure convergence as well as the  $L^1$ -convergence  $\mathcal{L}(X | \mathcal{F}_n) \rightarrow \mathcal{L}(X | \mathcal{F}_{-\infty})$ , i.e.*

$$\rho'(\mathcal{L}(X | \mathcal{F}_n), \mathcal{L}(X | \mathcal{F}_{-\infty})) \xrightarrow{n \rightarrow -\infty} 0 \quad \text{almost surely}$$

and

$$\mathbb{E} \left[ \rho'(\mathcal{L}(X | \mathcal{F}_n), \mathcal{L}(X | \mathcal{F}_{-\infty})) \right] \xrightarrow{n \rightarrow -\infty} 0.$$

*Proof.* By Lévy's reversed martingale convergence theorem, the convergence  $\mathbb{E}[f(X) | \mathcal{F}_n] \rightarrow \mathbb{E}[f(X) | \mathcal{F}_{-\infty}]$  holds almost surely for every continuous functions  $f: A \rightarrow \mathbb{R}$ . The almost sure weak convergence  $\mathcal{L}(X | \mathcal{F}_n) \rightarrow \mathcal{L}(X | \mathcal{F}_{-\infty})$  follows from Lemma 3.7. Since the Kantorovich distance metrizes the weak convergence, we get the almost sure convergence of  $\rho'(\mathcal{L}(X | \mathcal{F}_n), \mathcal{L}(X | \mathcal{F}_{-\infty}))$  to 0, as well as the  $L^1$ -convergence by the dominated convergence theorem.  $\square$

**Example (Poissonian Markov chain).** Consider Example 3.4. We are going to determine the conditional law  $\mathcal{L}(X_0 | \mathcal{F}_{-\infty})$ . For every  $n \leq -1$ , the conditional law  $\mathcal{L}(X_0 | \mathcal{F}_n)$  is the binomial distribution on  $\{0, \dots, X_n\}$  with success probability parameter  $\theta_n := \lambda_0 / \lambda_n$ . Since  $(X_n)_{n \leq 0}$  is decreasing,  $X_n$  almost surely goes to a random variable  $X_{-\infty}$  taking its values in  $\mathbb{N} \cup \{+\infty\}$ .

- **Case 1:**  $\lambda_n \rightarrow \lambda_{-\infty} < \infty$ . In this case, it is easy to see with the help of Fourier transforms that  $X_{-\infty}$  has the Poisson distribution with mean  $\lambda_{-\infty}$ . And by Lemma 3.8,  $\mathcal{L}(X_0 | \mathcal{F}_{-\infty})$  is the binomial distribution on  $\{0, \dots, X_{-\infty}\}$  with success probability parameter  $\lambda_0/\lambda_{-\infty}$ .
- **Case 2:**  $\lambda_n \rightarrow +\infty$ . In this case,  $X_n$  almost surely goes to  $+\infty$ . Indeed,  $\mathbb{P}(X_{-\infty} > K) \geq \mathbb{P}(X_n > K) \rightarrow 1$  for any  $K > 0$ . By the well-known Poisson approximation to the binomial distribution, it is expected that  $\mathcal{L}(X_0 | \mathcal{F}_n)$  should be well approximated by the Poisson distribution with mean  $X_n\theta_n$  and then that  $\mathcal{L}(X_0 | \mathcal{F}_{-\infty})$  should be the deterministic Poisson distribution with mean  $\lambda_0$  (that is, the law of  $X_0$ ). We prove it using Lemma 3.8. Let  $\mathcal{L}_n := \mathcal{L}(X_0 | \mathcal{F}_n)$ , denote by  $\mathcal{P}(\lambda)$  the Poisson distribution with mean  $\lambda$  and by  $\text{Bin}(k, \theta)$  the binomial distribution on  $\{0, \dots, k\}$  with success probability parameter  $\theta$ . Let  $\rho$  be the discrete distance on the state space  $\mathbb{N}$  of  $X_0$ . By introducing an appropriate coupling of  $\mathcal{P}(\lambda)$  and  $\text{Bin}(k, \theta)$ , as described in the introduction of [19], it is not difficult to prove that

$$\rho'(\text{Bin}(k, \theta), \mathcal{P}(k\theta)) \leq k\theta^2.$$

By applying this result,

$$\rho'(\mathcal{L}_n, \mathcal{P}(X_n\theta_n)) \leq X_n\theta_n^2 = \frac{X_n \lambda_0^2}{\lambda_n \lambda_n}.$$

Hence

$$\mathbb{E}[\rho'(\mathcal{L}_n, \mathcal{P}(X_n\theta_n))] \rightarrow 0. \quad (3.1)$$

On the other hand, for every  $\lambda \geq \lambda' > 0$ , using the fact that  $\mathcal{P}(\lambda) = \mathcal{P}(\lambda') * \mathcal{P}(\lambda - \lambda')$ , it is easy to derive the inequality

$$\rho'(\mathcal{P}(\lambda), \mathcal{P}(\lambda')) \leq 1 - \exp(\lambda' - \lambda) \leq |\lambda - \lambda'|.$$

Thus

$$\rho'(\mathcal{P}(X_n\theta_n), \mathcal{P}(\lambda_0)) \leq |X_n\theta_n - \lambda_0|.$$

Since  $\text{Var}(X_n\theta_n) = \theta_n^2\lambda_n = \lambda_0^2/\lambda_n \rightarrow 0$ , we get by Tchebychev's inequality,  $X_n\theta_n \rightarrow \lambda_0$  in probability, which implies that

$$\mathbb{E}[\rho'(\mathcal{P}(X_n\theta_n), \mathcal{P}(\lambda_0))] \rightarrow 0.$$

Together with (3.1), this yields

$$\mathbb{E}[\rho'(\mathcal{L}_n, \mathcal{P}(\lambda_0))] \rightarrow 0.$$

Comparing with Lemma 3.8, we get, as expected,  $\mathcal{L}(X_0 | \mathcal{F}_{-\infty}) = \mathcal{P}(\lambda_0)$ .

The second assertion of Theorem 3.6 shows that the Poissonian Markov chain generates a standard filtration when  $\lambda_n \rightarrow +\infty$ , and a non-Kolmogorovian filtration otherwise.

### 3.2.2 Tool 2: Ordered couplings and linear metrics

**Lemma 3.9.** *Let  $\mu, \nu$  and  $\eta$  be probability measures defined on an ordered set  $E$  such that  $\mu \stackrel{st}{\leq} \nu$  and  $\nu \stackrel{st}{\leq} \eta$ . Then we can find three random variables  $X_\mu, X_\nu, X_\eta$  on the same probability space, with respective distribution  $\mu, \nu$  and  $\eta$ , such that  $X_\mu \leq X_\nu \leq X_\eta$  a.s. In particular,  $\mu \stackrel{st}{\leq} \eta$ .*

*Proof.* Let us consider three copies  $E_1, E_2, E_3$  of  $E$ . Since  $\mu \stackrel{st}{\leq} \nu$ , we can find a probability measure  $\mathbb{P}_{\mu, \nu}$  on  $E_1 \times E_2$  which is a coupling of  $\mu$  and  $\nu$ , such that  $\mathbb{P}_{\mu, \nu}(\{(x_1, x_2) : x_1 \leq x_2\}) = 1$ . In the same way, we can find a probability measure  $\mathbb{P}_{\nu, \eta}$  on  $E_2 \times E_3$  which is a coupling

of  $\nu$  and  $\eta$ , such that  $\mathbb{P}_{\nu,\eta}(\{(x_2, x_3) : x_2 \leq x_3\}) = 1$ . We consider the relatively independent coupling of  $\mathbb{P}_{\mu,\nu}$  and  $\mathbb{P}_{\nu,\eta}$  over  $E_2$ , which is the probability measure on  $E_1 \times E_2 \times E_3$ , defined by

$$\mathbb{P}(A \times B \times C) := \int_B d\nu(x) \mathbb{P}_{\mu,\nu}(A \times E_2 | x_2 = x) \mathbb{P}_{\nu,\eta}(E_2 \times C | x_2 = x).$$

Under  $\mathbb{P}$ , the pair  $(x_1, x_2)$  follows  $\mathbb{P}_{\mu,\nu}$  and the pair  $(x_2, x_3)$  follows  $\mathbb{P}_{\nu,\eta}$ . In particular,  $x_1, x_2$  and  $x_3$  are respectively distributed according to  $\mu, \nu$  and  $\eta$ , and we have

$$\mathbb{P}(\{(x_1, x_2, x_3) : x_1 \leq x_2 \leq x_3\}) = 1.$$

□

**Definition 3.10.** A pseudometric on an ordered set is *linear* if  $\rho(a, c) = \rho(a, b) + \rho(b, c)$  for every  $a \leq b \leq c$ .

**Lemma 3.11.** Let  $\rho$  be a linear pseudometric on a totally ordered set  $A$ , and let  $\rho'$  be the associated Kantorovich pseudometric on the set of probability measures on  $A$ . Let  $(Y_\mu, Y_\nu)$  be an ordered coupling of two probability measures  $\mu$  and  $\nu$  on  $A$ . Then

$$\mathbb{E}[\rho(Y_\mu, Y_\nu)] = \rho'(\mu, \nu).$$

In other words, the Kantorovich distance is achieved by any ordered coupling.

Moreover, the Kantorovich pseudometric  $\rho'$  is linear for the stochastic order: if  $\mu \stackrel{st}{\leq} \nu \stackrel{st}{\leq} \eta$ , one has

$$\rho'(\mu, \eta) = \rho'(\mu, \nu) + \rho'(\nu, \eta). \quad (3.2)$$

*Proof.* Since  $\rho$  is linear and the set is totally ordered, we can find a non-decreasing map  $\varphi : A \rightarrow \mathbb{R}$  such that for all  $x, y \in A$ ,  $\rho(x, y) = |\varphi(x) - \varphi(y)|$ . Hence we can assume without loss of generality that  $A \subset \mathbb{R}$  and  $\rho(x, y) = |x - y|$ . Since  $(Y_\mu, Y_\nu)$  is an ordered coupling, we can also assume that  $Y_\mu \geq Y_\nu$  a.s. Thus,

$$\mathbb{E}[\rho(Y_\mu, Y_\nu)] = \mathbb{E}[Y_\mu] - \mathbb{E}[Y_\nu] \geq 0.$$

Now, consider any coupling  $(X_\mu, X_\nu)$  of  $\mu$  and  $\nu$ . Then

$$\mathbb{E}[\rho(X_\mu, X_\nu)] = \mathbb{E}[|X_\mu - X_\nu|] \geq \left| \mathbb{E}[X_\mu - X_\nu] \right| = \left| \mathbb{E}[X_\mu] - \mathbb{E}[X_\nu] \right| = \mathbb{E}[\rho(Y_\mu, Y_\nu)],$$

which proves the first assertion of the lemma.

Now, assuming that  $\mu \stackrel{st}{\leq} \nu \stackrel{st}{\leq} \eta$ , we consider an ordered coupling  $(Y_\mu, Y_\nu, Y_\eta)$  where  $Y_\mu \leq Y_\nu \leq Y_\eta$  a.s (see Lemma 3.9). Then,

$$\rho'(\mu, \eta) = \mathbb{E}[\rho(Y_\mu, Y_\eta)] = \mathbb{E}[\rho(Y_\mu, Y_\nu)] + \mathbb{E}[\rho(Y_\nu, Y_\eta)] = \rho'(\mu, \nu) + \rho'(\nu, \eta),$$

and the proof is over. □

In the next proposition,  $(X_n)_{n \leq 0}$  is a monotonic Markov process with a given increasing representation  $(f_n, U_n)$  (see Section 3.1), and we assume that all the state spaces  $A_n$  are totally ordered. Given a distance  $\rho_0$  on  $A_0$ , we iteratively define the pseudometrics  $\rho_n$  on  $A_n$  as in Section 2.2. As explained in Section 2.1, for any  $m \leq 0$ , for any  $x_m \in A_m$ , we denote by  $(Y_n(m, x_m))_{m \leq n \leq 0}$  the Propp-Wilson coupling starting at  $x_m$ .

This proposition is the main point in the demonstration of Theorem 3.6. It will also be used later to derive the intrinsic metrics on the Pascal and Euler graphs.

**Proposition 3.12.** Assume that  $\rho_0$  is a linear distance on  $A_0$ . Then for all  $n \leq 0$ ,  $\rho_n$  is a linear pseudometric on  $A_n$ . Moreover, for all  $(y, z)$  in  $A_n$ ,  $\rho_n(y, z)$  is the Kantorovich distance between  $\mathcal{L}(X_0 | X_n = y)$  and  $\mathcal{L}(X_0 | X_n = z)$  induced by  $\rho_0$  and

$$\forall y, z, \quad \rho_n(y, z) = \mathbb{E}[\rho_0(Y_0(n, y), Y_0(n, z))].$$

*Proof.* The statement of the lemma obviously holds for  $n = 0$ . Assume that it holds for  $n + 1$  ( $n \leq -1$ ). Since the updating functions  $f_n$  are increasing, for all  $(y, z)$  in  $A_n$ , the random pair  $(Y_{n+1}(n, y), Y_{n+1}(n, z))$  is an ordered coupling of  $\mathcal{L}(X_{n+1} | X_n = y)$  and  $\mathcal{L}(X_{n+1} | X_n = z)$ . Therefore by Lemma 3.11 and using the linearity of  $\rho_{n+1}$ ,

$$\rho_n(y, z) := (\rho_{n+1})'(\mathcal{L}(X_{n+1} | X_n = y), \mathcal{L}(X_{n+1} | X_n = z))$$

is a linear distance, and moreover

$$\rho_n(y, z) = \mathbb{E}[\rho_{n+1}(Y_{n+1}(n, y), Y_{n+1}(n, z))].$$

By induction, this is equal to

$$\mathbb{E}[\rho_0(Y_0(n+1, Y_{n+1}(n, y)), Y_0(n+1, Y_{n+1}(n, z)))].$$

Observe now that for any  $x$ , we have  $Y_0(n+1, Y_{n+1}(n, x)) = Y_0(n, x)$ . Hence,

$$\rho_n(y, z) = \mathbb{E}[\rho_0(Y_0(n, y), Y_0(n, z))].$$

Moreover, the random pair  $(Y_0(n, y), Y_0(n, z))$  is an ordered coupling of  $\mathcal{L}(X_0 | X_n = y)$  and  $\mathcal{L}(X_0 | X_n = z)$ . Therefore, by Lemma 3.11, since  $\rho_0$  is linear, we get that  $\rho_n(y, z)$  is the Kantorovich distance between  $\mathcal{L}(X_0 | X_n = y)$  and  $\mathcal{L}(X_0 | X_n = z)$  induced by  $\rho_0$ .  $\square$

### 3.3 Proof of Theorem 3.6

We are now ready to prove the equivalence between the conditions stated in Theorem 3.6.

We have seen at the end of Section 3.1 that there exists an increasing representation, thus  $(c) \implies (b)$  is obvious.  $(b) \implies (a)$  stems from Lemma 1.5.  $(a) \implies (e)$  is obvious and  $(e) \implies (d)$  stems from Lemma 3.8. The main point to show is  $(d) \implies (c)$ . Let  $(f_n, U_n)_{n \leq 0}$  be a parameterization of  $(X_n)_{n \leq 0}$  with increasing updating functions  $f_n$ . We denote by  $\rho$  the usual distance on  $\mathbb{R}$ .

By hypothesis, for each fixed  $n \leq 0$ ,  $\mathcal{L}(X_n | \mathcal{F}_{-\infty}) = \mathcal{L}(X_n)$ . Without loss of generality, we can assume that every  $X_n$  takes its values in a compact subset of  $\mathbb{R}$ . Lemma 3.8 then gives the  $L^1$ -convergence of  $\mathcal{L}(X_n | \mathcal{F}_m)$  to  $\mathcal{L}(X_n)$  as  $m$  goes to  $-\infty$ :

$$s_m := \mathbb{E}[\rho'(\mathcal{L}(X_n | \mathcal{F}_m), \mathcal{L}(X_n))] \xrightarrow{m \rightarrow -\infty} 0.$$

Hence, for each  $m$  there exists  $x_m$  in the state space of  $X_m$  such that

$$\rho'(\mathcal{L}(X_n | X_m = x_m), \mathcal{L}(X_n)) \leq s_m.$$

Consider the Propp-Wilson coupling construction of Section 2.1. Since  $\rho$  is a linear distance, and each  $f_n$  is increasing, we can apply Lemma 3.11 to get

$$\mathbb{E}[\rho(X_n, Y_n(m, x_m)) | \mathcal{F}_m] = \rho'(\mathcal{L}(X_n | \mathcal{F}_m), \mathcal{L}(X_n | X_m = x_m))$$

for every integer  $m < n \leq 0$ . Taking the expectation on both sides yields

$$\begin{aligned} \mathbb{E}[\rho(X_n, Y_n(m, x_m))] &= \mathbb{E}[\rho'(\mathcal{L}(X_n | \mathcal{F}_m), \mathcal{L}(X_n | X_m = x_m))] \\ &\leq \mathbb{E}[\rho'(\mathcal{L}(X_n | \mathcal{F}_m), \mathcal{L}(X_n))] + \mathbb{E}[\rho'(\mathcal{L}(X_n), \mathcal{L}(X_n | X_m = x_m))] \\ &\leq 2s_m \xrightarrow{m \rightarrow -\infty} 0. \end{aligned}$$

Then  $(c)$  follows from Lemma 2.1.

Now to prove 2) we take the sequence  $(x_m)$  for  $n = 0$  and we use again the Propp-Wilson coupling. Assuming that the Markov process is identifiable, the iterated Kantorovich pseudometrics  $\rho_n$  introduced in Section 2.2 with initial distance  $\rho_0 = \rho$  are metrics.

By Proposition 3.12, for every integer  $m \leq n \leq 0$ ,

$$\rho_n(X_n, Y_n(m, x_m)) = \mathbb{E} [\rho_0(X_0, Y_0(m, x_m)) \mid X_n, Y_n(m, x_m)].$$

We have seen in the first part of the proof that

$$\mathbb{E} [\rho_0(X_0, Y_0(m, x_m))] \xrightarrow{m \rightarrow -\infty} 0$$

under the assumption  $\mathcal{L}(X_0 \mid \mathcal{F}_{-\infty}) = \mathcal{L}(X_0)$ . Thus, for every  $n \leq 0$ , the expectation  $\mathbb{E} [\rho_n(X_n, Y_n(m, x_m))]$  goes to 0 as  $m \rightarrow -\infty$ , and Lemma 2.1 gives the result.

## 4 Multidimensional monotonic Markov processes

We now want to prove a multidimensional version of Theorem 3.6. However, as compared to the unidimensional case, the criterion we obtain only guarantee standardness of the filtration, but not the existence of a generating parameterization. In this section,  $(X_n)_{n \leq 0}$  is a Markov process taking its values in  $\mathbb{R}^d$  for some integer  $d \geq 1$  or  $d = \infty$ . For each  $n \leq 0$ , we denote by  $\mu_n$  the law of  $X_n$ , and by  $A_n$  the support of  $\mu_n$ .

### 4.1 Monotonicity for multidimensional Markov processes

We first have to extend the notion of monotonicity given in Definition 3.3 to the case of multidimensional Markov processes.

**Definition 4.1.** We say that  $(X_n)$  is *monotonic* if for each  $n < 0$ , for all  $x, x'$  in  $A_n$ , there exists a coupling  $(Y, Y')$  of  $\mathcal{L}(X_{n+1} \mid X_n = x)$  and  $\mathcal{L}(X_{n+1} \mid X_n = x')$ , whose distribution depends measurably on  $(x, x')$ , and which is *well-ordered with respect to  $(x, x')$* , which means that, for each  $1 \leq k \leq d$ ,

- $x(k) \leq x'(k) \implies \mathbb{P}(Y(k) \leq Y'(k)) = 1$ ,
- $x(k) \geq x'(k) \implies \mathbb{P}(Y(k) \geq Y'(k)) = 1$ .

For example,  $(X_n)_{n \leq 0}$  is a monotonic Markov process when the one-dimensional coordinate processes  $(X_n(k))_{n \leq 0}$  are independent monotonic Markov processes. But the definition does not require nor imply that the coordinate processes  $(X_n(k))_{n \leq 0}$  are Markovian.

Theorem 3.6 will be generalized to  $\mathbb{R}^d$ -valued monotonic processes in Theorem 4.5, except that we will not get the simpler criteria 2) under the identifiability assumption. This will be obtained with the help of Vershik's criterion (Lemma 2.5) in Theorem 4.8 for *strongly monotonic* Markov processes, defined below.

**Definition 4.2.** A Markov process  $(X_n)_{n \leq 0}$  taking its values in  $\mathbb{R}^d$  is said to be *strongly monotonic* if it is monotonic in the sense of the previous definition and if in addition, denoting by  $\mathcal{F}$  the filtration it generates and by  $\mathcal{F}(k)$  the filtration generated by the  $k$ -th coordinate process  $(X_n(k))_{n \leq 0}$ , the two following conditions hold:

- a) each process  $(X_n(k))_{n \leq 0}$  is Markovian,
- b) each filtration  $\mathcal{F}(k)$  is immersed in the filtration  $\mathcal{F}$ ,

Note that conditions a) and b) together mean that each process  $(X_n(k))_{n \leq 0}$  is Markovian with respect to  $\mathcal{F}$ .

The proof of the following lemma is left to the reader.

**Lemma 4.3.** *Let  $(X_n)_{n \leq 0}$  be a strongly monotonic Markov process taking its values in  $\mathbb{R}^d$ . Then each coordinate process  $(X_n(k))_{n \leq 0}$  is a monotonic Markov process.*

The converse of Lemma 4.3 is false, as shown by the example below.

**Example 4.4** (*Random walk on a square*). Let  $(X_n)_{n \leq 0}$  be the stationary random walk on the square  $\{-1, 1\} \times \{-1, 1\}$ , whose distribution is defined by:

- (*Instantaneous laws*) each  $X_n$  has the uniform distribution on  $\{-1, 1\} \times \{-1, 1\}$ ;
- (*Markovian transition*) at each time, the process jumps at random from one vertex of the square to one of its two connected vertices, more precisely, given  $X_n = (x_n(1), x_n(2))$ , the random variable  $X_{n+1}$  takes either the value  $(-x_n(1), x_n(2))$  or  $(x_n(1), -x_n(2))$  with equal probability.

Each of the two coordinate processes  $(X_n(1))_{n \leq 0}$  and  $(X_n(2))_{n \leq 0}$  is a sequence of independent random variables, therefore is a monotonic Markov process. It is not difficult to see in addition that each of them is Markovian with respect to the filtration  $\mathcal{F}$  of  $(X_n)_{n \leq 0}$ , hence the two conditions of Lemma 4.3 hold true. But one can easily check that the process  $(X_n)$  does not satisfy the conditions of Definition 4.1.

Note that the tail  $\sigma$ -field  $\mathcal{F}_{-\infty}$  is not degenerate because of the periodicity of  $(X_n)_{n \leq 0}$ , hence we obviously know that standardness does not hold for  $\mathcal{F}$ .

## 4.2 Standardness for monotonic multidimensional Markov processes

Since we are interested in the filtration generated by  $(X_n)_{n \leq 0}$ , one can assume without loss of generality that the support  $A_n$  of the law of  $X_n$  is included in  $[0, 1]^d$  for every  $n \leq 0$ . Indeed, applying a strictly increasing transformation on each coordinate of the process alters neither the Markov and the monotonicity properties, nor the  $\sigma$ -fields  $\sigma(X_n)$ .

**Theorem 4.5.** *Let  $(X_n)_{n \leq 0}$  be a  $d$ -dimensional monotonic Markov process, and  $\mathcal{F}$  the filtration it generates. The following conditions are equivalent.*

- $\mathcal{F}$  is Kolmogorovian.
- For every  $n \leq 0$ , the conditional law  $\mathcal{L}(X_n | \mathcal{F}_{-\infty})$  is almost surely equal to  $\mathcal{L}(X_n)$ .
- $\mathcal{F}$  is standard.

*Proof.* We only have to prove that (b) implies (c).

We consider a family  $(U_n^j)_{n \leq 0, j \geq 1}$  of independent random variables, uniformly distributed on  $[0, 1]$ . The standardness of the filtration generated by  $(X_n)_{n \leq 0}$  will be proved by constructing a copy  $(Z_n)_{n \leq 0}$  of  $(X_n)_{n \leq 0}$  such that

- For each  $n \leq 0$ ,  $Z_n$  is measurable with respect to the  $\sigma$ -algebra  $\mathcal{U}_n$  generated by  $(U_m^j)_{m \leq n, j \geq 1}$ . (Observe that the filtration  $\mathcal{U} := (\mathcal{U}_n)_{n \leq 0}$  is of product type.)
- The filtration generated by  $(Z_n)_{n \leq 0}$  is immersed in  $\mathcal{U}$ .

For each  $j \geq 1$ , using the random variables  $U_n^j$  we will construct inductively a process  $Z^j := (Z_n^j)_{n_j \leq n \leq 0}$ , where  $(n_j)_{j \geq 1}$  is a decreasing sequence of negative integers to be precised later. Each  $Z_n$  will then be obtained as an almost-sure limit, as  $j \rightarrow \infty$ , of the sequence  $(Z_n^j)$ .

### Construction of a sequence of processes

We consider as in Section 2 that the Markovian transitions are given by kernels  $P_n$ . For every  $n < 0$ , we take an updating function  $f_{n+1} : A_n \times [0, 1] \rightarrow A_{n+1}$  such that  $\mathcal{L}(f_n(x, U)) = P_n(x, \cdot)$  for every  $x \in A_n$  whenever  $U$  is uniformly distributed on  $[0, 1]$ .

To construct the first process  $Z^1$ , we choose an appropriate point  $x_{n_1} \in A_{n_1}$  (which is also to be precised later), and set  $Z_{n_1}^1 := x_{n_1}$ . Then for  $n_1 \leq n < 0$ , we inductively define

$$Z_{n+1}^1 := f_{n+1}(Z_n^1, U_{n+1}^1),$$

so that

$$\mathcal{L}(Z^1) = \mathcal{L}\left((X_n)_{n_1 \leq n \leq 0} \mid X_{n_1} = x_{n_1}\right).$$

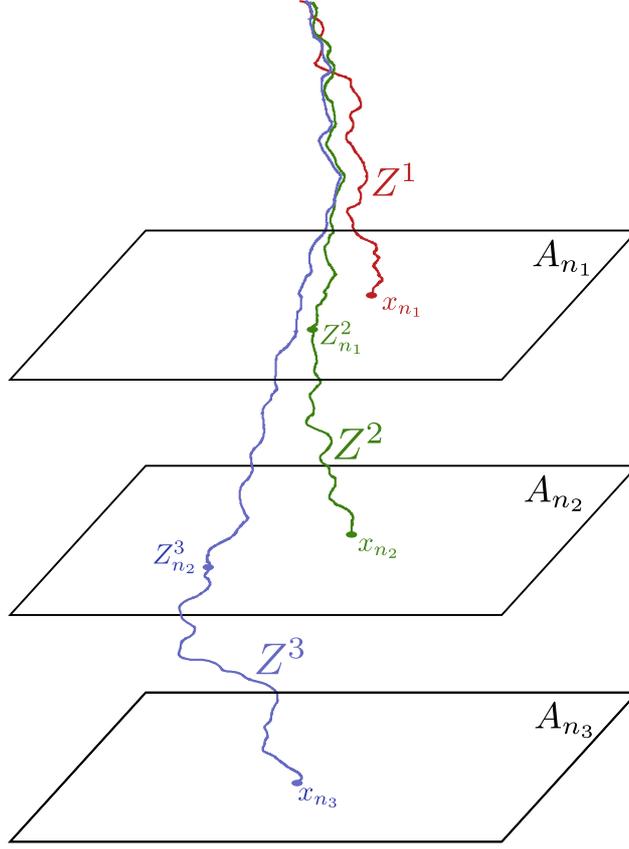


Figure 1: Construction of the sequence of processes  $(Z^j)$ . The processes  $Z^j$  and  $Z^{j+1}$  are coupled in a well-ordered way from time  $n_j$  to 0.

Assume that we have constructed the processes  $Z^i$  for all  $1 \leq i \leq j$ . Then we get the process  $Z^{j+1}$  by choosing an appropriate point  $x_{n_{j+1}} \in A_{n_{j+1}}$ , setting  $Z_{n_{j+1}}^{j+1} := x_{n_{j+1}}$ , and inductively

$$Z_{n+1}^{j+1} := \begin{cases} f_{n+1} \left( Z_n^{j+1}, U_{n+1}^{j+1} \right) & \text{for } n_{j+1} \leq n < n_j, \\ f_{n+1}^{j+1} \left( Z_n^j, Z_n^{j+1}, Z_{n+1}^j, U_{n+1}^{j+1} \right) & \text{for } n_j \leq n < 0. \end{cases}$$

where the function  $f_{n+1}^{j+1}$  is recursively obtained as follows. Let  $(Z, Z', Z_+, Z'_+)$  be a random four-tuple such that  $\mathcal{L}(Z, Z') = \mathcal{L}(Z_n^j, Z_n^{j+1})$  and  $\mathcal{L}((Z_+, Z'_+) | Z, Z') = \Lambda_{Z_1, Z'_1}$ , where  $\Lambda_{x, x'}$  is the well-ordered coupling of Definition 4.1. Recall that the first and second margins of  $\Lambda_{x, x'}$  are  $P_{n+1}(x, \cdot)$  and  $P_{n+1}(x', \cdot)$ . Now, consider a kernel  $Q$  being a regular version of the conditional distribution  $\mathcal{L}(Z'_+ | Z, Z', Z_+)$ , and then take  $f_{n+1}^{j+1}$  such that  $\mathcal{L}(f_{n+1}^{j+1}(z, z', z_+, U)) = Q((z, z', z_+), \cdot)$  for every  $(z, z', z_+) \in A_n \times A_n \times A_{n+1}$  whenever  $U$  is uniformly distributed on  $[0, 1]$ .

In this way, we get by construction

$$\forall n_j \leq n < 0, \quad \mathcal{L}(Z_{n+1}^{j+1} | Z_n^j, Z_n^{j+1}) = P_{n+1}(Z_n^{j+1}, \cdot). \quad (4.1)$$

Moreover, we easily prove by induction that  $Z_n^j$  is measurable with respect to  $\sigma(U_{m,i}; m \leq n, 1 \leq i \leq j) \subset \mathcal{U}_n$  for all possible  $n$  and  $j$ . Now, we want to prove that, for all  $j \geq 1$  and all  $n_j \leq n < 0$ ,

$$\mathcal{L}(Z_{n+1}^j | \mathcal{U}_n) = P_{n+1}(Z_n^j, \cdot). \quad (4.2)$$

This equality stems from the definition of  $f_{n+1}$  for  $j = 1$ . Assuming the equality holds for  $j$ , we show that it holds for  $j + 1$  as follows. When  $n_{j+1} \leq n < n_j$ , this comes again from the definition of  $f_{n+1}$ . If  $n_j \leq n < 0$ , since  $Z_{n+1}^{j+1} = f_{n+1}^{j+1}(Z_n^j, Z_n^{j+1}, Z_{n+1}^j, U_{n+1}^{j+1})$ , where

$U_{n+1}^{j+1}$  is independent of  $(Z_n^j, Z_n^{j+1}, Z_{n+1}^j)$ , we get

$$\mathcal{L}(Z_{n+1}^{j+1} | \mathcal{U}_n \vee Z_{n+1}^j) = \mathcal{L}(Z_{n+1}^{j+1} | Z_n^j, Z_n^{j+1}, Z_{n+1}^j).$$

Using the induction hypothesis, we know that  $\mathcal{L}(Z_{n+1}^j | \mathcal{U}_n) = \mathcal{L}(Z_{n+1}^j | Z_n^j)$ , and we can write

$$\mathcal{L}(Z_{n+1}^{j+1} | \mathcal{U}_n) = \mathcal{L}(Z_{n+1}^{j+1} | Z_n^j, Z_n^{j+1}).$$

Recalling (4.1), we conclude that (4.2) holds for  $j + 1$ .

From (4.2), it follows that

$$\mathcal{L}(Z^j) = \mathcal{L}\left((X_n)_{n_j \leq n \leq 0} | X_{n_j} = x_{n_j}\right)$$

for every  $j \geq 1$ . Moreover, given  $Z_{n_j}^{j+1}$ , the processes  $Z^j$  and  $Z^{j+1}$  are coupled from  $n_j$  in a well-ordered way with respect to  $(x_{n_j}, Z_{n_j}^{j+1})$ . (See Figure 1.)

### Choice of the sequences $(n_j)$ and $(x_{n_j})$

In this part we explain how we can choose the sequences  $(n_j)$  and  $(x_{n_j})$  so that

$$\forall n \leq 0, Z_n^j \text{ converges almost surely as } j \rightarrow \infty. \quad (4.3)$$

Moreover, to ensure that the filtration generated by the limit process  $(Z_n)_{n \leq 0}$  is immersed in  $\mathcal{U}$ , we will also require the following convergence:

$$\forall n \leq -1, \mathcal{L}(Z_{n+1}^j | Z_n^j) \xrightarrow[j \rightarrow \infty]{a.s.} \mathcal{L}(Z_{n+1} | Z_n). \quad (4.4)$$

Recall we assumed that  $A_n \subset [0, 1]^d$ . Let us define the distance  $\rho$  on  $A_n$  by  $\rho(x, x') := \sum_{k=1}^d a_k |x(k) - x'(k)|$ , where, in order to handle the case when  $d = \infty$ , we take a sequence  $(a_k)_{k=1}^d$  of positive numbers satisfying  $\sum a_k = 1$ . For any  $j \geq 1$ , we also define the distance  $\Delta_j$  on  $(\mathbb{R}^d)^j$  by

$$\Delta_j((x_1, \dots, x_j), (y_1, \dots, y_j)) := \max_{1 \leq \ell \leq j} \rho(x_\ell, y_\ell).$$

Let us introduce, for  $j \geq 1$ , and  $\ell \leq -j$ , the measurable subset of  $A_\ell$

$$M_\ell^j := \left\{ x \in A_\ell : \Delta_j' \left( \mathcal{L}\left((X_n)_{-j < n \leq 0} | X_\ell = x\right), \mathcal{L}\left((X_n)_{-j < n \leq 0}\right) \right) > 2^{-j} \right\}.$$

Applying Lemma 3.8 and using hypothesis (b) of Theorem 4.5, for each  $j \geq 1$ ,

$$\mu_\ell(M_\ell^j) \xrightarrow[\ell \rightarrow -\infty]{} 0. \quad (4.5)$$

For each  $n \leq -1$ , we denote by  $\mathcal{M}_1(A_{n+1})$  the set of probability measures on  $A_{n+1}$ , equipped with the Kantorovich distance  $\rho'$ . We also consider  $\varphi_n : A_n \rightarrow \mathcal{M}_1(A_{n+1})$ , defined by

$$\varphi_n(z) := \mathcal{L}(X_{n+1} | X_n = z).$$

Since  $\varphi_n$  is a measurable function, we can apply Lusin Theorem to get the existence, for any  $k \geq 1$ , of a continuous approximation  $\varphi_n^k$  of  $\varphi_n$ , such that

$$\mu_n(\varphi_n \neq \varphi_n^k) < 2^{-k}. \quad (4.6)$$

Let us choose  $n_1$  and  $x_{n_1}$ : By (4.5), we can choose  $|n_1|$  large enough so that  $\mu_{n_1}(M_{n_1}^1) < 2^{-1}$ , and then choose  $x_{n_1} \in A_{n_1} \setminus M_{n_1}^1$ .

Assume now that for some  $j \geq 2$  we have already chosen  $n_{j-1}$  such that  $\mu_{n_{j-1}}(M_{n_{j-1}}^{j-1}) < 2^{-(j-1)}$  and  $x_{n_{j-1}} \in A_{n_{j-1}} \setminus M_{n_{j-1}}^{j-1}$ . By Lemma 3.8 and using hypothesis (b), we get

$$\mathbb{P}\left(X_{n_{j-1}} \in M_{n_{j-1}}^{j-1} | X_\ell\right) \xrightarrow[\ell \rightarrow -\infty]{a.s.} \mu_{n_{j-1}}(M_{n_{j-1}}^{j-1}) < 2^{-(j-1)},$$

and for each  $n, k, -j \leq n \leq 0, 1 \leq k \leq j$ ,

$$\mathbb{P}\left(\varphi_n(X_n) \neq \varphi_n^k(X_n) \mid X_\ell\right) \xrightarrow[\ell \rightarrow -\infty]{a.s.} \mu_n\left(\varphi_n \neq \varphi_n^k\right) < 2^{-k}.$$

Therefore, using also (4.5), if  $|n_j|$  is large enough, we will have

$$\mu_{n_j}\left(M_{n_j}^j\right) < 2^{-j},$$

and there exists  $x_{n_j} \in A_{n_j} \setminus M_{n_j}^j$  such that

$$\mathbb{P}\left(X_{n_{j-1}} \in M_{n_{j-1}}^{j-1} \mid X_{n_j} = x_{n_j}\right) < 2^{-(j-1)}, \quad (4.7)$$

as well as

$$\forall n, k, -j \leq n \leq 0, 1 \leq k \leq j, \mathbb{P}\left(\varphi_n(X_n) \neq \varphi_n^k(X_n) \mid X_{n_j} = x_{n_j}\right) < 2^{-k}. \quad (4.8)$$

### Convergence of the sequence of processes

We want to prove that, for each  $n \leq 0$ , with the above choice of  $(n_j)$  and  $(x_{n_j})$ , the sequence  $(Z_n^j)_{j \geq -n}$  is almost surely a Cauchy sequence.

Since we used well-ordered couplings in the construction of the processes  $Z^j$ , and since the distance  $\delta$  defined by the absolute value on  $\mathbb{R}$  is linear, by application of Lemma 3.11, we have, when  $-j \leq n < 0$

$$\begin{aligned} \mathbb{E}\left[\rho\left(Z_n^j, Z_n^{j+1}\right) \mid Z_{n_j}^{j+1}\right] &= \sum_{k=1}^d a_k \mathbb{E}\left[\left|Z_n^j(k) - Z_n^{j+1}(k)\right| \mid Z_{n_j}^{j+1}\right] \\ &= \sum_{k=1}^d a_k \delta'\left(\mathcal{L}\left(Z_n^j(k) \mid Z_{n_j}^{j+1}\right), \mathcal{L}\left(Z_n^{j+1}(k) \mid Z_{n_j}^{j+1}\right)\right) \\ &\leq \rho'\left(\mathcal{L}\left(Z_n^j \mid Z_{n_j}^{j+1}\right), \mathcal{L}\left(Z_n^{j+1} \mid Z_{n_j}^{j+1}\right)\right) \\ &= \rho'\left(\mathcal{L}(X_n \mid X_{n_j} = x_{n_j}), \mathcal{L}(X_n \mid X_{n_j} = Z_{n_j}^{j+1})\right), \end{aligned} \quad (4.9)$$

the inequality coming from the fact that the minimum of a sum is larger than the sum of the minima. Note that, since the converse inequality is obvious by definition of the Kantorovich distance  $\rho'$ , the above inequality is in fact an equality. Then, by the triangular inequality, we can bound  $\mathbb{E}\left[\rho\left(Z_n^j, Z_n^{j+1}\right) \mid Z_{n_j}^{j+1}\right]$  by the sum

$$\rho'\left(\mathcal{L}(X_n \mid X_{n_j} = x_{n_j}), \mathcal{L}(X_n)\right) + \rho'\left(\mathcal{L}(X_n), \mathcal{L}(X_n \mid X_{n_j} = Z_{n_j}^{j+1})\right). \quad (4.10)$$

Recall we chose  $x_{n_j} \in A_{n_j} \setminus M_{n_j}^j$ , which ensures by definition of  $M_{n_j}^j$  that the first term of (4.10) is bounded by  $2^{-j}$ . Moreover, the second term of (4.10) can be bounded by

$$\mathbb{1}_{Z_{n_j}^{j+1} \in M_{n_j}^j} + 2^{-j} \mathbb{1}_{Z_{n_j}^{j+1} \notin M_{n_j}^j}.$$

By (4.7), for each  $-j < n \leq 0$ ,

$$\mathbb{P}\left(Z_{n_j}^{j+1} \in M_{n_j}^j\right) = \mathbb{P}\left(X_{n_j} \in M_{n_j}^j \mid X_{n_{j+1}} = x_{n_{j+1}}\right) < 2^{-j}.$$

Thus, by integrating with respect to  $Z_{n_j}^{j+1}$ , we obtain that  $\mathbb{E}\left[\rho\left(Z_n^j, Z_n^{j+1}\right)\right]$  is bounded above by

$$2^{-j} + 2^{-j} + \mathbb{P}\left(Z_{n_j}^{j+1} \in M_{n_j}^j\right) \leq 3 \times 2^{-j}.$$

Therefore, for each fixed  $n \leq 0$ ,  $(Z_n^j)_{j > -n}$  is almost surely a Cauchy sequence and converges almost surely to some limit  $Z_n$ , which is measurable with respect to the  $\sigma$ -algebra  $\mathcal{U}_n$  generated by  $(U_m^j)_{m \leq n, j \geq 1}$ .

Observe that for any fixed  $m \leq 0$ , since  $x_{n_j}$  has been chosen in  $A_{n_j} \setminus M_{n_j}^j$ ,

$$\mathcal{L}\left(\left(Z_n^j\right)_{m \leq n \leq 0}\right) = \mathcal{L}\left(\left(X_n\right)_{m \leq n \leq 0} \mid X_{n_j} = x_{n_j}\right) \xrightarrow{j \rightarrow \infty} \mathcal{L}\left(\left(X_n\right)_{m \leq n \leq 0}\right).$$

Hence, we conclude that  $(Z_n)_{n \leq 0}$  is a copy of  $(X_n)_{n \leq 0}$ .

### Proof of the immersion of $(Z_n)_{n \leq 0}$ in $\mathcal{U}$

We need to prove that for all  $n < 0$ ,  $\mathcal{L}(Z_{n+1} | \mathcal{U}_n) = \mathcal{L}(Z_{n+1} | Z_n)$ . We have already seen that

$$\mathcal{L}(Z_{n+1}^j | \mathcal{U}_n) = \mathcal{L}(X_{n+1} | X_n = Z_n^j) = \mathcal{L}(Z_{n+1}^j | Z_n^j).$$

We now want to take the limit as  $j \rightarrow \infty$ . For any continuous function  $g$  on  $A_{n+1}$ , we have

$$\mathbb{E}[g(Z_{n+1}^j) | \mathcal{U}_n] \xrightarrow[j \rightarrow \infty]{a.s.} \mathbb{E}[g(Z_{n+1}) | \mathcal{U}_n]$$

by the conditional dominated convergence theorem. Therefore, by Lemma 3.7,

$$\mathcal{L}(Z_{n+1}^j | \mathcal{U}_n) = \mathcal{L}(Z_{n+1}^j | Z_n^j) \xrightarrow[j \rightarrow \infty]{a.s.} \mathcal{L}(Z_{n+1} | \mathcal{U}_n).$$

By the dominated convergence theorem, we then get

$$\begin{aligned} \mathbb{E}\left[\rho'\left(\mathcal{L}(Z_{n+1}^j | Z_n^j), \mathcal{L}(Z_{n+1} | Z_n)\right)\right] \\ \xrightarrow[j \rightarrow \infty]{} \mathbb{E}\left[\rho'\left(\mathcal{L}(Z_{n+1} | \mathcal{U}_n), \mathcal{L}(Z_{n+1} | Z_n)\right)\right]. \end{aligned} \quad (4.11)$$

On the other hand, the LHS of the preceding formula can be rewritten as  $\mathbb{E}[\rho'(\varphi_n(Z_n^j), \varphi_n(Z_n))]$ , and bounded by the sum of the three following terms:

$$\begin{aligned} T_1 &:= \mathbb{E}\left[\rho'\left(\varphi_n(Z_n^j), \varphi_n^k(Z_n^j)\right)\right], \\ T_2 &:= \mathbb{E}\left[\rho'\left(\varphi_n^k(Z_n^j), \varphi_n^k(Z_n)\right)\right], \\ T_3 &:= \mathbb{E}\left[\rho'\left(\varphi_n^k(Z_n), \varphi_n(Z_n)\right)\right]. \end{aligned}$$

Using (4.6),  $T_3 \leq 2^{-k}$  which can be made arbitrarily small by fixing  $k$  large enough. Once  $k$  has been fixed,  $T_2 \xrightarrow[j \rightarrow \infty]{} 0$  by continuity of  $\varphi_n^k$  and dominated convergence. Then, remembering (4.8), we get  $T_1 < 2^{-k}$  as soon as  $j \geq |n|$  and  $j \geq k$ . This proves that

$$\mathbb{E}\left[\rho'\left(\varphi_n(Z_n^j), \varphi_n(Z_n)\right)\right] \xrightarrow[j \rightarrow \infty]{} 0.$$

Comparing with (4.11), we get the desired equality

$$\mathcal{L}(Z_{n+1} | \mathcal{U}_n) = \mathcal{L}(Z_{n+1} | Z_n).$$

□

### 4.3 Computation of iterated Kantorovich metrics

Here we assume that  $(X_n)_{n \leq 0}$  is a *strongly* monotonic Markov process (Definition 4.2). As before, we assume without loss of generality that it takes its values in  $[0, 1]^d$  equipped with the distance  $\rho$  on  $[0, 1]^d$  defined by  $\rho(x, x') := \sum_{k=1}^d a_k |x(k) - x'(k)|$ , where  $(a_k)_{k=1}^d$  is a sequence of positive numbers satisfying  $\sum a_k = 1$ , whose role is to handle the case when  $d = \infty$ .

The purpose of this section is to establish a connection between the iterated Kantorovich metrics  $\rho_n$  initiated by  $\rho$  and those associated to the Markov processes  $(X_n(k))_{n \leq 0}$ , initiated by the distance  $\delta$  defined by the absolute value on  $\mathbb{R}$ . Then, with the help of Vershik's criterion (Lemma 2.5), we will establish the analogue of criterion 2) in Theorem 3.6.

**Lemma 4.6.** *For each  $\ell \leq 0$ , and each  $n \in \{\ell, \dots, 0\}$ ,*

$$\begin{aligned} \rho'(\mathcal{L}(X_n | X_\ell = x_\ell), \mathcal{L}(X_n | X_\ell = x'_\ell)) \\ = \sum_{k=1}^d a_k \delta'(\mathcal{L}(X_n(k) | X_\ell(k) = x_\ell(k)), \mathcal{L}(X_n(k) | X_\ell(k) = x'_\ell(k))). \end{aligned}$$

*Proof.* Let  $x_\ell$  and  $x'_\ell$  be two points in  $A_\ell$ . As in the proof of Theorem 4.5, we can construct two processes  $(Z_n)_{n \geq \ell}$  and  $(Z'_n)_{n \geq \ell}$  such that

- $\mathcal{L}\left((Z_n)_{n \geq \ell}\right) = \mathcal{L}\left((X_n)_{n \geq \ell} \mid X_\ell = x_\ell\right)$ ,
- $\mathcal{L}\left((Z'_n)_{n \geq \ell}\right) = \mathcal{L}\left((X_n)_{n \geq \ell} \mid X_\ell = x'_\ell\right)$ ,
- for each  $n \geq \ell$ , the coupling  $(Z_n, Z'_n)$  is well-ordered with respect to  $(x_\ell, x'_\ell)$ .

By similar arguments as those used in (4.9), relying on Lemma 3.11, we get

$$\begin{aligned} \rho'(\mathcal{L}(X_n \mid X_\ell = x_\ell), \mathcal{L}(X_n \mid X_\ell = x'_\ell)) &= \rho'(\mathcal{L}(Z_n), \mathcal{L}(Z'_n)) \\ &= \sum_{k=1}^d a_k \delta'(\mathcal{L}(Z_n(k)), \mathcal{L}(Z'_n(k))). \end{aligned}$$

But  $\mathcal{L}(Z_n(k)) = \mathcal{L}(X_n(k) \mid X_\ell = x_\ell)$ , and since the process  $(X_n(k))_{n \leq 0}$  is Markovian with respect to the filtration  $\mathcal{F}$ , the latter is also equal to  $\mathcal{L}(X_n(k) \mid X_\ell(k) = x_\ell(k))$ .  $\square$

**Proposition 4.7.** *Let  $(\rho_n)_{n \leq 0}$  be the sequence of iterated Kantorovich pseudometrics associated to the Markov process  $(X_n)_{n \leq 0}$ , initiated by  $\rho$  on  $A_0$ . Then for any  $x_n, x'_n$  in  $A_n$ ,  $\rho_n(x_n, x'_n)$  is the Kantorovich distance between  $\mathcal{L}(X_0 \mid X_n = x_n)$  and  $\mathcal{L}(X_0 \mid X_n = x'_n)$  for every  $n \leq -1$ , and it is given by*

$$\rho_n(x_n, x'_n) = \sum_{k=1}^d a_k \delta_n(x_n(k), x'_n(k))$$

where  $\delta_n$  is the iterated Kantorovich pseudometric associated to the Markov process  $(X_n(k))_{n \leq 0}$ , initiated by  $\delta$ .

*Proof.* By Lemma 3.11, the Kantorovich pseudometrics  $\delta'$  in Lemma 4.6 are linear. Therefore, we can iteratively use Lemma 4.6 to get the  $n$ -th iterated Kantorovich pseudometrics: For any  $x_n, x'_n$  in  $A_n$ ,

$$\rho_n(x_n, x'_n) = \sum_{k=1}^d a_k \delta_n(x_n(k), x'_n(k)).$$

By Lemma 4.3, each process  $(X_n(k))_{n \leq 0}$  is monotonic. Thus we can apply Proposition 3.12 (the unidimensional case), which gives that  $\delta_n(x_n(k), x'_n(k))$  is the Kantorovich pseudometric between  $\mathcal{L}(X_0(k) \mid X_n(k) = x_n(k))$  and  $\mathcal{L}(X_0(k) \mid X_n(k) = x'_n(k))$ . Then  $\rho_n(x_n, x'_n)$  is the Kantorovich distance between  $\mathcal{L}(X_0 \mid X_n = x_n)$  and  $\mathcal{L}(X_0 \mid X_n = x'_n)$  by Lemma 4.6.  $\square$

**Theorem 4.8.** *Let  $(X_n)_{n \leq 0}$  be an  $\mathbb{R}^d$ -valued strongly monotonic Markov process. If it is identifiable, then the equivalent conditions of Theorem 4.5 are also equivalent to  $\mathcal{L}(X_0 \mid \mathcal{F}_{-\infty}) = \mathcal{L}(X_0)$ .*

*Proof.* This is a consequence of Proposition 4.7, Lemma 2.5, and Lemma 3.8.  $\square$

## 5 Standardness of adic filtrations

Standardness of adic filtrations associated to Bratteli graphs has become an important topic since the recent discoveries of Vershik [28, 29]. As we will explain in Section 5.1, these are the filtrations induced by ergodic central measures on the path space of a Bratteli graph.

We will apply Theorem 3.6 to derive standardness of some well-known examples of adic filtrations, namely those corresponding to the Pascal and the Euler graphs (Sections 6 and 7).

Actually, as we will see, it is straightforward from our Theorem 3.6 that every ergodic central probability measure on the one-dimensional Pascal graph induces a standard filtration (by  $(e) \implies (a)$ ). But Theorem 3.6 is also practical to check the ergodicity of the random walk (using  $(d)$  or  $2)$ ). For the Euler graph we cannot directly apply Theorem 3.6 because of multiple edges. Lemma 5.3 will allow us to deal with this situation.

In Section 8 we will apply Theorem 4.5 to get standardness of the adic filtrations corresponding to the multidimensional Pascal graph.

### 5.1 Adic filtrations and other filtrations on Bratteli graphs

Some examples of Bratteli graphs are shown in Figure 2. Usually Bratteli graphs are graded by the nonnegative integers  $\mathbb{N}$  but for our purpose it is more convenient to consider the nonpositive integers  $-\mathbb{N}$  as the index set of the levels of the graphs. Thus, the set of vertices  $\mathbf{V}$  and the set of edges  $\mathbf{E}$  of a Bratteli graph  $B = (\mathbf{V}, \mathbf{E})$  have the form  $\mathbf{V} = \cup_{n \leq 0} \mathbf{V}_n$  and  $\mathbf{E} = \cup_{n \leq 0} \mathbf{E}_n$  where  $\mathbf{V}_n$  denotes the set of vertices at level  $n$  and  $\mathbf{E}_n$  denotes the set of edges connecting levels  $n - 1$  and  $n$ . The 0-th level set of vertices  $\mathbf{V}_0 = \{v_0\}$  actually consists of a single vertex  $v_0$ . Each vertex of level  $n$  is assumed to be connected to at least one vertex at level  $n - 1$  and, if  $n < -1$ , to at least one vertex at level  $n + 1$ .

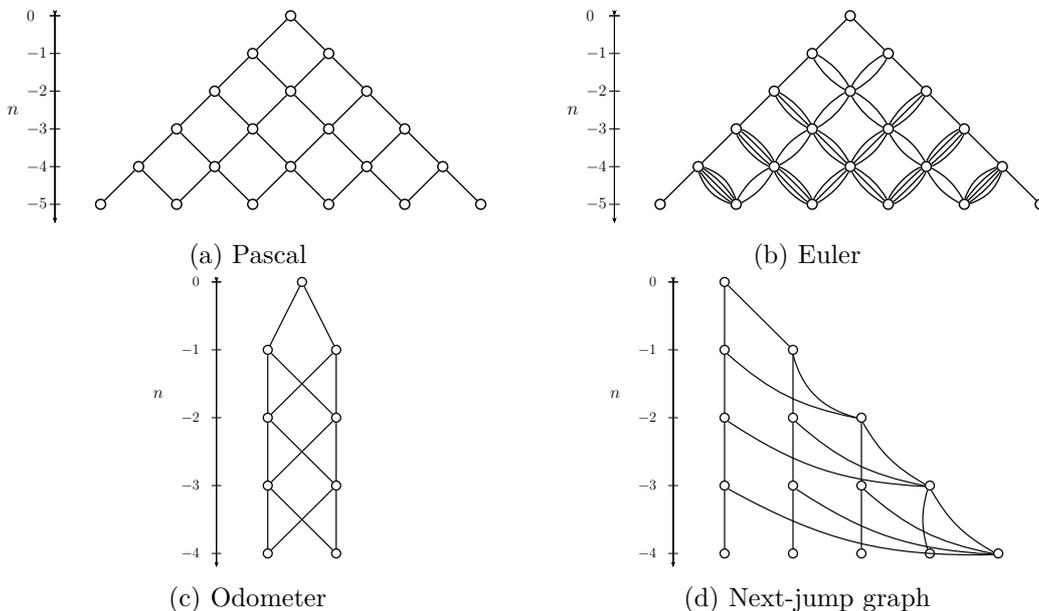


Figure 2: Four Bratteli graphs

There can also exist multiple edges connecting two vertices (see Euler graph). For every vertex  $v \in \mathbf{V}_n$ ,  $n < 0$ , we put *labels* on the set of edges connecting  $v$  to level  $n + 1$  (see Figure 3).

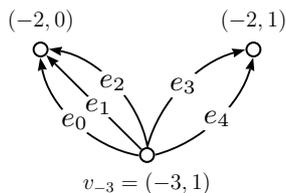


Figure 3: Labeling edges in the Euler graph

We denote by  $\Gamma_B$  the set of infinite paths, where, as usual, an infinite path is a sequence  $\gamma = (\gamma_n)_{n \leq 0} \in \prod_{n \leq 0} \mathbf{E}_n$  of connected edges starting at  $v_0$ , and passing through exactly one vertex at each level  $n \leq -1$ . The path space  $\Gamma_B$  has a natural Borel structure and any probability  $\mu$  on  $\Gamma_B$  can be interpreted as the law of a random path  $(G_n)_{n \leq 0}$ . The filtration  $\mathcal{G}$  generated by  $(G_n)_{n \leq 0}$  is also the filtration generated by the stochastic process

$(V_n, \varepsilon_n)_{n \leq 0}$  where  $V_n$  is the vertex at level  $n$  of the random path  $(G_n)_{n \leq 0}$  and  $\varepsilon_n$  is the label of the edge connecting the vertices  $V_{n-1}$  and  $V_n$ . When the graph has no multiple edges then  $\mathcal{G}$  is also the filtration generated by the random walk on the vertices  $(V_n)_{n \leq 0}$ .

By Rokhlin's correspondence (see [9]), and up to measure algebra isomorphism, the filtration  $\mathcal{G}$  corresponds to the increasing sequence of measurable partitions  $(\xi_n)_{n \leq 0}$  on  $(\Gamma_B, \mu)$ , where  $\xi_n$  is the measurable partition of  $\Gamma_B$  into the equivalence classes of the equivalence relation  $\mathcal{R}_n$  defined by  $\gamma \mathcal{R}_n \gamma'$  if  $\gamma_m = \gamma'_m$  for all  $m \leq n$ . The probabilistic definition of centrality of the probability measure  $\mu$ , given below, amounts to say that  $\mu$  is invariant for the tail equivalence relation  $\mathcal{R}_{-\infty}$  defined by  $\gamma \mathcal{R}_{-\infty} \gamma'$  if  $\gamma_m = \gamma'_m$  for  $|m|$  large enough.

**Definition 5.1.** The probability measure  $\mu$  on  $\Gamma_B$  is *central* if for each  $n < 0$ , the conditional distribution of  $(G_{n+1}, \dots, G_0)$  given  $\mathcal{G}_n$  is uniform on the set of paths connecting the vertex  $V_n$  to the root of the graph.

The probabilistic property of  $\mathcal{G}$  corresponding to ergodicity of this tail equivalence relation with respect to  $\mu$  is the degeneracy of the tail  $\sigma$ -field:

**Definition 5.2.** The probability measure  $\mu$  on  $\Gamma_B$  is *ergodic* if  $\mathcal{G}$  is Kolmogorovian.

When  $\mu$  is central then the process  $(V_n, \varepsilon_n)_{n \leq 0}$  as well as the random walk on the vertices  $(V_n)_{n \leq 0}$  are Markovian. More precisely,  $(V_n)_{n \leq 0}$  is Markovian with respect to the filtration  $\mathcal{G}$  generated by  $(V_n, \varepsilon_n)_{n \leq 0}$ ; in other words, the filtration  $\mathcal{F}$  generated by  $(V_n)_{n \leq 0}$  is immersed in  $\mathcal{G}$ . Furthermore the conditional distribution of  $V_{n+1}$  given  $V_n = v_n$  is given by

$$\mathbb{P}(V_{n+1} = v_{n+1} | V_n = v_n) = m(v_n, v_{n+1}) \frac{\dim(v_{n+1})}{\dim(v_n)} \quad (5.1)$$

where  $m(v_n, v_{n+1})$  is the number of edges connecting  $v_n$  and  $v_{n+1}$ , and  $\dim(v)$  denotes the number of paths from vertex  $v$  to the final vertex  $v_0$ .

Centrality and ergodicity of  $\mu$  also correspond to invariance and ergodicity of the so-called *adic transformation*  $T$  on  $\Gamma_B$ , and in this case the tail equivalence relation  $\mathcal{R}_{-\infty}$  defines the partition of  $\Gamma_B$  into the orbits of the adic transformation. Standardness of  $\mathcal{G}$  is stronger than ergodicity of  $\mu$ , but note that standardness of  $\mathcal{G}$  under a central ergodic measure  $\mu$  is not a priori a property about the corresponding adic transformation, since the adic transformation on a Bratteli graph is possibly isomorphic to the adic transformation on another Bratteli graph, and these two different Bratteli graphs can generate non-isomorphic filtrations. For example the dyadic odometer is isomorphic to an adic transformation on the graph shown on Figure 2c as well as an adic transformation on the graph shown on Figure 2d. The usual adic representation of the dyadic odometer is given by the graph shown in Figure 2c. One easily sees that there is a unique central probability measure, and that the corresponding Markov process  $(V_n)_{n \leq 0}$  is actually a sequence of i.i.d. random variables having the uniform distribution on  $\{0, 1\}$ . Therefore  $\mathcal{G}$  is obviously a standard filtration. The Bratteli graph of Figure 2d shows another possible adic representation of the dyadic odometer. Standardness of the corresponding filtration  $\mathcal{G}$  has been studied in [16] and [17] in the case when  $\mu$  is any independent product of Bernoulli measures on the path space, and this includes all the central ergodic measures. In Sections 6 and 7 we will use Theorem 3.6 to study the case of the Pascal graph (Figure 2a) and the case of the Euler graph (Figure 2b).

The lemma below is useful to establish standardness in the case of a graph with multiple edges, such as the Euler graph. Note that the conditional independence assumption  $\mathcal{L}(\varepsilon_n | V_{n-1}) = \mathcal{L}(\varepsilon_n | \mathcal{G}_{n-1})$  of this lemma implies that  $(V_n)_{n \leq 0}$  is Markovian, and this assumption is always fulfilled for a central measure.

**Lemma 5.3.** *Let  $\mathcal{G}$  be the filtration associated to a probability measure on the path space of a Bratteli graph, and denote by  $(V_n, \varepsilon_n)_{n \leq 0}$  the stochastic process generating  $\mathcal{G}$ , where  $V_n$  is the vertex at level  $n$  and  $\varepsilon_n$  is the label of the edge connecting  $V_{n-1}$  to  $V_n$ . Assume that  $\mathcal{L}(\varepsilon_n | V_{n-1}) = \mathcal{L}(\varepsilon_n | \mathcal{G}_{n-1})$ , that is to say  $\varepsilon_n$  is conditionally independent of  $\mathcal{G}_{n-1}$  given  $V_{n-1}$ . Denote by  $\mathcal{F}$  the filtration of the random walk  $(V_n)_{n \leq 0}$  on the vertices.*

*Then*

- 1) there exists a parameterization  $(U_n)_{n \leq 0}$  of  $\mathcal{F}$  which is also a parameterization of  $\mathcal{G}$ , and such that the parametric extension of  $\mathcal{F}$  with  $(U_n)_{n \leq 0}$  (Definition 1.4) is also the parametric extension of  $\mathcal{G}$  with  $(U_n)_{n \leq 0}$ ;
- 2) assuming  $\mathbf{V}_n \subset \mathbb{R}$  and  $(V_n)_{n \leq 0}$  monotonic, there exists a monotonic parametric representation  $(f_n, U_n)_{n \leq 0}$  of  $(V_n)_{n \leq 0}$  with a parameterization  $(U_n)_{n \leq 0}$  satisfying the above properties.

*Proof.* Assume without loss of generality that the labels of the edges are real numbers. Denote by  $\phi_n$  a measurable function such that  $V_n = \phi_n(V_{n-1}, \varepsilon_n)$ , and denote by  $h_n(v_{n-1}, \cdot)$  the right-continuous inverse of the cumulative distribution function of the conditional law  $\mathcal{L}(\varepsilon_n | V_{n-1} = v_n)$ . Then the function  $f_n$  defined by

$$f_n(v_{n-1}, \cdot) = \phi_n(v_{n-1}, h_n(v_{n-1}, \cdot))$$

is an updating function of the Markov kernel  $\mathbb{P}(V_n \in \cdot | V_{n-1} = v_{n-1})$ .

Consider a copy  $(V'_n)_{n \leq 0}$  of the process  $(V_n)_{n \leq 0}$  given by a parametric representation  $(f_n, U'_n)_{n \leq 0}$  with these updating functions  $f_n$ , and set  $\varepsilon'_n = h_n(V'_{n-1}, U'_n)$ . Then it is not difficult to see that the process  $(V'_n, \varepsilon'_n)_{n \leq 0}$  is a copy of  $(V_n, \varepsilon_n)_{n \leq 0}$ . Moreover, denoting by  $\mathcal{G}'$  its filtration,  $U'_n$  is independent of  $\mathcal{G}'_{n-1}$ , and

$$\mathcal{G}'_n = \mathcal{G}'_{n-1} \vee \sigma(\varepsilon'_n) \subset \mathcal{G}'_{n-1} \vee \sigma(U'_n),$$

thereby showing that  $(U'_n)_{n \leq 0}$  is a parameterization of  $\mathcal{G}'$ . This proves 1).

Assuming now  $\mathbf{V}_n \subset \mathbb{R}$ , it is always possible to take right-continuous increasing functions  $\phi_n(v_{n-1}, \cdot)$ . With such a choice, the function  $f_n$  constructed above is the quantile updating function (2.1), and then the representation is monotonic whenever  $(V_n)_{n \leq 0}$  is monotonic.  $\square$

We cannot deduce from result 1) of Lemma 5.3 that  $\mathcal{G}$  admits a generating parameterization whenever  $\mathcal{F}$  admits a generating parameterization. But thanks to this result and to Proposition 6.1 in [14], which says that standardness is hereditary under parametric extension, we know that  $\mathcal{F}$  is standard if and only if  $\mathcal{G}$  is standard. This result is not used in the present paper but it is useful for the study of other Bratteli graphs.

## 5.2 Vershik's intrinsic metrics

Given a probability measure  $\mu$  on  $\Gamma_B$ , for which the process  $(V_n)_{n \leq 0}$  is Markovian, we can consider the iterated Kantorovich pseudometrics  $\rho_n$  defined as in Section 2.2. But since  $\mathbf{V}_0$  is always reduced to a singleton, we start from a metric  $\rho_{-1}$  defined on the set  $\mathbf{V}_{-1}$  instead of a metric  $\rho_0$  on  $\mathbf{V}_0$ . Each pseudometric  $\rho_n$ ,  $n \leq -1$  is then defined on the set  $\mathbf{V}_n$  of vertices of level  $n$ . These pseudometrics only depend on the Markov kernels  $P_n$ , in particular all central probability measures will give rise to the same sequence of pseudometrics. The pseudometrics  $\rho_n$  obtained in the case of a central measure have been introduced by Vershik in [29], who called them *intrinsic pseudometrics*. In the next sections we will provide the intrinsic metrics  $\rho_n$  for the Pascal graph and the Euler graph with the help of Proposition 3.12, and for the higher dimensional Pascal graph with the help of Proposition 4.7.

Applying the theorems of [29] about the identification of the ergodic central measures is beyond the scope of this paper. This is based on the intrinsic pseudometric defined on the whole set of vertices  $\cup_{n \leq 0} \mathbf{V}_n$  and extending all the  $\rho_n$ , which we will not explicit here. Our derivation of the  $\rho_n$  provides a helpful starting point for further work in this direction.

Recall that the  $\rho_n$  are metrics under the identifiability of the associated Markov process  $(V_n)_{n \leq 0}$  (Definition 2.4), and identifiability is easy to check in the case of central measures. It is equivalent to the following property: *For each  $n < -1$ , for any two different vertices  $v, v' \in \mathbf{V}_n$ , there exists at least one vertex  $w \in \mathbf{V}_{n+1}$  such that the number of edges connecting  $v$  and  $w$  is different from the number of edges connecting  $v'$  and  $w$ .* For a graph without multiple edge, this simply means that  $v$  and  $v'$  are not connected to the same set of vertices at level  $n + 1$ .

## 6 Pascal filtration

Consider the  $(-\mathbb{N})$ -graded Pascal graph shown in Figure 4a. At each level  $n$ , we label the vertices  $0, 1, \dots, |n|$ . Then a vertex can be identified by the pair  $(n, k)$  consisting in its level  $n$  and its label  $k$ , but when the level is understood we simply use the label as the identifier. Each vertex  $v$  at level  $n$  is connected to vertices  $v$  and  $v + 1$  at level  $n - 1$ . There is no multiple edge and a random path in the graph corresponds to a random walk  $(V_n)_{n \leq 0}$  on the vertices of the graph, where  $V_n$  is a vertex at level  $n$  and  $(V_n, V_{n-1})$  are connected.

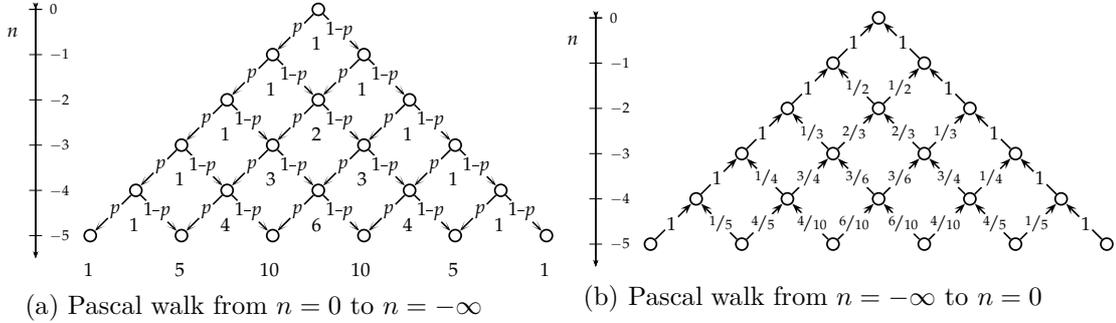


Figure 4: Pascal random walk

The path space of the Pascal graph is naturally identified with  $\{0, 1\}^{-\mathbb{N}}$ . Under any central probability measure, the process  $(V_n)_{n \leq 0}$  obviously is a monotonic and identifiable Markov process (definitions 3.3 and 2.4). Its Markovian transition distributions  $\mathcal{L}(V_n | V_{n-1} = v)$  are easy to derive with the help of formula (5.1). They are shown in Figure 4b for  $n = 0$  to  $n = -4$ . The only thing we will need is the conditional law  $\mathcal{L}(V_{-1} | V_n = v_n)$  and it is not difficult to see that it is the distribution on  $\{0, 1\}$  given by  $\mathbb{P}(V_{-1} = 1 | V_n = v_n) = \frac{v_n}{|n|}$ .

### 6.1 Standardness

It has been shown (see e.g. [20]) that the ergodic central probability measures are those for which the reverse random walk  $(V_0, V_{-1}, \dots)$  is Markovian with a constant Markovian transition  $(p, 1 - p)$  as shown in Figure 4a. In other words the ergodic central probability measures are the infinite product Bernoulli measures  $(p, 1 - p)$ . Then  $V_n$  has the binomial distribution  $\text{Bin}(|n|, p)$ .

Using Theorem 3.6, we can directly show standardness of the filtration  $\mathcal{F}$  generated by  $(V_n)$  under these infinite product Bernoulli measures.

**Proposition 6.1.** *When  $\mu$  is an infinite product Bernoulli measure  $(p, 1 - p)$  then the random walk  $(V_n)_{n \leq 0}$  is a monotonic Markov process generating a standard filtration. In particular, this measure is ergodic.*

*Proof.* Obviously,  $(V_n)_{n \leq -1}$  is a monotonic and identifiable Markov process (see last paragraph in Section 5.2). We check criterion 2) in Theorem 3.6. The conditional distribution  $\mu_{v_n} := \mathcal{L}(V_{-1} | V_n = v_n)$  is the law on  $\{0, 1\}$  given by  $\mu_{v_n}(1) = \frac{v_n}{|n|}$ , thus the conditional law  $\mathcal{L}(V_{-1} | \mathcal{F}_n)$  goes to  $\mathcal{L}(V_{-1})$  by the law of large numbers and then Theorem 3.6 applies in view of Lemma 3.7.  $\square$

In fact, as long as the process  $(V_n)_{n \leq 0}$  is a Markov process for some probability measure on  $\Gamma_B$ , it is easy to see that it is necessarily a monotonic Markov process. We then get the following consequence of Theorem 3.6 (by  $(e) \implies (a)$ ).

**Theorem 6.2.** *For any ergodic probability measure on  $\Gamma_B$  under which  $(V_n)_{n \leq 0}$  is a Markov process, the filtration  $\mathcal{F}$  generated by  $(V_n)_{n \leq 0}$  admits a generating parameterization, hence is standard.*

## 6.2 Intrinsic metrics on the Pascal graph

We did not need to resort to Vershik's standardness criterion (Lemma 2.5) to prove standardness of the Pascal adic filtrations (Proposition 6.1). However, as we mentioned in Section 5.2, it is interesting to have a look at the intrinsic metrics  $\rho_n$  on the state space  $\mathbf{V}_n = \{0, \dots, |n|\}$  of  $V_n$ , starting from the 0-1 distance on  $\mathbf{V}_{-1}$ . The  $\rho_n$  are easily obtained by Proposition 3.12: the distance  $\rho_n(v_n, v'_n)$  is nothing but the Kantorovich distance between  $\mathcal{L}(V_{-1} | V_n = v_n)$  and  $\mathcal{L}(V_{-1} | V_n = v'_n)$ , and then

$$\rho_n(v_n, v'_n) = \frac{|v_n - v'_n|}{|n|},$$

wherefrom it is not difficult to apply Lemma 2.5 to get standardness of  $\mathcal{F}$ . The space  $(\mathbf{V}_n, \rho_n)$  is isometric to the subset  $\{\frac{k}{|n|}, k = 0, \dots, |n|\}$  of the unit interval  $[0, 1]$ . Figure 5 shows an embedding of the Pascal graph in the plane such that  $\rho_n$  is given by the Euclidean distance at each level  $n$ .

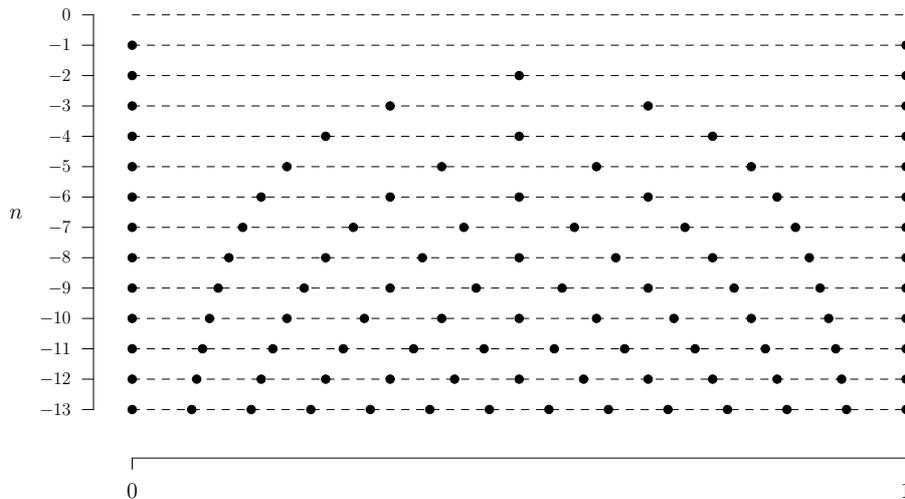


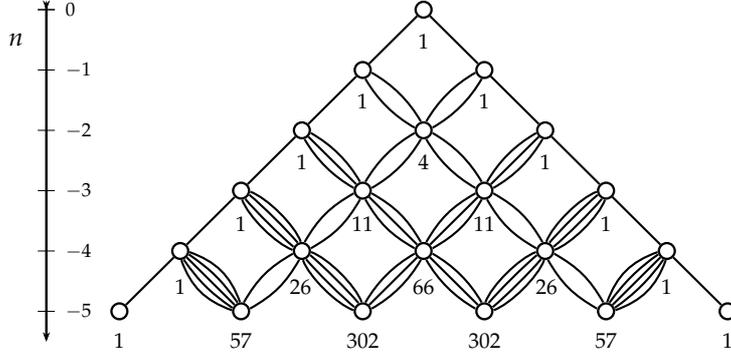
Figure 5: The Pascal graph under the intrinsic metrics

## 7 Euler filtration

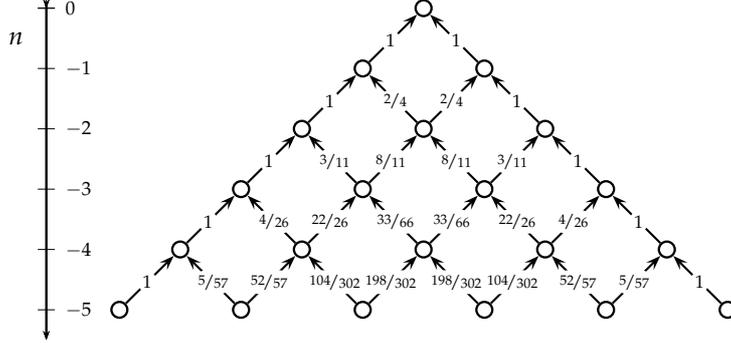
The Euler graph, shown on Figure 6a from level  $n = 0$  to level  $n = -5$ , has the same vertex set as the Pascal graph, but has multiple edges: Vertex  $v$  of level  $n$  is connected to vertex  $v$  of level  $n - 1$  by  $v + 1$  edges, and to vertex  $v + 1$  of level  $n - 1$  by  $|n| + 1 - v$  edges. We refer to [3, 5, 21] for properties of this graph. In particular, the number of paths connecting vertex  $v$  of level  $n \leq -1$  to the root vertex at level 0 is the *Eulerian number*  $A(|n| + 1, v)$ .

It is shown in [11] that there exist countably many ergodic central measures on  $\Gamma_B$  for this graph. However, only one of them, called the *symmetric measure*, has full support, as shown in [3] (the others are concentrated on paths whose distance to one of the sides of the triangle is bounded).

Given a probability measure on  $\Gamma_B$ , as explained in Section 5, we consider a stochastic process  $(G_n)_{n \leq 0}$  distributed on  $\Gamma_B$  according to  $\mu$ , where  $G_n$  is the edge at level  $n$ , and we are interested in the filtration  $\mathcal{G}$  it generates. This filtration is also generated by the process  $(V_n, \varepsilon_n)_{n \leq 0}$ , where  $V_n$  is the vertex at level  $n$  and  $\varepsilon_n$  the label connecting  $V_{n-1}$  to  $V_n$ . Under the symmetric measure, the process  $(G_n)_{n \leq 0}$  is Markovian and the conditional distribution of  $G_{n-1}$  given  $G_n$  is the uniform distribution among the  $|n| + 2$  edges in  $\mathbf{E}_n$  connected to  $G_n$ . We will derive standardness of the filtration  $\mathcal{G}$  under the symmetric measure. The explicit conditional distributions  $\mathcal{L}(V_{-1} | V_n = v_n)$  can be derived from Equation (1.1) in [21], but to show standardness we will only use the following result



(a) Euler graph



(b) Walk on the vertices

Figure 6: Euler random walk

coming from Equation (1.3) in [21]:

$$\lim_{n \rightarrow -\infty} \mathbb{P}(V_{-1} = 1 \mid V_n = v_n) = \frac{1}{2}$$

for every sequence  $(v_n)_{n \leq 0}$  of vertices  $v_n \in \mathbf{V}_n$  such that both  $v_n$  and  $|n| - v_n$  go to infinity as  $n \rightarrow -\infty$ .

## 7.1 Standardness

For the Euler filtration we have to deal with multiple edges:  $\mathcal{G}$  is generated by the Markov process  $(V_n, \varepsilon_n)_{n \leq 0}$  (Section 5) and Theorem 3.6 can only provide a generating parameterization of the smaller filtration  $\mathcal{F}$  generated by the random walk on the vertices  $(V_n)_{n \leq 0}$ . A generating parameterization of  $\mathcal{G}$  will be derived by applying Theorem 3.6 to  $(V_n)_{n \leq 0}$  and then by applying Lemma 5.3.

**Lemma 7.1.** *Under the symmetric central measure  $\mu$ , we have*

$$V_n \xrightarrow[n \rightarrow -\infty]{a.e.} \infty, \quad \text{and} \quad |n| - V_n \xrightarrow[n \rightarrow -\infty]{a.e.} \infty.$$

*Proof.* Consider the Markov process  $(\tilde{V}_n)_{n \leq 0}$  where  $\tilde{V}_n$  takes its values in  $\mathbf{V}_n$ , defined by the conditional distribution

$$\mathbb{P}(\tilde{V}_{n-1} = v \mid \tilde{V}_n = v) = \mathbb{P}(\tilde{V}_{n-1} = v + 1 \mid \tilde{V}_n = v) = \frac{1}{2}.$$

The process  $(\tilde{V}_0, \tilde{V}_{-1}, \dots)$  is nothing but the well-known simple symmetric random walk. By the law of large numbers, the property claimed for  $V_n$  obviously holds for  $\tilde{V}_n$ . Moreover, we can easily construct a coupling of the two Markov processes for which, for all  $n \leq 0$ ,

$$\left| V_n - \frac{|n|}{2} \right| \leq \left| \tilde{V}_n - \frac{|n|}{2} \right| \quad \text{a.s.}$$

Consequently  $(V_n)$  inherits of the same property. □

**Proposition 7.2.** *For the symmetric central measure  $\mu$ , the Euler filtration  $\mathcal{G}$  admits a generating parameterization, hence is standard. In particular,  $\mu$  is ergodic.*

*Proof.* We first check criterion 2) in Theorem 3.6 for  $(V_n)_{n \leq -1}$  which obviously is a monotonic and identifiable Markov process. As we previously mentioned, it follows from Equation (1.3) in [21] that  $\mu_{v_n}(1) \rightarrow \frac{1}{2}$  whenever  $(v_n)$  is a sequence of vertices such that  $v_n \in \mathbf{V}_n$  and both  $v_n$  and  $|n| - v_n$  go to infinity as  $n \rightarrow -\infty$ . We recognize the distribution of  $V_{-1}$  under  $\mu$ , and using Lemma 7.1 we see that criterion 2) in Theorem 3.6 is fulfilled. Now, by (c) in Theorem 3.6 and 2) in Lemma 5.3,  $\mathcal{F}$  and  $\mathcal{G}$  admit a common generating parameterization. It follows by Lemma 1.5 that  $\mathcal{G}$  is standard.  $\square$

Similarly to Theorem 6.2 about the Pascal graph, one has the following theorem for the Euler graph.

**Theorem 7.3.** *Under an ergodic probability measure on  $\Gamma_B$  and under the conditional independence assumption  $\mathcal{L}(\varepsilon_n | V_{n-1}) = \mathcal{L}(\varepsilon_n | \mathcal{G}_{n-1})$ , the filtration  $\mathcal{G}$  admits a generating parameterization, hence is standard.*

*Proof.* Under the conditional independence assumption, the process  $(V_n)_{n \leq 0}$  is Markovian, and the filtration  $\mathcal{F}$  it generates admits a generating parameterization by Theorem 6.2. We conclude similarly to the proof of Proposition 7.2, combining Theorem 3.6 and Lemma 5.3.  $\square$

## 7.2 Intrinsic metrics on the Euler graph

Similarly to the Pascal case, the intrinsic metrics  $\rho_n$  on the state space  $\mathbf{V}_n = \{0, \dots, |n|\}$  of  $V_n$ , starting from the discrete distance on  $\mathbf{V}_{-1}$ , are easily obtained by Proposition 3.12: The distance  $\rho_n(v_n, v'_n)$  is nothing but the Kantorovich distance between  $\mathcal{L}(V_{-1} | V_n = v_n)$  and  $\mathcal{L}(V_{-1} | V_n = v'_n)$ . We can explicit these conditional laws using the formula provided by Equation (1.1) in [21], which gives the number of paths connecting a vertex  $v_n$  at some level  $n \leq -2$  to the right vertex at level  $-1$ . The number of such paths is the generalized Eulerian number

$$A_{0,1}(|n| - v_n, v_n - 1) = \sum_{t=0}^{|n|-v_n} (-1)^{|n|-v_n-t} \binom{t+2}{t} \binom{|n|+2}{|n|-v_n-t} (1+t)^{|n|-1}.$$

Recalling that the total number of paths connecting vertex  $v_n$  of level  $n$  to the root of the graph is the classical Eulerian number  $A(|n| + 1, v_n)$ , we get the conditional law  $\mathcal{L}(V_{-1} | V_n = v_n)$  under the centrality assumption: It is the probability on  $\{0, 1\}$  given by

$$\mathbb{P}(V_{-1} = 1 | V_n = v_n) = \frac{A_{0,1}(|n| - v_n, v_n - 1)}{A(|n| + 1, v_n)}.$$

From this, we can derive the following formula giving the intrinsic metric at level  $n$ :

$$\rho_n(v_n, v'_n) = \left| \frac{A_{0,1}(|n| - v_n, v_n - 1)}{A(|n| + 1, v_n)} - \frac{A_{0,1}(|n| - v'_n, v'_n - 1)}{A(|n| + 1, v'_n)} \right|.$$

We also know by Proposition 3.12 that the space  $(\mathbf{V}_n, \rho_n)$  is isometric a subset of the unit interval  $[0, 1]$ . Figure 7 shows an embedding of the Euler graph in the plane such that  $\rho_n$  is given by the Euclidean distance at each level  $n$ .

## 8 Multidimensional Pascal filtration

Now we introduce the  $d$ -dimensional Pascal graph. The Pascal graph of Section 6 corresponds to the case  $d = 2$ . We will provide three different proofs that the filtration is standard for any dimension  $d \geq 2$  under the known ergodic central measures. The first proof is an application of Theorem 4.5. The second proof is an application of Theorem 4.8, using Proposition 4.7 to derive the intrinsic metrics  $\rho_n$ . These two proofs only

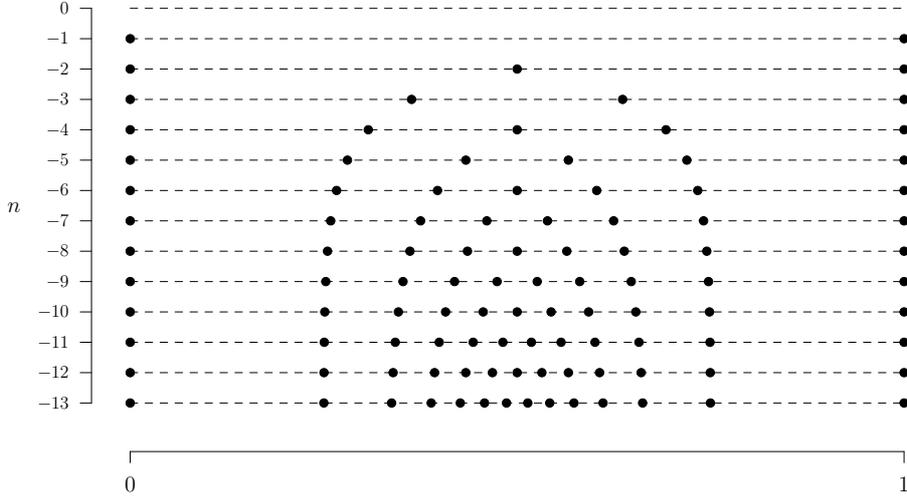


Figure 7: The Euler graph under the intrinsic metrics

provides standardness, not a generating parameterization. In the third proof we construct a generating parameterization with the help of Theorem 3.6.

Let  $d \geq 2$  be an integer or  $d = \infty$ . Vertices of the  $d$ -dimensional Pascal graph are points  $(i_1, \dots, i_d) \in \mathbb{N}^d$  when  $d < \infty$ . When  $d = \infty$ , the vertices are the sequences  $(i_1, i_2, \dots) \in \mathbb{N}^\infty$  with finitely many nonzero terms. The set of vertices at level  $n$  is

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

and two vertices  $(i_1, \dots, i_d) \in \mathbf{V}_n^d$  and  $(j_1, \dots, j_d) \in \mathbf{V}_{n-1}^d$  are connected if and only if  $\sum |i_k - j_k| = 1$ .

Since there is no multiple edge in the graph, for any central probability measure, the corresponding adic filtration  $\mathcal{G}$  is generated by the Markovian random walk on the vertices. Temporarily denoting by  $(V_n)_{n \leq 0}$  this random walk, centrality means that the Markovian transition from  $n$  to  $n + 1$  is given by

$$\mathcal{L}(V_{n+1} \mid V_n = v) = \sum_{i=1}^d \frac{v(i)}{|n|} \delta_{v-e_i}, \quad (8.1)$$

where  $e_i$  is the vector whose  $i$ -th term is 1 and all the other ones are 0.

It is known (see [4], Theorem 5.3) that a central measure is ergodic if and only if there is a probability vector  $(\theta_1, \dots, \theta_d)$  such that every Markov transition from  $n$  to  $n - 1$  is given by

$$\mathbb{P}(V_{n-1} = v_n + e_i \mid V_n = v_n) = \theta_i \quad \text{for all } i.$$

Under this ergodic central measure,  $V_n$  has the multinomial distribution with parameter  $(\theta_1, \dots, \theta_d)$  (see Figure 8). For this reason, let us term the ergodic central measures as the *multinomial central measures*. It is not difficult to check that the multinomial central measures are ergodic, but in our second and third proofs of standardness we will not use ergodicity.

From now on, we denote by  $(V_n^{d,\theta})_{n \leq 0}$  the Markovian random walk corresponding to  $(\theta_1, \dots, \theta_d)$ . We write  $V_n^{d,\theta} = (V_n^{d,\theta}(1), \dots, V_n^{d,\theta}(d))$ . Each process  $(V_n^{d,\theta}(i))_{n \leq 0}$  is the random walk on the vertices of the Pascal graph as in Section 6, and is Markovian with respect to  $\mathcal{G}$ , that is, the filtration  $\mathcal{G}(i)$  generated by the process  $(V_n^{d,\theta}(i))_{n \leq 0}$  is immersed in  $\mathcal{G}$  (thus the multidimensional process satisfies conditions a) and b) of Definition 4.2).

It is worth mentioning that standardness of  $\mathcal{G}$  cannot be deduced from the equality  $\mathcal{G} = \mathcal{G}(1) \vee \dots \vee \mathcal{G}(d)$  and from the fact that the filtrations  $\mathcal{G}(i)$  are standard and jointly immersed: This is a consequence of theorem 3.9 in [13], but Example 4.4 also provides a counter-example, and more precisely it shows that even the degeneracy of  $\mathcal{G}_{-\infty}$  cannot be deduced from the degeneracy of each  $\mathcal{G}_{-\infty}(i)$ .

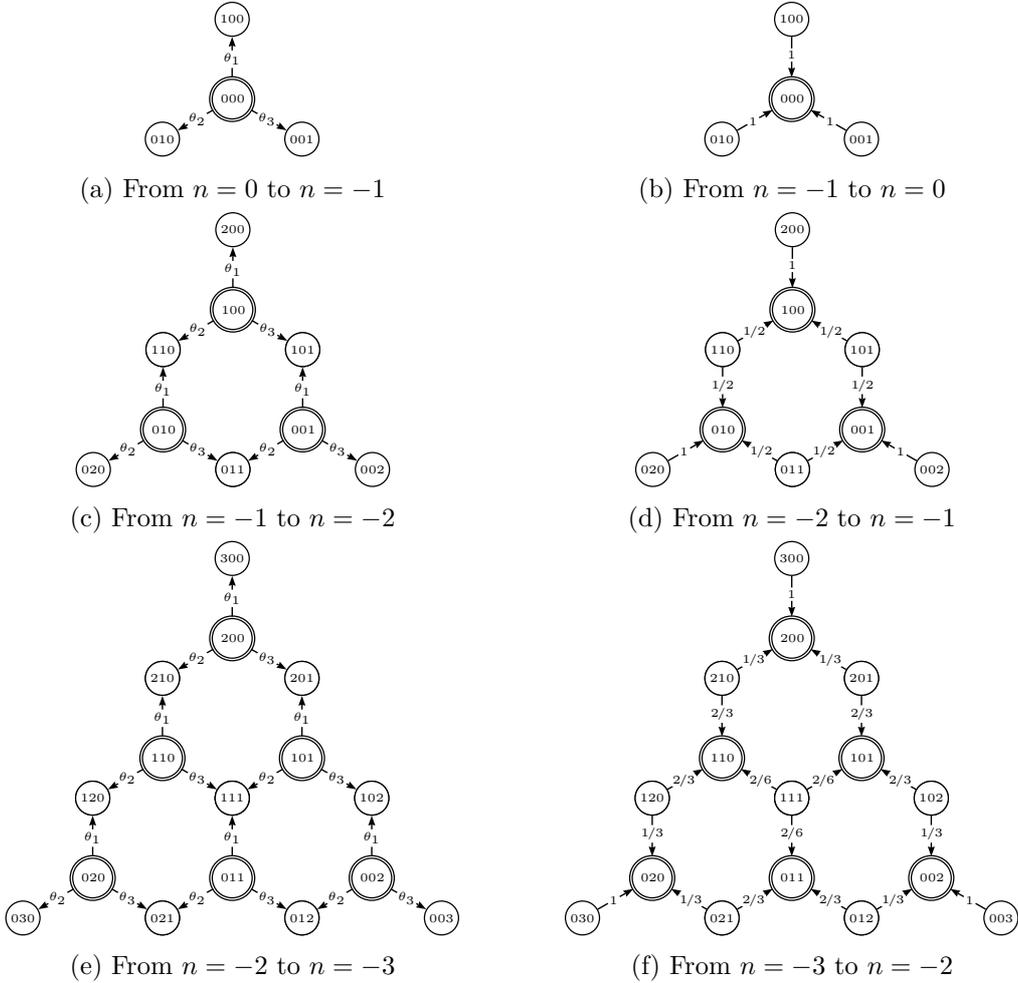


Figure 8: Random walk on the Pascal pyramid ( $d = 3$ )

**Proposition 8.1.** *The  $d$ -dimensional Pascal filtration generated by the process  $(V_n^{d,\theta})_{n \leq 0}$  is standard for any  $d \geq 2$  under any multinomial central measure. Consequently, the multinomial central measures are ergodic.*

We now provide our three different proofs of the above proposition. Another proof is provided in [18], by immersing the filtration in a filtration shown to be standard.

### 8.1 First proof of standardness, using monotonicity of multidimensional Markov processes

Our first proof is an application of Theorem 4.5. Since we know that the tail sigma-algebra  $\mathcal{G}_{-\infty}$  is trivial, it remains to show the monotonicity of the Markov process  $(V_n^{d,\theta})_{n \leq 0}$ . Let us consider two points  $v$  and  $v'$  in  $\mathbf{V}_n^d$ : They satisfy

$$v(1) + \dots + v(d) = v'(1) + \dots + v'(d) = |n|.$$

We want to construct a coupling of  $\mathcal{L}(V_{n+1}^{d,\theta} | V_n^{d,\theta} = v)$  and  $\mathcal{L}(V_{n+1}^{d,\theta} | V_n^{d,\theta} = v')$  which is well-ordered with respect to  $(v, v')$ . We will get this coupling in the form  $(Y, Y') = f_{n+1}(v, v', U)$ , where  $U$  is a uniform random variable on  $\{1, \dots, |n|\}$ . We can easily construct two partitions  $(A_i)_{1 \leq i \leq d}$  and  $(A'_i)_{1 \leq i \leq d}$  of  $\{1, \dots, |n|\}$  such that, for each  $1 \leq i \leq d$ ,  $|A_i| = v(i)$ ,  $|A'_i| = v'(i)$ , and  $A_i = A'_i$  whenever  $v(i) = v'(i)$ . Now, for each  $u \in \{1, \dots, |n|\}$ , there exists a unique pair  $(i, i')$  such that  $u \in A_i \cap A'_{i'}$  and we set  $f_{n+1}(v, v', u) := (v - e_i, v' - e_{i'})$ . In this way, we respect the conditional distribution given in (8.1). Moreover, by construction it is clear that this coupling is well-ordered, since  $v(i) = v'(i)$  implies  $Y(i) = Y'(i)$ , and at each step, coordinates never decrease by more than one unit. Thus, since we know that  $\mathcal{G}_{-\infty}$  is trivial, Theorem 4.5 applies and show that  $\mathcal{G}$  is standard.

## 8.2 Second proof of standardness, computing intrinsic metrics

In the preceding proof, we admitted the degeneracy of  $\mathcal{G}_{-\infty}$ . Here we provide an alternative short proof of standardness of the filtration which does not use this result. We have seen in the preceding proof that the Markov process is monotonic. It is even strongly monotonic (Definition 4.2), thus we can use the tools of Section 4.3. Moreover the Markov process is identifiable (see Section 5.2), hence Theorem 4.8 applies and then in order to derive standardness it suffices to check that  $\mathcal{L}(V_{-1} | \mathcal{F}_n) \rightarrow \mathcal{L}(V_{-1})$ , which is a straightforward consequence of the law of large numbers.

We can use Proposition 4.7 to derive the intrinsic metrics  $\rho_n$ , starting at level  $-1$ . Remembering the unidimensional case, we get

$$\rho_n(v'_n, v''_n) = \sum_{i=1}^d a_i \frac{|v'_n(i) - v''_n(i)|}{|n|}.$$

The  $V'$  property of  $X_{-1}$  (Definition 2.5),

$$\lim_{n \rightarrow -\infty} \mathbb{E}[\rho_n(V'_n, V''_n)] = 0,$$

is, similarly to  $\mathcal{L}(V_{-1} | \mathcal{F}_n) \rightarrow \mathcal{L}(V_{-1})$ , a straightforward consequence of the law of large numbers.

## 8.3 Third proof of standardness, constructing a generating parameterization

The third proof is a little bit longer, but it is self-contained (it does not use the degeneracy of  $\mathcal{G}_{-\infty}$ , nor Theorem 4.8). Moreover, it provides a generating parameterization of the  $d$ -dimensional Pascal filtration.

We start by giving a natural parameterized representation of the Markov process  $(V_n^{d,\theta})_{n \leq 0}$  and we will see that it is generating. We first introduce the notation

$$\bar{v}(i) = \sum_{k=1}^i v(k)$$

for each  $v \in \mathbf{V}_n^d$ , any  $n \leq 0$  and  $i \in \{1, \dots, d\}$ . Recalling the Markovian transition from  $n$  to  $n+1$ , we can easily construct a parameterized representation  $(f_n, U_n)_{n \leq 0}$  for the Markov process  $(V_n^{d,\theta})_{n \leq 0}$  by taking the uniform distribution on  $\{1, \dots, |n|\}$  as the law of  $U_{n+1}$  and by defining the updating functions by

$$f_{n+1}(v, u) := v - e_i,$$

where  $i$  is the unique index such that  $u \in ]\bar{v}(i-1), \bar{v}(i)]$

Now, we point out that, for each  $1 \leq i \leq d-1$ , the process  $(\bar{V}_n^{d,\theta}(i))_{n \leq 0}$  is a Markov process with the same distribution as the process arising in the two-dimensional Pascal graph, that is, with our notations,

$$\mathcal{L}(\bar{V}_n^{d,\theta}(i))_{n \leq 0} = \mathcal{L}(V_n^{2,(p_i, 1-p_i)}(i))_{n \leq 0}$$

where  $p_i := \theta_1 + \dots + \theta_i$ . Moreover, the above parameterized representation of the Markov process  $(V_n^{d,\theta})_{n \leq 0}$  provides a parameterized representation of the Markov process  $(\bar{V}_n^{d,\theta}(i))_{n \leq 0}$ :

$$\bar{V}_{n+1}^{d,\theta}(i) = f_{n+1}^{(i)}(\bar{V}_n^{d,\theta}(i), U_{n+1}) := \begin{cases} \bar{V}_n^{d,\theta}(i) - 1 & \text{if } U_{n+1} \leq \bar{V}_n^{d,\theta}(i) \\ \bar{V}_n^{d,\theta}(i) & \text{otherwise.} \end{cases}$$

This parameterization coincides with the increasing representation of the process that we used in the classical Pascal graph corresponding to  $d = 2$ , hence as we have shown in

Section 6, Theorem 3.6 proves that it is a generating parameterization. It follows that for each  $1 \leq i \leq d$  and each  $n \leq 0$ ,  $\bar{V}_n^{d,\theta}(i)$  is measurable with respect to the  $\sigma$ -algebra generated by  $U_n, U_{n-1}, \dots$ . Thus

$$V_n^{d,\theta}(i) = \bar{V}_n^{d,\theta}(i) - \bar{V}_n^{d,\theta}(i-1)$$

is itself measurable with respect to the same  $\sigma$ -algebra, and the parameterized representation of the Markov process  $(V_n^{d,\theta})_{n \leq 0}$  is generating. Lemma 1.5 then allows us to conclude that the  $d$ -dimensional Pascal filtration is standard.

## References

- [1] H. Crauel. Random Probability Measures on Polish Spaces. CRC Press, 2003.
- [2] S. Bailey Frick, Limited scope adic transformations [arXiv:0708.1328v1](https://arxiv.org/abs/0708.1328v1)
- [3] S. Bailey Frick, K.Petersen. Random permutations and unique fully supported ergodicity for the Euler adic transformation. (English) Ann. Inst. Henri Poincaré, Probab. Stat. 44 (2008), no. 5, 876–885.
- [4] S. Bailey Frick, K.Petersen. Reinforced random walks and adic transformations. J. Theoret. Probab. 23 (2010), no. 3, 920–943.
- [5] S. Bailey Frick, M. Keane, K. Petersen and I. Salama. Ergodicity of the adic transformation on the Euler graph. Math. Proc. Cambridge Philos. Soc. 141 (2006), 231–238.
- [6] P. Berti, L. Pratelli, and P. Rigo. Almost sure weak convergence of random probability measures, Stochastics 78 (2006), no. 2, 91–97.
- [7] S. Bezuglyi, J. Kwiatkowski, K. Medynets, and B. Solomyak. Invariant measures on stationary Bratteli diagrams. Ergod. Th. & Dynam. Sys., 30 (2010), no.4, 973–1007.
- [8] S. Bezuglyi, J. Kwiatkowski, K. Medynets and B. Solomyak. Finite rank Bratteli diagrams: Structure of invariant measures. Trans. Amer. Math. Soc. 365 (2013), 2637–2679
- [9] Coudène, Y.: Une version mesurable du théorème de Stone-Weierstrass. Gazette des mathématiciens 91, 10–17 (2002)
- [10] Émery, M., Schachermayer, W.: On Vershik’s standardness criterion and Tsirelson’s notion of cosiness. Séminaire de Probabilités XXXV, Springer Lectures Notes in Math. 1755, 265–305 (2001)
- [11] Alexander Gnedin and Grigori Olshanski. The boundary of the Eulerian number triangle, Mosc. Math. J. 6 (2006), 461–475.
- [12] É. Janvresse, T. de la Rue. The Pascal Adic Transformation is Loosely Bernoulli. Annales de l’IHP, Probab. Stat., 40 (2004), 133-139.
- [13] Laurent, S.: On standardness and I-cosiness. Séminaire de Probabilités XLIII, Springer Lecture Notes in Mathematics 2006, 127–186 (2010)
- [14] Laurent, S.: 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)
- [15] Laurent, S.: Further comments on the representation problem for stationary processes. Statist. Probab. Lett. 80, 592–596 (2010).
- [16] Laurent, S: Standardness and nonstandardness of next-jump time filtrations. Electronic Communications in Probability 18 (2013), no. 56, 1–11.

- [17] Laurent, S: Uniform entropy scalings of filtrations. Preprint 2014. <https://hal.archives-ouvertes.fr/hal-01006337> <hal-01006337v2>
- [18] Laurent, S: Filtrations of the erased-word processes. Preprint 2014. <https://hal.archives-ouvertes.fr/hal-00999719v2> <hal-00999719v2> (to appear in Séminaire de Probabilités).
- [19] Lindvall, T.: Introduction to the coupling method. Wiley, New-York (1992).
- [20] X. Méla and K. Petersen. Dynamical properties of the Pascal adic transformation. *Ergodic Theory Dynam. Systems*, 25(1):227–256, 2005.
- [21] Karl Petersen and Alexander Varchenko. The Euler Adic Dynamical System and Path Counts in the Euler Graph. *Tokyo J. of Math.* Volume 33, Number 2 (2010), 283–551
- [22] J.G. Propp and D.B. Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures and Algorithms*, 9(1 & 2):223–252, 1996.
- [23] N. Sidorov, Arithmetic dynamics, *Topics in dynamics and ergodic theory*, Lond. Math. Soc. Lect. Note Ser. 310 (2003), 145–189. Cambridge Univ. Press.
- [24] Thorrisson, H.: *Coupling, Stationarity, and Regeneration*. Springer, New York (2000).
- [25] Vershik, A.M.: *Approximation in measure theory* (in Russian). PhD Dissertation, Leningrad University (1973). Expanded and updated version: [27].
- [26] A. M. Vershik. A Theorem on the Markov periodical approximation in ergodic theory. *J. Soviet Math.* 28 (1985), 667–673.
- [27] Vershik, A.M.: 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)
- [28] Vershik, A.M.: Smooth and non-smooth  $AF$ -algebras and problem on invariant measures. arXiv:1304.2193 (2013) <http://arxiv.org/abs/1304.2193>
- [29] A. M. Vershik, *Intrinsic metric on graded graphs, standardness, and invariant measures*. *Zapiski Nauchn. Semin. POMI* 421, 58–67 (2014).
- [30] A.M. Vershik: *The problem of describing central measures on the path spaces of graded graphs*. *Funct. Anal. Appl.* 48, No. 4, 1–20 (2014).