

# On Vershikian and I-cosy random variables and filtrations

Stéphane Laurent \*

July 23, 2010

## Abstract

We prove that the equivalence between Vershik's standardness criterion and the I-cosiness criterion for a filtration in discrete, negative time, holds separately for each random variable. This gives a strengthening and a more direct proof of the global equivalence between these two criteria. We also provide more elementary original propositions on Vershik's standardness criterion, while emphasizing that similar statements for I-cosiness are sometimes not so obvious.

**Acknowledgments.** I am indebted to the IAP research network for financial support (grant nr. P6/03 of the Belgian government, Belgian Science Policy), and to M. Émery for lively discussions on this work and helpful comments on earlier drafts of this paper.

## Contents

|   |   |    |
|---|---|----|
| 1 | Introduction                              | 2  |
| 2 | Preliminary notions                       | 4  |
| 3 | Vershik's progressive predictions         | 11 |
| 4 | The Vershik property                      | 16 |
| 5 | Vershik's theorem on lacunary isomorphism | 22 |
| 6 | Parametric extensions                     | 23 |
| 7 | The key lemma                             | 27 |
| 8 | Vershik property and I-cosiness           | 31 |
| 9 | Coparameterized joinings                  | 33 |

---

\*Université catholique de Louvain, Louvain-la-Neuve, Belgium.

# 1 Introduction

One of Vershik's achievements in his theory of filtrations in discrete, negative time, was to provide a criterion (sometimes called the *combinatorial standardness criterion*) for a *homogeneous* filtration to be generated by independent random variables ([Ver70, Ver73a, Ver94]).

In [Ver70] he devised an equivalent formulation of his criterion in terms of *universal projectors*; we call them *Vershik's progressive predictions*. This property, named *Vershik's standardness criterion* in [ES01] (but we simply call it the *Vershik property*), makes sense not only for homogeneous filtrations, but for arbitrary ones, and defines an invariant of filtrations. Vershik announced in [Ver94] that, under the natural assumption that the final  $\sigma$ -field  $\mathcal{F}_0$  is essentially separable (i.e., countably generated modulo null events), the Vershik property characterizes filtrations  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  in discrete, negative time, which can be *immersed* in the filtration generated by independent random variables, termed as *standard filtrations*. A proof of this theorem is provided in [ES01].

The present paper, in which homogeneous filtrations are not even defined, only deals with the Vershik property for arbitrary filtrations. We provide in particular a self-contained proof that the Vershik property is equivalent to the *I-cosiness criterion* for a filtration with an essentially separable final  $\sigma$ -field. We actually prove a stronger result: the Vershik property and the I-cosiness criterion are equivalent not only for a filtration, but for a random variable. This is the content of assertion (1.1), to be explained below.

Moreover we relax the assumption that the final  $\sigma$ -field of the filtration is essentially separable: we only assume the filtration to be *locally separable*, i.e., to admit essentially separable increments. This generalization from global separability to local separability has no practical interest, but it turns out that our proofs only need this weaker hypothesis, so we have no reason to use the stronger one. If we assumed an essentially separable final  $\sigma$ -field, all our results would remain just as interesting, and our proofs would not become simpler. Hereafter in this introduction, it is understood that we consider the context of locally separable filtrations, but nothing is lost by restricting discussion to filtrations with an essentially separable final  $\sigma$ -field.

I-cosiness is a variant of Tsirelson's notion of *cosiness* which he defined in [Tsi97] in the framework of continuous time. It was introduced by Émery and Schachermayer in [ES01], and proved to be equivalent to the Vershik property (when the final  $\sigma$ -field is essentially separable). Whereas it is obvious from

the definitions that I-cosiness implies the Vershik property, the converse is less straightforward. The proof given by [ES01] uses other non-obvious theorems as intermediate steps. Compared to that proof, ours is much more direct, and in fact is a corollary of the equivalence between both properties *separately for each random variable*, in the following sense. These two properties have the same structure: they require a certain assertion,  $I(X)$  or  $V(X)$ , to hold for any test random variable  $X$  measurable with respect to the final  $\sigma$ -field  $\mathcal{F}_0$ . Hence, the equivalence between both properties has the form

$$\left(\forall X \quad I(X)\right) \iff \left(\forall X \quad V(X)\right),$$

whereas we will establish the stronger assertion

$$\forall X \quad \left[ I(X) \iff V(X) \right]. \tag{1.1}$$

The proof that  $I(X)$  implies  $V(X)$  is easily deduced from the definitions. Our proof of the converse is a direct application of a technical lemma. Another application of this lemma, concerning *joinings* of filtrations, is given at the end.

We will also provide more or less elementary results on the Vershik property for random variables and filtrations. For example, we will show that the property  $V(X)$  that the random variable  $X$  is Vershikian, is equivalent to some filtration being Vershikian, namely, the filtration generated by the *progressive predictions*  $\pi_n X$  of  $X$ , which lie at the heart of the Vershik property.

Another property is shown to be equivalent to  $I(X)$  in [Lau10]<sup>1</sup>, where the Vershik property is not used (and not even stated), and where standardness and I-cosiness are showed to be equivalent. The present paper and [Lau10] complement each other; however we will not refer to [Lau10], except for some elementary lemmas, most of them displayed as preliminaries in the next section. Throughout the paper, in light of the new results, we will sometimes comment on the results of [Lau10].

In particular, having in view our main result (1.1), we will emphasize the benefits of the Vershik property as compared to I-cosiness, by noticing that some results proved with the help of  $V(X)$ , would be more difficult or less natural to establish via  $I(X)$ . For example we give a short proof of *Vershik's theorem on lacunary isomorphism*, which asserts that, if a filtration  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  has a degenerate tail  $\sigma$ -field  $\mathcal{F}_{-\infty} = \bigcap_n \mathcal{F}_n$  and an essentially

---

<sup>1</sup> This is Corollary 4.8 in [Lau10]. It involves *Vershik's first level criterion*, a property also due to Vershik, which is not stated in the present paper.

separable final  $\sigma$ -field  $\mathcal{F}_0$ , then, for some sub-sequence  $n \mapsto \phi(n)$ , the extracted filtration  $(\mathcal{F}_{\phi(n)})_{n \leq 0}$  is Vershikian. We do not know such an easy proof of this theorem using  $I(X)$  instead of  $V(X)$ .

## 2 Preliminary notions

We first give as preliminaries the main notations and conventions of this paper. Next we introduce the notion of *extension* of a filtration. Finally we give some notions and elementary lemmas borrowed from [Lau10]: we define the property of *local separability* for a filtration and give its useful characterization in terms of a *parameterization* for this filtration, and we define the *I-cosiness criterion* as a particular *self-joining criterion*. There are two other self-joining criteria in [Lau10], namely *Rosenblatt's self-joining criterion* and *Vershik's self-joining criterion*. Other ones are defined in [ES01].

### Main notations and conventions

By a probability space, we always mean a triple  $(\Omega, \mathcal{A}, \mathbb{P})$  where the  $\sigma$ -field  $\mathcal{A}$  is  $\mathbb{P}$ -complete. By a  $\sigma$ -field  $\mathcal{C} \subset \mathcal{A}$  we always mean an  $(\mathcal{A}, \mathbb{P})$ -complete  $\sigma$ -field. By a random variable on  $(\Omega, \mathcal{A}, \mathbb{P})$ , we mean the  $\mathbb{P}$ -equivalence class of some measurable map from  $\Omega$  to a separable metric space. The Borel  $\sigma$ -field on a separable metric space  $E$  is denoted by  $\mathfrak{B}_E$ . A  $\sigma$ -field  $\mathcal{C}$  is *essentially separable* if it is generated by a random variable; equivalently,  $\mathcal{C}$  is countably generated up to negligible sets. By convention, the  $\sigma$ -field generated by an empty family of random variables is the trivial  $\sigma$ -field  $\{\emptyset, \Omega\}$  up to negligible sets. When  $X$  is a random variable taking values in a Polish space  $E$ , the existence of the conditional law of  $X$  given any  $\sigma$ -field  $\mathcal{C}$  is guaranteed; we denote it by  $\mathcal{L}[X | \mathcal{C}]$ . The space  $E'$  of all probability measures on  $E$  is itself Polish for the *weak topology* inherited from the Polish topology of  $E$  (see [Bil]), and then  $\mathcal{L}[X | \mathcal{C}]$  is a  $\mathcal{C}$ -measurable  $E'$ -valued random variable. We use the notation  $L^0(\mathcal{C}; (E, \rho))$  or, shorter  $L^0(\mathcal{C}; E)$ , to denote the metrizable topological space of all  $\mathcal{C}$ -measurable random variables taking their values in a separable metric space  $(E, \rho)$ ; the space  $L^0(\mathcal{C}; E)$  is endowed with the topology of convergence in probability; when  $E = \mathbb{R}$  we simply call it  $L^0(\mathcal{C})$ . We use the same notations for  $L^1$  spaces; the space  $L^1(\mathcal{C}; E)$  is the set of all  $\mathcal{C}$ -measurable random variables  $X$  taking their values in  $E$  and such that  $\mathbb{E}[\rho(X, x)]$  is finite for some ( $\Leftrightarrow$  for all)  $x \in E$ ; the space  $L^1(\mathcal{C}; E)$  is endowed with the metric  $(X, Y) \mapsto \mathbb{E}[\rho(X, Y)]$ . We will implicitly use the

fact that  $\bigcup_m L^1(\mathcal{B}_m; E)$  is dense in  $L^1(\bigvee_m \mathcal{B}_m; E)$  for any increasing sequence  $(\mathcal{B}_m)_{m \in \mathbb{N}}$  of  $\sigma$ -fields (see lemma 2.12 in [Lau10]).

*Filtrations in discrete, negative time* — On an underlying probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ , a filtration is an increasing sequence of sub- $\sigma$ -fields of  $\mathcal{A}$  indexed by a time-axis. Most filtrations encountered in this paper are indexed by the time axis  $-\mathbb{N} = \{\dots, -2, -1, 0\}$ . If the time axis of a filtration  $\mathcal{F}$  is not specified, it will be understood that  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  is a filtration in discrete, negative time. We say that a filtration  $\mathcal{F}$  is *essentially separable* if the final  $\sigma$ -field  $\mathcal{F}_0$  (or equivalently, each  $\sigma$ -field  $\mathcal{F}_n$ ) is essentially separable. We say that a filtration  $\mathcal{F}$  is *Kolmogorovian* if the tail  $\sigma$ -field  $\mathcal{F}_{-\infty} := \bigcap_{n \leq 0} \mathcal{F}_n$  equals the trivial  $\sigma$ -field  $\{\emptyset, \Omega\}$  up to negligible sets. A filtration  $\mathcal{F}$  is *included* in a filtration  $\mathcal{G}$ , and we write  $\mathcal{F} \subset \mathcal{G}$ , if  $\mathcal{F}_n \subset \mathcal{G}_n$  for each  $n \leq 0$ . The *supremum*  $\mathcal{F} \vee \mathcal{G}$  of two filtrations  $\mathcal{F}$  and  $\mathcal{G}$  is the smallest filtration containing both  $\mathcal{F}$  and  $\mathcal{G}$ ; it is given by  $(\mathcal{F} \vee \mathcal{G})_n = \mathcal{F}_n \vee \mathcal{G}_n$ .

*Isomorphic  $\sigma$ -fields and filtrations* — An *embedding*  $\Psi$  from a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$  into a probability space  $(\Omega', \mathcal{A}', \mathbb{P}')$  is a (necessarily injective) map from the quotient  $\sigma$ -field  $\mathcal{A}/\mathbb{P}$  to the quotient  $\sigma$ -field  $\mathcal{A}'/\mathbb{P}'$  that preserves the  $\sigma$ -field structure and the probabilities. We write shortly  $\Psi: \mathcal{A} \rightarrow \mathcal{A}'$ . If in addition  $\Psi$  is onto then it is called an *isomorphism*. Up to isomorphism, an essentially separable  $\sigma$ -field is characterized by the descending sequence (possibly empty, finite, or denumerable) of the masses of its atoms. An embedding  $\Psi$  extends uniquely to random variables taking their values in a Polish space  $E$ , and thus defines a map from  $L^0(\mathcal{A}; E)$  to  $L^0(\mathcal{A}'; E)$  preserving the law of random variables. Given a random variable  $X \in L^0(\mathcal{A}; E)$  we call  $\Psi(X)$  the image of  $X$  under this map, and we call it the *copy* of  $X$ . One has  $\Psi(f(X)) = f(\Psi(X))$  for any measurable function  $f$  from a Polish space to another one. Details are provided in Annex A of [Lau10]. We will use the following lemmas and propositions from this Annex.

**Lemma 2.1.** *Let  $(\Omega, \mathcal{B}, \mathbb{P})$  and  $(\Omega', \mathcal{A}', \mathbb{P}')$  be two probability spaces and  $\Psi: \mathcal{B} \rightarrow \mathcal{A}'$  an embedding. Let  $E$  be a Polish space,  $X \in L^0(\mathcal{B}; E)$  and  $\mathcal{C} \subset \mathcal{B}$  be a  $\sigma$ -field. Then  $\Psi(\mathcal{L}[X | \mathcal{C}]) = \mathcal{L}[\Psi(X) | \Psi(\mathcal{C})]$ .*

**Proposition 2.2.** *Let  $(\Omega, \mathcal{B}, \mathbb{P})$  be a probability space and  $\mathcal{C}_1, \mathcal{C}_2$  two sub- $\sigma$ -fields of  $\mathcal{B}$ . Let  $(\Omega', \mathcal{A}', \mathbb{P}')$  be a probability space, and  $\Psi_1: \mathcal{C}_1 \rightarrow \mathcal{A}'$ ,  $\Psi_2: \mathcal{C}_2 \rightarrow \mathcal{A}'$  two embeddings. There exists an isomorphism  $\Psi: \mathcal{C}_1 \vee \mathcal{C}_2 \rightarrow \Psi_1(\mathcal{C}_1) \vee \Psi_2(\mathcal{C}_2)$  which simultaneously extends  $\Psi_1$  and  $\Psi_2$  if and only if one has  $\Psi_1(\mathcal{L}[C_2 | \mathcal{C}_1]) = \mathcal{L}[\Psi_2(C_2) | \Psi_1(\mathcal{C}_1)]$  for every  $\mathcal{C}_2$ -measurable random variable  $C_2$ .*

**Proposition 2.3.** *Let  $(\Omega, \mathcal{B}, \mathbb{P})$  be a probability space,  $\mathcal{C} \subset \mathcal{B}$  a  $\sigma$ -field, and  $V$  a  $\mathcal{B}$ -measurable random variable taking values in some Polish space  $E$ . Let  $(\Omega', \mathcal{A}', \mathbb{P}')$  be a probability space,  $\Psi_0: \mathcal{C} \rightarrow \mathcal{A}'$  an embedding, and  $V'$  a random variable taking values in  $E$ . Then there exists an isomorphism  $\Psi: \mathcal{C} \vee \sigma(V) \rightarrow \Psi_0(\mathcal{C}) \vee \sigma(V')$  extending  $\Psi_0$  and sending  $V$  to  $V'$  if and only if one has  $\Psi_0(\mathcal{L}[V | \mathcal{C}]) = \mathcal{L}[V' | \Psi_0(\mathcal{C})]$ .*

**Lemma 2.4.** *On a probability space  $(\Omega, \mathcal{C}, \mathbb{P})$ , let  $\mu$  be a random probability on a Polish space  $E$ . We define the probability  $\hat{\mathbb{P}} := \mathbb{P} \otimes \mu$  on the measurable space  $(\hat{\Omega}, \hat{\mathcal{B}}) := (\Omega \times E, \mathcal{C} \otimes \mathfrak{B}_E)$  by*

$$\hat{\mathbb{P}}[\hat{B}] = \mathbb{E} \left[ \int \mathbb{1}_{\hat{B}}(\cdot, t) d\mu(t) \right].$$

*Then the identification with the first factor  $\iota: \mathcal{C} \rightarrow \hat{\mathcal{B}}$  is an embedding from  $(\Omega, \mathcal{C}, \mathbb{P})$  into  $(\hat{\Omega}, \hat{\mathcal{B}}, \hat{\mathbb{P}})$ , and we have  $\iota(\mu) = \mathcal{L}[\hat{V} | \hat{\mathcal{C}}]$  where  $\hat{\mathcal{C}} = \iota(\mathcal{C})$  and  $\hat{V}$  is the random variable defined by  $\hat{V}(\omega, t) = t$ .*

The definition of isomorphic  $\sigma$ -fields extends naturally to filtrations as follows. Two filtrations  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  and  $\mathcal{F}' = (\mathcal{F}'_n)_{n \leq 0}$ , defined on possibly different probability spaces, are said to be *isomorphic* if there is an isomorphism  $\Psi: \mathcal{F}_0 \rightarrow \mathcal{F}'_0$  such that  $\Psi(\mathcal{F}_n) = \mathcal{F}'_n$  for every  $n \leq 0$ . We shortly say that  $\Psi: \mathcal{F} \rightarrow \mathcal{F}'$  is an isomorphism. We denote by  $\Psi(\mathcal{F})$  the filtration  $\mathcal{F}'$  and we call it the *copy* of the filtration  $\mathcal{F}$  by the isomorphism  $\Psi$ . A typical example of isomorphic filtrations  $\mathcal{F}$  and  $\mathcal{F}'$  is when  $\mathcal{F}$  and  $\mathcal{F}'$  are respectively generated by two processes  $(X_n)_{n \leq 0}$  and  $(X'_n)_{n \leq 0}$  having the same law.

### Extension of a filtration

We say that a filtration  $\mathcal{F}$  is *immersed* in a filtration  $\mathcal{G}$  if  $\mathcal{F} \subset \mathcal{G}$  and if every  $\mathcal{F}$ -martingale is a  $\mathcal{G}$ -martingale; the notation  $\mathcal{F} \stackrel{m}{\subset} \mathcal{G}$  means that  $\mathcal{F}$  is immersed in  $\mathcal{G}$ . Usual characterizations of immersion are given in lemma 2.5. These facts are elementary exercises given without proof. We also say that  $\mathcal{G}$  is an *extension* of  $\mathcal{F}$ . More generally, an *extension* of  $\mathcal{F}$  is a filtration  $\mathcal{G}'$  not necessarily defined on the same probability space as  $\mathcal{F}$ , such that  $\mathcal{F}$  has an isomorphic copy  $\mathcal{F}'$  immersed in  $\mathcal{G}'$ . We also say that  $\mathcal{F}$  is *immersible* in  $\mathcal{G}'$ . We say that two filtrations  $\mathcal{F}'$  and  $\mathcal{F}''$  are *jointly immersed* if they are both immersed in  $\mathcal{F}' \vee \mathcal{F}''$ , or, equivalently, in some filtration.

**Lemma 2.5.** *Let  $\mathcal{F}$  and  $\mathcal{G}$  be two filtrations on a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . The following conditions are equivalent:*

- (i)  $\mathcal{F}$  is immersed in  $\mathcal{G}$ ;

- (ii)  $\mathcal{F}$  is included in  $\mathcal{G}$  and for each  $n \leq 0$ , the  $\sigma$ -fields  $\mathcal{F}_0$  and  $\mathcal{G}_n$  are conditionally independent given  $\mathcal{F}_n$ ;
- (iii) for every  $\mathcal{F}_0$ -measurable random variable  $Y$  taking its values in a Polish space, one has  $\mathcal{L}[Y | \mathcal{G}_n] = \mathcal{L}[Y | \mathcal{F}_n]$  for each  $n \leq 0$ ;
- (iv) for every random variable  $X \in L^1(\mathcal{F}_0)$ , one has  $\mathbb{E}[X | \mathcal{G}_n] = \mathbb{E}[X | \mathcal{F}_n]$  for each  $n \leq 0$ .

Note also that immersion of  $\mathcal{F}$  in  $\mathcal{G}$  implies  $\mathcal{F}_n = \mathcal{F}_0 \cap \mathcal{G}_n$  for all  $n \leq 0$ . The three lemmas below are copied verbatim from [Lau10]. The third one is a direct consequence of the first two ones.

**Lemma 2.6.** *A filtration  $\mathcal{F}$  is immersed in a filtration  $\mathcal{G}$  if and only if  $\mathcal{F} \subset \mathcal{G}$  and for every integer  $n < 0$ , the  $\sigma$ -fields  $\mathcal{F}_{n+1}$  and  $\mathcal{G}_n$  are conditionally independent given  $\mathcal{F}_n$ .*

**Lemma 2.7.** *If  $\mathcal{B}$ ,  $\mathcal{C}$  and  $\mathcal{D}$  are three  $\sigma$ -fields such that  $\mathcal{B}$  and  $\mathcal{C}$  are conditionally independent given  $\mathcal{D}$ , then  $\mathcal{D} \vee \mathcal{B}$  and  $\mathcal{C}$  are also conditionally independent given  $\mathcal{D}$ .*

**Lemma 2.8.** *Let  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  and  $\mathcal{G} = (\mathcal{G}_n)_{n \leq 0}$  be two filtrations such that  $\mathcal{F} \subset \mathcal{G}$ . Let  $(V_n)_{n \leq 0}$  be a process such that  $\mathcal{F}_n \subset \mathcal{F}_{n-1} \vee \sigma(V_n)$  for every  $n \leq 0$ . If  $V_n$  is conditionally independent of  $\mathcal{G}_{n-1}$  given  $\mathcal{F}_{n-1}$  for every  $n \leq 0$ , then  $\mathcal{F}$  is immersed in  $\mathcal{G}$ .*

As a consequence of the lemma 2.8, a process  $(V_n)_{n \leq 0}$  is Markovian with respect to a filtration  $\mathcal{G}$  if and only if it is Markovian and its generated filtration  $\mathcal{F}$  is immersed in  $\mathcal{G}$ .

### Superinnovations and locally separable filtrations.

The standing assumption on filtrations considered in this paper is *local separability*, defined below. It is clear from this definition that essentially separable filtrations, that is, filtrations with an essentially separable final  $\sigma$ -field, are always locally separable.

**Definition 2.9.** A filtration  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  is *locally separable* if for each  $n \leq 0$ , there exists a random variable  $V_n$  such that  $\mathcal{F}_n = \mathcal{F}_{n-1} \vee \sigma(V_n)$ .

A convenient tool for dealing with locally separable filtrations is the existence of a *parameterization* for such a filtration, stated in lemma 2.13 which is borrowed from [Lau10].

**Definition 2.10.** Let  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  be a filtration and  $n_0 < 0$  an integer or  $n_0 = -\infty$ . A *superinnovation* of  $\mathcal{F}$  from  $n_0$  to 0 is a sequence of random variables  $\mathbf{V} = (V_n)_{n \in ]n_0, 0] \cap \mathbb{Z}}$  such that for each  $n \in ]n_0, 0] \cap \mathbb{Z}$ , the random variable  $V_n$  takes its values in a Polish space, is independent of  $\mathcal{F}_{n-1} \vee \sigma(V_m; n_0 < m \leq n-1)$ , and satisfies  $\mathcal{F}_n \subset \mathcal{F}_{n-1} \vee \sigma(V_n)$ . We also say that  $\mathbf{V}$  is a *parameterization* in the case when each  $V_n$  is uniformly distributed on  $[0, 1]$ . We say more specifically that  $\mathbf{V}$  is a *local superinnovation* if  $n_0 > -\infty$  and a *global superinnovation* if  $n_0 = -\infty$ .

**Lemma 2.11.** Let  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  be a filtration and  $\mathbf{V} = (V_n)_{n \in ]n_0, 0] \cap \mathbb{Z}}$  a superinnovation of  $\mathcal{F}$  where  $n_0 < 0$  is an integer or  $n_0 = -\infty$ . Let  $\mathcal{V} = (\mathcal{V}_n)_{n \leq 0}$  be the filtration generated by  $\mathbf{V}$  on the time-axis  $-\mathbb{N}$  (in particular  $\mathcal{V}_n$  is degenerate for  $n \leq n_0$ ). Then  $\mathcal{F}$  and  $\mathcal{V}$  are jointly immersed (in  $\mathcal{F} \vee \mathcal{V}$ ).

*Proof.* This is an easy consequence of lemma 2.6 and lemma 2.7.  $\square$

**Definition 2.12.** The filtration  $\mathcal{F} \vee \mathcal{V}$  defined in the previous lemma is called the *extension of  $\mathcal{F}$  with the superinnovation  $\mathbf{V}$* , and is called a *parametric extension* of  $\mathcal{F}$ .

Note that, as  $\mathcal{F}$  is immersed in its extension  $\mathcal{F} \vee \mathcal{V}$ , any joining (to be soon defined) of  $\mathcal{F} \vee \mathcal{V}$  induces a joining of  $\mathcal{F}$ . We shall see in section 9 that conversely, in some cases, a joining of  $\mathcal{F}$  can be extended to a joining of  $\mathcal{F} \vee \mathcal{V}$  (proposition 9.2).

**Lemma 2.13.** Let  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  be a filtration. The following conditions are equivalent:

- (i)  $\mathcal{F}$  is locally separable;
- (ii) for each  $n \leq 0$ , one has  $\mathcal{F}_n \subset \mathcal{F}_{n-1} \vee \sigma(W_n)$  for some random variable  $W_n$ ;
- (iii)  $\mathcal{F}$  admits a global parameterization, that is, there exists a global parameterization of a filtration isomorphic to  $\mathcal{F}$ ;
- (iv)  $\mathcal{F}$  admits a global superinnovation.

For a filtration  $\mathcal{F}$  generated by a Markovian 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 that  $\sigma(X_n) \subset \sigma(X_{n-1}, U_n)$ . This fact is a particular case of lemma 2.15 below, borrowed from [Lau10].



**Definition 2.14.** Let  $(X_n)_{n \leq 0}$  be a process on  $(\Omega, \mathcal{A}, \mathbb{P})$ , and let  $\phi: -\mathbb{N} \rightarrow -\mathbb{N}$  be a strictly increasing map with  $\phi(0) = 0$ .

1. We say that  $\phi$  is a *sequence of memory-loss times of type I* for  $(X_n)_{n \leq 0}$  if  $X_n$  is conditionally independent of  $\sigma(X_m; m < n)$  given  $(X_{\phi(k-1)}, \dots, X_{n-1})$  for every  $k, n \in -\mathbb{N}$  satisfying  $\phi(k-1) < n \leq \phi(k)$ .
2. Let  $\mathcal{F}$  be the filtration generated by  $(X_n)_{n \leq 0}$ . We say that  $\phi$  is a *sequence of memory-loss times of type II* for  $(X_n)_{n \leq 0}$  if there exist a probability space  $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$ , an embedding  $\Psi: \mathcal{F}_0 \rightarrow \bar{\mathcal{A}}$ , and a parameterization  $(U'_n)_{n \leq 0}$  of the filtration  $\mathcal{F}' := \Psi(\mathcal{F})$  such that

$$\Psi(\sigma(X_n)) \subset \Psi(\sigma(X_{\phi(k-1)})) \vee \sigma(U'_{\phi(k-1)+1}, \dots, U'_n)$$

for every  $k, n \in -\mathbb{N}$  satisfying  $\phi(k-1) < n \leq \phi(k)$ .

Obviously, a process is Markovian if and only if it admits the identity map  $\phi: -\mathbb{N} \rightarrow -\mathbb{N}$  as a sequence of memory-loss times of type I.

**Lemma 2.15.** *If  $\phi$  is a sequence of memory-loss times of type I for  $(X_n)_{n \leq 0}$ , then  $\phi$  is a sequence of memory-loss times of type II for  $(X_n)_{n \leq 0}$ .*

### Joinings, self-joining criterion, I-cosiness

A *joining* of a filtration  $\mathcal{F}$  is a pair  $(\mathcal{F}', \mathcal{F}'')$  of filtrations isomorphic to  $\mathcal{F}$ , defined on the same probability space and *jointly immersed*. (Recall that joint immersion means that both  $\mathcal{F}'$  and  $\mathcal{F}''$  are immersed in  $\mathcal{F}' \vee \mathcal{F}''$ , or, equivalently, in some filtration.) For example,  $(\mathcal{F}', \mathcal{F}'')$  is a joining of  $\mathcal{F}$  when  $\mathcal{F}'$  and  $\mathcal{F}''$  are two independent filtrations both isomorphic to  $\mathcal{F}$ .

Thus, rigorously, a joining is a pair  $(\Psi'(\mathcal{F}), \Psi''(\mathcal{F}))$  given by a probability space  $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$  and two embeddings  $\Psi': \mathcal{F}_0 \rightarrow \bar{\mathcal{A}}$  and  $\Psi'': \mathcal{F}_0 \rightarrow \bar{\mathcal{A}}$  with the additional property of joint immersion. Given a filtration  $\mathcal{E}$  immersed in  $\mathcal{F}$ , a joining  $(\mathcal{F}', \mathcal{F}'')$  of  $\mathcal{F}$  induces a joining  $(\mathcal{E}', \mathcal{E}'')$  of  $\mathcal{E}$ .

Below we define the notion of *self-joining criterion* for a filtration. This notion generalizes an idea due to Tsirelson in the framework of continuous time (*cosiness*, introduced in [Tsi97]); it does not appear elsewhere as such in the literature, but it includes several known criteria. The main one is the *I-cosiness criterion*. Our motivation to define the general notion of a self-joining criterion is that many elementary properties of I-cosiness are actually true for any self-joining criterion. Two other self-joining criteria are

considered in [Lau10], namely *Rosenblatt's self-joining criterion* and *Ver-shik's self-joining criterion*. Contrary to the I-cosiness self-joining criterion, these ones are not *hereditary* according to definition 2.17. Some hereditary self-joining criteria are defined in [ES01] in addition to I-cosiness.

**Definition 2.16.** Let  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  be a filtration on a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . Given a property  $(\mathcal{S})$  on the collection of all joinings  $(\mathcal{F}', \mathcal{F}'')$  of  $\mathcal{F}$ , we define the corresponding *self-joining criterion* for  $\mathcal{F}$  as follows.

1. Let  $(E, \rho)$  be a Polish metric space and  $X \in L^1(\mathcal{F}_0; E)$ . The self-joining criterion holds for  $X$  (with respect to  $\mathcal{F}$ ) if for each real number  $\delta > 0$ , there exist two filtrations  $\mathcal{F}'$  and  $\mathcal{F}''$  defined on a probability space  $(\overline{\Omega}, \overline{\mathcal{A}}, \overline{\mathbb{P}})$  such that:
  - (i)  $(\mathcal{F}', \mathcal{F}'')$  is a joining of  $\mathcal{F}$  satisfying property  $(\mathcal{S})$  (this will be denoted by  $\mathcal{S}(\mathcal{F}', \mathcal{F}'')$ );
  - (ii) one has  $\overline{\mathbb{E}}[\rho(X', X'')] < \delta$ , where  $X'$  and  $X''$  are the respective copies of  $X$  in  $\mathcal{F}'$  and in  $\mathcal{F}''$ .
2. We say that a  $\sigma$ -field  $\mathcal{E}_0 \subset \mathcal{F}_0$  satisfies this self-joining-criterion (with respect to  $\mathcal{F}$ ) if every random variable  $X \in L^1(\mathcal{E}_0)$  satisfies this self-joining-criterion with respect to  $\mathcal{F}$ .
3. We say that the filtration  $\mathcal{F}$  satisfies this self-joining-criterion if the final  $\sigma$ -field  $\mathcal{F}_0$  satisfies this self-joining-criterion with respect to  $\mathcal{F}$ .

**Definition 2.17.** Let  $(\mathcal{S})$  be a property on the collection of joinings of filtrations (thus  $(\mathcal{S})$  defines a self-joining criterion for any filtration). We say that  $(\mathcal{S})$  defines a *hereditary self-joining criterion* if for any filtrations  $\mathcal{E}$  and  $\mathcal{F}$  such that  $\mathcal{E} \stackrel{m}{\subset} \mathcal{F}$ , the implication  $\mathcal{S}(\mathcal{F}', \mathcal{F}'') \implies \mathcal{S}(\mathcal{E}', \mathcal{E}'')$  holds for any joining  $(\mathcal{F}', \mathcal{F}'')$  of  $\mathcal{F}$ .

**Definition 2.18.** The *I-cosiness criterion* is the hereditary self-joining criterion defined by

$$\mathcal{S}(\mathcal{F}', \mathcal{F}'') = \text{“}\mathcal{F}' \text{ and } \mathcal{F}'' \text{ are independent in small time”},$$

in the sense that the  $\sigma$ -fields  $\mathcal{F}'_{n_0}$  and  $\mathcal{F}''_{n_0}$  are independent for some integer  $n_0 \leq 0$ . We also say that the joining  $(\mathcal{F}', \mathcal{F}'')$  is *independent up to*  $n_0$ .

A random variable, or a  $\sigma$ -field, or a filtration which satisfies this criterion will be called *I-cosy*.

The following lemma is obvious from the definition of a hereditary self-joining criterion. Corollary 8.4 asserts that the converse (of the first assertion) of this lemma holds for I-cosiness under the hypothesis of local separability of the filtrations. In fact we shall see that the similar property and its converse are obviously true for the Vershik property (lemma 4.3), and we shall see that the Vershik property is equivalent to I-cosiness for locally separable filtrations (assertion (1.1) and theorem 8.2).

**Lemma 2.19.** *Let  $\mathcal{E} \overset{m}{\subset} \mathcal{F}$ . Any  $\mathcal{E}_0$ -measurable random variable satisfying a hereditary self-joining criterion with respect to  $\mathcal{F}$  satisfies this self-joining criterion with respect to  $\mathcal{E}$  too. Consequently  $\mathcal{E}$  satisfies this self-joining criterion whenever  $\mathcal{F}$  satisfies it.*

The following lemma is proved in [Lau10] in the particular case of I-cosiness, but its proof is actually valid for any self-joining criterion.

**Lemma 2.20.** *A random variable satisfies a self-joining criterion if and only if the  $\sigma$ -field  $\sigma(X)$  satisfies this self-joining criterion. Consequently, for the filtration  $\mathcal{F}$  to satisfy a self-joining criterion, it suffices that there exists a random variable  $X$  that generates  $\mathcal{F}_0$  and satisfies this self-joining criterion.*

We will also use the following lemma which easily stems from lemma 2.20.

**Lemma 2.21.** *Let  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  be a filtration and consider a self-joining criterion for  $\mathcal{F}$ . Let  $\mathcal{E}_0 \subset \mathcal{F}_0$  be a  $\sigma$ -field. Then  $\mathcal{E}_0$  satisfies the self-joining criterion if and only if each random variable  $X \in L^1(\mathcal{E}_0; [0, 1])$  satisfies the self-joining criterion.*

### 3 Vershik's progressive predictions

The Vershik property of a filtration  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  involves *Vershik's progressive predictions*  $\pi_n X$  which we introduce here. They are equivalently termed as *universal projectors* in [Ver70] and [Ver94]. Given a random variable  $X$  measurable with respect to the final  $\sigma$ -field  $\mathcal{F}_0$ , the  $n$ -th progressive prediction  $\pi_n X$  of  $X$  is our “best knowledge” about  $X$  at time  $n$ . We shall see that the process  $(\pi_n X)_{n \leq 0}$  of progressive predictions generates the smallest filtration immersed in  $\mathcal{F}$  and making  $X$  measurable, in the sense that  $X$  is measurable with respect to the final  $\sigma$ -field (lemma 3.7).

Let  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  be a filtration and  $X \in L^0(\mathcal{F}_0; E)$  where  $E$  is a Polish space. The existence of the conditional law  $\mathcal{L}[X | \mathcal{F}_{-1}]$  is guaranteed by the

Polish assumption, and  $\mathcal{L}[X | \mathcal{F}_{-1}]$  is then a random variable in the space  $E'$  of probability measures on  $E$ , which is itself Polish when endowed with the weak topology on  $E'$ . So it is possible to define the conditional law  $\mathcal{L}[\mathcal{L}[X | \mathcal{F}_{-1}] | \mathcal{F}_{-2}]$  of the conditional law  $\mathcal{L}[X | \mathcal{F}_{-1}]$  given  $\mathcal{F}_{-2}$ , which is a random variable in the Polish space  $(E')'$  of all probability measures on  $E'$ . Following Vershik ([Ver70, Ver71, Ver94]), the progressive predictions  $\pi_n X$  are recursively defined by iterating this procedure: we put  $\pi_0 X = X$ , an  $\pi_{n-1} X = \mathcal{L}[\pi_n X | \mathcal{F}_{n-1}]$ ; thus, the  $n$ -th progressive prediction  $\pi_n X$  of  $X$  with respect to  $\mathcal{F}$  is a random variable taking its values in the Polish space  $E_n$ , which is recursively defined by  $E_0 = E$  and  $E_{n-1} = (E_n)'$ , where  $E'$  denotes the space of probability measures on  $E$ . We will sometimes denote the  $n$ -th progressive prediction  $\pi_n X$  by  $\pi_n^{\mathcal{F}} X$  when the filtration  $\mathcal{F}$  needs to be specified.

Below we list some elementary lemmas about the progressive predictions  $\pi_n X$ . In lemmas 3.1 to 3.8, we fix a Polish space  $E$ , a filtration  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  on a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ , and a random variable  $X \in L^0(\mathcal{F}_0; E)$ .

**Lemma 3.1.** *Let  $(\Omega', \mathcal{A}', \mathbb{P}')$  be a probability space and  $\Psi: \mathcal{F}_0 \rightarrow \mathcal{A}'$  be an embedding. Then  $\Psi(\pi_n^{\mathcal{F}} X) = \pi_n^{\Psi(\mathcal{F})} \Psi(X)$ .*

*Proof.* Straightforward from lemma 2.1. □

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

**Lemma 3.2.** *Let  $f: E \rightarrow \mathbb{R}$  be a measurable function such that  $f(X) \in L^1$ . Then for each  $n \leq 0$ , one has  $\mathbb{E}[f(X) | \mathcal{F}_n] = (\pi_n X)(f^n)$  where the functions  $f^n: E_n \rightarrow \mathbb{R}$  are recursively defined by  $f^0 = f$  and  $f^{n-1}(\mu) = \mu(f^n)$ .*

**Lemma 3.3.** *For each  $n \leq 0$ , the  $\sigma$ -field generated by the  $n$ -th progressive prediction  $\pi_n X$  contains the  $\sigma$ -field generated by the conditional law  $\mathcal{L}[X | \mathcal{F}_n]$ ; this inclusion is strict in general.*

*Proof.* Lemma 3.2 shows that  $\mathbb{E}[f(X) | \mathcal{F}_n]$  is  $\sigma(\pi_n X)$ -measurable for any suitable function  $f$ . It follows that  $\mathcal{L}[X | \mathcal{F}_n]$  is  $\sigma(\pi_n X)$ -measurable. An example of strict inclusion is provided in [ES01]. □

The difference between the  $\sigma$ -field generated by  $\pi_n X$  and the  $\sigma$ -field generated by  $\mathcal{L}[X | \mathcal{F}_n]$  is highlighted by noting that each random variable of the form  $\mathbb{E}[g(\mathbb{E}[f(X) | \mathcal{F}_{-1}]) | \mathcal{F}_{-2}]$  for suitable functions  $f$  and  $g$ , is measurable with respect to the  $\sigma$ -field generated by  $\pi_{-2} X$ , whereas it is not measurable with respect to the  $\sigma$ -field generated by  $\mathcal{L}[X | \mathcal{F}_{-2}]$  in general. Verification of the following lemma is left to the reader.

**Lemma 3.4.** *Let  $F$  be a Polish space and  $g: E \rightarrow F$  a measurable function. Then  $\pi_n g(X) = g^{(n)}(\pi_n X)$  where the functions  $g^{(n)}: E_n \rightarrow F_n$  are recursively defined by  $g^{(0)} = g$  and for each probability  $\mu \in E_n = (E_{n+1})'$ , the probability  $g^{(n)}(\mu) \in F_n = (F_{n+1})'$  is defined as the image of  $\mu$  under the mapping  $g^{(n+1)}$ .*

**Lemma 3.5.** *Let  $\mathcal{E}$  be a filtration immersed in  $\mathcal{F}$ . If  $X$  is measurable with respect to  $\mathcal{E}_0$ , then  $\pi_n^{\mathcal{F}} X = \pi_n^{\mathcal{E}} X$  for every  $n \leq 0$ .*

*Proof.* This is a straightforward consequence of the third characterization of immersion in lemma 2.5.  $\square$

The filtration generated by the process  $(\pi_n X)_{n \leq 0}$  will now be studied. Thanks to the following lemma, we first notice that this filtration depends only on the  $\sigma$ -field generated by  $X$ .

**Lemma 3.6.** *If two random variables  $X$  and  $Y$  with values in Polish spaces generate the same  $\sigma$ -field, the processes  $(\pi_n X)_{n \leq 0}$  and  $(\pi_n Y)_{n \leq 0}$  generate the same filtration.*

*Proof.* Obviously, it suffices to show that  $\mathcal{L}[X | \mathcal{F}_{-1}]$  and  $\mathcal{L}[Y | \mathcal{F}_{-1}]$  generate the same  $\sigma$ -field. Thanks to the Doob property of Polish spaces, one has  $X = h(Y)$  for some measurable function  $h$ . Therefore,  $\sigma(\mathcal{L}[X | \mathcal{F}_{-1}]) \subset \sigma(\mathcal{L}[Y | \mathcal{F}_{-1}])$  owing to the fact that  $\mathbb{E}[f(X) | \mathcal{F}_{-1}] = \mathbb{E}[f \circ h(Y) | \mathcal{F}_{-1}]$  for all suitable functions  $f$ . The reverse inclusion follows by symmetry.  $\square$

Lemma 3.7 below says that the filtration generated by the process  $(\pi_n X)_{n \leq 0}$  is the smallest filtration  $\mathcal{E} = (\mathcal{E}_n)_{n \leq 0}$  immersed in  $\mathcal{F}$  and such that  $X$  is  $\mathcal{E}_0$ -measurable.

**Lemma 3.7.** *The filtration generated by the process  $(\pi_n X)_{n \leq 0}$  is immersed in any filtration  $\mathcal{E} = (\mathcal{E}_n)_{n \leq 0}$  immersed in  $\mathcal{F}$  and such that  $X$  is  $\mathcal{E}_0$ -measurable. In particular it is immersed in  $\mathcal{F}$ .*

*Proof.* The filtration generated by the process  $(\pi_n X)_{n \leq 0}$  is immersed in  $\mathcal{F}$  as a straightforward consequence of lemma 2.8. It follows from lemma 3.5 that it is also immersed in  $\mathcal{E}$  if  $\mathcal{E}$  is immersed in  $\mathcal{F}$  and if  $X$  is  $\mathcal{E}_0$ -measurable.  $\square$

**Lemma 3.8.** *If  $X$  generates the  $\sigma$ -field  $\mathcal{F}_0$ , the process  $(\pi_n X)_{n \leq 0}$  generates the filtration  $\mathcal{F}$ .*

*Proof.* It suffices to show that  $\pi_{-1} X$  generates  $\mathcal{F}_{-1}$ . Any random variable  $S \in L^1(\mathcal{F}_0)$  can be written as  $S = f(X)$  where  $f$  is Borelian. Thus  $\mathbb{E}[S | \mathcal{F}_{-1}] =$

$(\pi_{-1}X)(f)$ . In particular  $S = (\pi_{-1}X)(f)$  if  $S$  is  $\mathcal{F}_{-1}$ -measurable, which shows that  $\pi_{-1}X$  generates  $\mathcal{F}_{-1}$ .  $\square$

**Lemma 3.9.** *The filtration  $\mathcal{F}$  generated by a martingale  $(M_n)_{n \leq 0}$  is also generated by the process  $(\pi_n^{\mathcal{F}}M_0)_{n \leq 0}$ .*

*Proof.* Obviously, the filtration generated by  $(\pi_n M_0)_{n \leq 0}$  is included in  $\mathcal{F}$ . The reverse inclusion holds as we have  $M_n = \mathbb{E}[M_0 | \mathcal{F}_n]$  and thus  $\sigma(M_n) \subset \sigma(\pi_n M_0)$  in view of lemma 3.3.  $\square$

### Iterated Kantorovich-Rubinstein metrics

Given a filtration  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  and an  $\mathcal{F}_0$ -measurable random variable  $X$  with values in a compact metric space  $K$ , we will define the  $n$ -th iterated Kantorovich-Rubinstein metric on the state space  $K_n$  of the  $n$ -th progressive prediction  $\pi_n X$  of  $X$ . The sequence of iterated Kantorovich-Rubinstein metrics has been introduced by Vershik to state his standardness criterion. Before turning to the Vershik property, we will give some applications of this sequence of metrics in connection with self-joining criteria (definition 2.16).

When  $(K, \rho)$  is a compact metric space, the weak topology on the set  $K'$  of probability measures on  $K$  is itself compact and is induced by the Kantorovich-Rubinstein metric  $\rho'$  on  $K'$  (see [Dud], [Ver04]). The Kantorovich-Rubinstein distance  $\rho'(\mu, \nu)$  between two probability measures  $\mu$  and  $\nu$  on  $K$  is defined as

$$\rho'(\mu, \nu) = \inf_{\Lambda} \iint \rho(x, y) d\Lambda(x, y),$$

where the infimum is taken over all probability measures  $\Lambda$  on  $K \times K$  whose first and second marginal measures are  $\mu$  and  $\nu$  respectively. Hence, we can define the Kantorovich-Rubinstein metric  $(\rho')'$  associated to  $\rho'$  which is in turn a metric on the space  $(K')'$  of probability measures on  $K'$ . And so on: the construction can be iterated, we end up with a sequence of compact metric spaces  $(K_n, \rho_n)_{n \leq 0}$  recursively defined by  $(K_0, \rho_0) = (K, \rho)$  and  $(K_{n-1}, \rho_{n-1}) = (K'_n, \rho'_n)$ .

Note that the Kantorovich-Rubinstein distance between two degenerate measures  $\mu = \delta_x$  and  $\nu = \delta_y$  is obviously given by

$$\rho'(\delta_x, \delta_y) = \rho(x, y). \tag{3.1}$$

In lemma 3.10, lemma 3.11, and proposition 3.12, we shall consider a filtration  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  and a compact metric space  $(K, \rho)$ . Recall that the  $n$ -th

progressive prediction  $\pi_n X$  of a random variable  $X \in L^0(\mathcal{F}_0; K) = L^1(\mathcal{F}_0; K)$  is a random variable taking its values in the compact metric space  $(K_n, \rho_n)$ . The following lemma easily follows from the definition of the iterated metrics  $\rho_n$ ; its proof is left to the reader.

**Lemma 3.10.** *Let  $X, Y \in L^1(\mathcal{F}_0; K)$ . Then the process  $(\rho_n(\pi_n X, \pi_n Y))_{n \leq 0}$  is a submartingale with respect to  $\mathcal{F}$ . In particular  $\mathbb{E}[\rho_n(\pi_n X, \pi_n Y)]$  decreases as  $n$  decreases.*

This lemma admits the following consequence. Recall that a joining of  $\mathcal{F}$  is a pair of jointly immersed filtrations isomorphic to  $\mathcal{F}$ .

**Lemma 3.11.** *Let  $X \in L^1(\mathcal{F}_0; K)$ . Let  $(\mathcal{F}', \mathcal{F}'')$  be a joining of  $\mathcal{F}$  on  $(\overline{\Omega}, \overline{\mathcal{A}}, \overline{\mathbb{P}})$ ; call  $X'$  and  $X''$  the respective copies of  $X$  by the embeddings defining the joining. Then the process  $(\rho_n(\pi_n X', \pi_n X''))_{n \leq 0}$  is a submartingale with respect to  $\mathcal{F}' \vee \mathcal{F}''$ . Consequently,  $\overline{\mathbb{E}}[\rho_n(\pi_n X', \pi_n X'')] decreases as  $n$  decreases.$*

In the context of this lemma, remark that, thanks to lemma 3.1,  $\pi_n X'$  denotes the  $n$ -th progressive prediction  $\pi_n^{\mathcal{F}'} X'$  of the copy  $X'$  of  $X$  as well as the copy of the  $n$ -th progressive prediction  $\pi_n^{\mathcal{F}} X$  of  $X$ .

*Proof of lemma 3.11.* As  $(\mathcal{F}', \mathcal{F}'')$  is a joining of  $\mathcal{F}$ , we know from lemma 3.5 that  $\pi_n^{\mathcal{F}'} X' = \pi_n^{\mathcal{F}' \vee \mathcal{F}''} X'$  and  $\pi_n^{\mathcal{F}''} X'' = \pi_n^{\mathcal{F}' \vee \mathcal{F}''} X''$ , so the result follows from lemma 3.10.  $\square$

**Proposition 3.12.** *Consider a self-joining criterion for  $\mathcal{F}$  (definition 2.16). If  $X \in L^1(\mathcal{F}_0; K)$  satisfies this self-joining criterion, then the  $\sigma$ -field generated by  $(\pi_n X)_{n \leq 0}$  satisfies this self-joining criterion. If the self-joining criterion is hereditary (definition 2.17), the filtration generated by  $(\pi_n X)_{n \leq 0}$  satisfies this self-joining criterion too.*

*Proof.* By lemma 2.20, this amounts to saying that the random variable  $R := (\dots, \pi_{-1} X, \pi_0 X)$  satisfies this self-joining criterion, where  $R$  is considered as a random variable taking its values in the Polish space  $\prod_{n \leq 0} E_n$  equipped with the distance  $d(e', e'') = \sum_{n \leq 0} 2^{n-1} \rho_n(e'_n, e''_n)$  for  $e' = (\dots, e'_{-1}, e'_0) \in \prod_{n \leq 0} E_n$  and  $e'' = (\dots, e''_{-1}, e''_0) \in \prod_{n \leq 0} E_n$ . This follows without difficulty from lemma 3.11. The last assertion is obvious from the definition of a hereditary self-joining criterion and from the fact that the filtration generated by  $(\pi_n X)_{n \leq 0}$  is immersed in  $\mathcal{F}$  (lemma 3.7).  $\square$

**Corollary 3.13.** *The filtration generated by a martingale  $(M_n)_{n \leq 0}$  satisfies a self-joining criterion if and only if the random variable  $M_0$  satisfies this self-joining criterion.*

*Proof.* This stems from proposition 3.12 and lemma 3.9. □

## 4 The Vershik property

The Vershik property is defined with the help of the notion of *dispersion* of a random variable defined as follows. Given a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$  and  $X \in L^1(\mathcal{A}; E)$  where  $(E, \rho)$  is a separable metric space, and calling  $\mu$  the law of  $X$ , the *dispersion* of  $X$  is the positive number

$$\text{disp } X = \iint \rho(x_1, x_2)(\mu \otimes \mu)(dx_1, dx_2) ,$$

which depends only upon the law  $\mu$ . Equivalently,

$$\text{disp } X = \overline{\mathbb{E}} [\rho(X', X'')] ,$$

where  $X'$  and  $X''$  are two independent copies of  $X$  defined on some probability space  $(\overline{\Omega}, \overline{\mathcal{A}}, \overline{\mathbb{P}})$ . Given two random variables  $X, Y \in L^1(\mathcal{A}; E)$ , it is easy to establish the inequality

$$|\text{disp } X - \text{disp } Y| \leq 2\mathbb{E} [\rho(X, Y)] , \quad (4.1)$$

thereby showing that  $\text{disp}$  is a continuous function on  $L^1(\mathcal{A}; E)$ .

Hereafter, given a filtration  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ , a compact metric space  $(K, \rho)$  and a random variable  $X \in L^1(\mathcal{F}_0; K)$ , it is understood that we consider the  $n$ -th iterated Kantorovich-Rubinstein distance  $\rho_n$  as the ambient metric on the compact space  $K_n$  in which the  $n$ -th progressive prediction  $\pi_n X$  of  $X$  takes its values.

**Definition 4.1.** Let  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  be a filtration on a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ .

1. Let  $K$  be a compact metric space and  $X \in L^1(\mathcal{F}_0; K)$ . We say that the random variable  $X$  is *Vershikian* (with respect to  $\mathcal{F}$ ) if  $\text{disp}(\pi_n X) \rightarrow 0$  as  $n$  goes to  $-\infty$ .
2. We say that a  $\sigma$ -field  $\mathcal{E}_0 \subset \mathcal{F}_0$  is *Vershikian* (with respect to  $\mathcal{F}$ ) if each random variable  $X \in L^1(\mathcal{E}_0; [0, 1])$  is Vershikian with respect to  $\mathcal{F}$ .
3. We say that the filtration  $\mathcal{F}$  is *Vershikian* if the  $\sigma$ -field  $\mathcal{F}_0$  is Vershikian with respect to  $\mathcal{F}$ .



When there is no ambiguity, we will omit the specification *with respect to*  $\mathcal{F}$  in this definition. In view of lemma 3.1, it is clear that the Vershik property is preserved by isomorphism. One of our main goals is to prove the converse of the following proposition.

**Proposition 4.2.** *Let  $\mathcal{F}$  be a filtration and  $K$  a compact metric space. Every I-cosy random variable  $X \in L^1(\mathcal{F}_0; K)$  is Vershikian. Consequently, every I-cosy filtration is Vershikian.*

*Proof.* This easily follows from the definition of I-cosiness, from lemma 3.11, and, with the notations of this lemma, from the fact that  $\overline{\mathbb{E}}[\rho_n(\pi_n X', \pi_n X'')] = \text{disp } \pi_n X$  if  $\pi_n X'$  and  $\pi_n X''$  are two independent copies of  $\pi_n X$ . The consequence stems from lemma 2.21.  $\square$

The proof of the converse of proposition 4.2 is the object of section 8. The rest of the current section is devoted to elementary properties of the Vershik property.

**Lemma 4.3.** *Let  $\mathcal{F}$  be a filtration,  $\mathcal{E}$  a filtration immersed in  $\mathcal{F}$ , and  $K$  a compact metric space. A random variable  $X \in L^1(\mathcal{E}_0; K)$  is Vershikian with respect to  $\mathcal{F}$  if and only if it is Vershikian with respect to  $\mathcal{E}$ . Consequently, if the filtration  $\mathcal{F}$  is Vershikian, then so does also  $\mathcal{E}$ .*

*Proof.* This follows from  $\pi_n^{\mathcal{F}} X = \pi_n^{\mathcal{E}} X$  (lemma 3.5).  $\square$

Whereas each point of this lemma is obvious, on the contrary the similar statement with I-cosiness is not obvious. However we shall see that it holds for locally separable filtrations as a consequence of the converse of proposition 4.2; but, to our knowledge, this result cannot be directly proved from the I-cosiness criterion.

We are now going to show that a random variable  $X$  is Vershikian if and only if the  $\sigma$ -field  $\sigma(X)$  is Vershikian (proposition 4.7). We shall see a stronger result in corollary 6.5, which asserts that  $X$  is Vershikian with respect to a filtration  $\mathcal{F}$  if and only if the filtration  $\mathcal{F}^X$  generated by  $(\pi_n X)_{n \leq 0}$  is Vershikian; in view of lemma 4.3, this amounts to say that the  $\sigma$ -field  $\sigma(\pi_n X; n \leq 0)$  is Vershikian with respect to  $\mathcal{F}$ , since  $\mathcal{F}^X$  is immersed in  $\mathcal{F}$  (lemma 3.7).

**Lemma 4.4.** *Let  $\mathcal{F}$  be a filtration and  $K$  be a compact metric space. The set of Vershikian random variables  $X \in L^1(\mathcal{F}_0; K)$  is a closed subset of  $L^1(\mathcal{F}_0; K)$ .*

*Proof.* This results from inequality (4.1) and from lemma 3.10.  $\square$

**Lemma 4.5.** *Let  $(\Omega, \mathcal{A}, \mathbb{P})$  be a probability space and  $X \in L^1(\mathcal{A}; E)$  where  $(E, \rho)$  is a separable metric space. The set of all random variables of the form  $f(X)$  where  $f: E \rightarrow \mathbb{R}$  is Lipschitz function, is a dense subset of  $L^1(\sigma(X))$ .*

*Proof.* See lemma 2.15 in [Lau10].  $\square$

Verification of the following lemma is left to the reader. It easily results from the definition of Kantorovich-Rubinstein metrics.

**Lemma 4.6.** *Let  $E$  and  $F$  be two compact metric spaces and  $g: E \rightarrow F$  be  $c$ -Lipschitz. Let  $g^{(n)}: E_n \rightarrow F_n$  be the functions such that  $\pi_n g(X) = g^{(n)}(\pi_n X)$  defined in lemma 3.4. For each  $n \leq 0$ , the function  $g^{(n)}$  is  $c$ -Lipschitz for the iterated Kantorovich-Rubinstein metrics on  $E_n$  and  $F_n$ .*

**Proposition 4.7.** *Let  $\mathcal{F}$  be a filtration and  $K$  a compact metric space. A random variable  $X \in L^1(\mathcal{F}_0; K)$  is Vershikian if and only if the  $\sigma$ -field  $\sigma(X)$  is Vershikian. Consequently, for the filtration  $\mathcal{F}$  to be Vershikian, it suffices that there exists a Vershikian random variable  $X$  that generates  $\mathcal{F}_0$ .*

*Proof.* By lemmas 4.4 and 4.5, it suffices to show that the Vershik property is satisfied for every random variable of the form  $g(X)$  where  $g: K \rightarrow \mathbb{R}$  is  $c$ -Lipschitz. To do so, it suffices to notice that, thanks to lemma 4.6, one has  $\text{disp } \pi_n g(X) \leq c \text{disp } \pi_n X$  for every  $n \leq 0$ .  $\square$

**Corollary 4.8.** *Let  $\mathcal{F}$  be a filtration,  $\mathcal{E}_0 \subset \mathcal{F}_0$  a  $\sigma$ -field and  $K$  a compact metric space. If  $\mathcal{E}_0$  is Vershikian, every random variable  $X \in L^1(\mathcal{E}_0; K)$  is Vershikian.*

*Proof.* Assume that  $\mathcal{E}_0$  is Vershikian. It is clear that so is also any sub- $\sigma$ -field of  $\mathcal{E}_0$ . Hence, if  $X \in L^1(\mathcal{E}_0; K)$ , the  $\sigma$ -field  $\sigma(X)$  is Vershikian, and  $X$  is Vershikian by proposition 4.7.  $\square$

The following results will be used in section 6. Proposition 4.11 says that the Vershik property is asymptotic; it shall be proved with the help of the two following lemmas, in the proofs and the statements of which we use the following notations and relations. Given a filtration  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  and an integer  $N \in -\mathbb{N}$ , we denote by  $\mathcal{F}^{[N]}$  the truncated filtration  $(\mathcal{F}_{N+n})_{n \leq 0}$ . We will use the trivial relation

$$\pi_{N+n}^{\mathcal{F}} X = \pi_n^{\mathcal{F}^{[N]}} \pi_N^{\mathcal{F}} X, \quad (4.2)$$

for every  $\mathcal{F}_0$ -measurable random variable  $X$  taking its values in a compact metric space, wherefrom we see that  $X$  is Vershikian with respect to  $\mathcal{F}$  if and only if  $\pi_N^{\mathcal{F}}X$  is Vershikian with respect to  $\mathcal{F}^N$ . In addition, we will use the canonical isometries

$$i_n: (K, \rho) \rightarrow (K_n, \rho_n), \quad n \in -\mathbb{N},$$

recursively defined by  $i_0(x) = x$  and  $i_{n-1}(x) = \delta_{i_n(x)}$  for every compact metric space  $K$ . Each  $i_n$  is an isometry owing to equality (3.1). One obviously has  $\pi_n X = i_n(X)$  whenever  $X$  is  $\mathcal{F}_n$ -measurable. In lemma 4.9 and lemma 4.10, we consider a filtration  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  and a random variable  $X \in L^1(\mathcal{F}_0; K)$  for some compact metric space  $K$ .

**Lemma 4.9.** *Let  $N \in -\mathbb{N}$ . If  $X$  is measurable with respect to  $\mathcal{F}_N$ , then the Vershik property of  $X$  holds with respect to  $\mathcal{F}$  if and only if it holds with respect to  $\mathcal{F}^N$ .*

*Proof.* Let  $N \in -\mathbb{N}$ . Thanks to (4.2), the Vershik property for  $X$  with respect to  $\mathcal{F}$  is equivalent to the Vershik property of  $\pi_N^{\mathcal{F}}X$  with respect to  $\mathcal{F}^N$ . But  $\pi_N^{\mathcal{F}}X = i_N(X)$  in the case when  $X$  is measurable with respect to  $\mathcal{F}_N$ , where  $i_N: K \rightarrow K_N$  is the canonical isometry. Thus the Vershik property of  $X$  with respect to  $\mathcal{F}$  is finally equivalent to the Vershik property of  $X$  with respect to  $\mathcal{F}^N$ .  $\square$

**Lemma 4.10.** *The following propositions are equivalent:*

- (i)  $X$  is Vershikian;
- (ii) for every  $n \in -\mathbb{N}$ ,  $\pi_n X$  is Vershikian;
- (iii) there exists  $n \in -\mathbb{N}$  such that  $\pi_n X$  is Vershikian.

*Proof.* Let  $N \in -\mathbb{N}$ . Thanks to (4.2), the Vershik property of  $X$  with respect to  $\mathcal{F}$  is equivalent to the Vershik property of  $\pi_N^{\mathcal{F}}X$  with respect to  $\mathcal{F}^N$ . Thanks to lemma 4.9, this is equivalent to the Vershik property of  $\pi_N^{\mathcal{F}}X$  with respect to  $\mathcal{F}$ .  $\square$

**Proposition 4.11.** *Let  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  be a filtration. The following conditions are equivalent:*

- (i)  $\mathcal{F}$  is Vershikian;

- (ii) for every  $N \in -\mathbb{N}$ , the truncated filtration  $(\mathcal{F}_{N+n})_{n \leq 0}$  is Vershikian;
- (iii) there exists  $N \in -\mathbb{N}$  such that the truncated filtration  $(\mathcal{F}_{N+n})_{n \leq 0}$  is Vershikian.

Note that no assumption is made on the filtration  $\mathcal{F}$  of this proposition, whereas the analogous proposition for I-cosiness given in [Lau10], needs local separability.

*Proof of proposition 4.11.* (ii)  $\implies$  (iii) is obvious. (i)  $\implies$  (ii) follows from lemma 4.9. Now we prove that (iii)  $\implies$  (i). Let  $X \in L^1(\mathcal{F}_0; [0, 1])$ . If  $\mathcal{F}^N$  is Vershikian, then, thanks to corollary 4.8, the random variable  $\pi_N^{\mathcal{F}} X$  is Vershikian with respect to  $\mathcal{F}^N$ . Owing to lemma 4.9, this amounts to say that  $\pi_N^{\mathcal{F}} X$  is Vershikian with respect to  $\mathcal{F}$ , and, owing to lemma 4.10, this amounts to say that  $X$  is Vershikian with respect to  $\mathcal{F}$ .  $\square$

The rest of this section is devoted to proving the following remark.

**Remark 4.12.** If  $X$  and  $Y$  are two Vershikian random variables, it is not true in general that the pair  $(X, Y)$  is a Vershikian random variable.

This remark will be proved by using three elementary facts. The first two ones are the contents of the following lemmas.

**Lemma 4.13.** *Any Vershikian filtration is Kolmogorovian.*

*Proof.* Let  $S \in L^1(\mathcal{F}_{-\infty}; [0, 1])$ . One has  $\pi_n S = i_n(S)$  for every  $n \leq 0$  where the  $i_n$  are the canonical isometries introduced above lemma 4.9. Hence one has  $\text{disp}(\pi_n S) = \text{disp}(S)$  for every  $n \leq 0$ . Consequently one has  $\text{disp}(S) = 0$  under the assumption that  $\mathcal{F}$  is Vershikian, which means that  $S$  is constant.  $\square$

As in [Lau10], we say that a filtration is *of product type* if it is generated by a sequence of independent random variables.

**Lemma 4.14.** *Any filtration of product type is Vershikian.*

*Proof.* On  $(\Omega, \mathcal{A}, \mathbb{P})$ , let  $(V_n)_{n \leq 0}$  be a sequence of independent random variables and  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  the filtration it generates. In order for  $\mathcal{F}$  to be Vershikian, it suffices, in view of lemma 4.4, that the  $\sigma$ -fields  $\sigma(V_{n_0+1}, \dots, V_0)$  are Vershikian for all integers  $n_0 < 0$ . For a given integer  $n_0 < 0$ , we introduce the filtration  $\mathcal{E} = (\mathcal{E}_n)_{n \leq 0}$  generated by  $(V_{n_0+1}, \dots, V_0)$  on the time-axis

$-\mathbb{N}$ , defined by

$$\mathcal{E}_n = \begin{cases} \{\emptyset, \Omega\} & \text{if } n \leq n_0, \\ \sigma(V_{n_0+1}, \dots, V_n) & \text{if } n_0 < n \leq 0. \end{cases}$$

By lemma 2.8,  $\mathcal{E}$  is immersed in  $\mathcal{F}$ . Therefore, for each random variable  $X \in L^1(\mathcal{E}_0; [0, 1])$ , one has  $\pi_n^{\mathcal{F}}X = \pi_n^{\mathcal{E}}X$  for every  $n \leq 0$  (lemma 3.5). Obviously,  $\pi_n^{\mathcal{E}}X$  is a degenerate (deterministic) random variable for  $n \leq n_0$ , hence we finally have  $\text{disp}(\pi_n X) = 0$  for  $n \leq n_0$ .  $\square$

The third elementary fact we use to prove remark 4.12 is theorem 3.9 of [Lau10], derived from [Par89], asserting in particular that *any essentially separable filtration is immersible in the supremum of two jointly immersed filtrations of product type*.

*Proof of remark 4.12.* Suppose that  $\mathcal{F}$  is an essentially separable filtration which is not Kolmogorovian. Theorem 3.9 of [Lau10] shows that  $\mathcal{F}$  is immersible in a filtration of the form  $\mathcal{G} \vee \mathcal{H}$  where  $\mathcal{G}$  and  $\mathcal{H}$  are jointly immersed filtrations of product type. In particular  $\mathcal{G} \vee \mathcal{H}$  is not Kolmogorovian either, and hence not Vershikian (lemma 4.13). Let  $X \in L^1(\mathcal{G}_0; [0, 1])$  and  $Y \in L^1(\mathcal{H}_0; [0, 1])$  be two random variables respectively generating  $\mathcal{G}_0$  and  $\mathcal{H}_0$ . By lemma 4.14,  $X$  and  $Y$  are Vershikian with respect to  $\mathcal{G}$  and  $\mathcal{H}$  respectively. As  $\mathcal{G}$  and  $\mathcal{H}$  are jointly immersed,  $X$  and  $Y$  are also Vershikian with respect to  $\mathcal{G} \vee \mathcal{H}$  due to lemma 4.3. However the pair  $(X, Y)$  cannot be Vershikian, because it generates the  $\sigma$ -field  $\mathcal{G}_0 \vee \mathcal{H}_0$  and thus the filtration  $\mathcal{G} \vee \mathcal{H}$  would be Vershikian by proposition 4.7.  $\square$

Before turning to the next section, we make a remark (which can be skipped) on the preceding proof. Observe that, with the notations of the above proof, the filtrations  $\mathcal{G}$  and  $\mathcal{H}$  are respectively generated by the processes  $(\pi_n X)_{n \leq 0}$  and  $(\pi_n Y)_{n \leq 0}$  in view of lemma 3.8. In a general context, given a filtration  $\mathcal{F}$  and two  $\mathcal{F}_0$ -measurable random variables  $X$  and  $Y$ , the filtrations  $\mathcal{E}^X$  and  $\mathcal{E}^Y$  generated by  $(\pi_n X)_{n \leq 0}$  and  $(\pi_n Y)_{n \leq 0}$  are always jointly immersed (lemma 3.5). We shall see in corollary 6.5 that the Vershik property for a random variable is equivalent to the Vershik property for the filtration generated by the progressive predictions of this random variable. However, contrary to the case in the above proof, it is not clear that in general, the Vershik property for  $(X, Y)$  is related to the filtration  $\mathcal{E}^X \vee \mathcal{E}^Y$ , which is in general

strictly smaller than the filtration  $\mathcal{E}^{(X,Y)}$  generated by  $(\pi_n(X, Y))_{n \leq 0}$ . To construct an example of such a strict inclusion, consider a  $\sigma$ -field  $\mathcal{F}_{-1}$  and two random variables  $X$  and  $Y$  such that the joint conditional law  $\mathcal{L}[X, Y | \mathcal{F}_{-1}]$  generates a  $\sigma$ -field strictly bigger than the one generated by the marginal conditional laws  $\mathcal{L}[X | \mathcal{F}_{-1}]$  and  $\mathcal{L}[Y | \mathcal{F}_{-1}]$ , define  $\mathcal{F}_0 = \mathcal{F}_{-1} \vee \sigma(X, Y)$  and set  $\mathcal{F}_n$  to be the degenerate  $\sigma$ -field for every  $n \leq -2$ .

## 5 Vershik's theorem on lacunary isomorphism

*Vershik's theorem on lacunary isomorphism* is the content of theorem 5.2 below. Although all main ideas are already present in Vershik's original proof ([Ver68, Ver94]), the one we give is more direct and less technical. (With the terminologies of [ES01], which translates Vershik's proof into probabilistic language, or of [Lau10], the technical step we do not need is the extraction of a filtration of product type from any essentially separable, Kolmogorovian, and conditionally non-atomic filtration.)

**Theorem 5.1.** *Let  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  be a filtration and  $X \in L^1(\mathcal{F}_0; K)$  where  $K$  is a compact metric space. If the filtration generated by  $(\pi_n X)_{n \leq 0}$  is Kolmogorovian, there exists a strictly increasing map  $\phi : -\mathbb{N} \rightarrow -\mathbb{N}$  such that  $X$  is Vershikian with respect to the extracted filtration  $(\mathcal{F}_{\phi(n)})_{n \leq 0}$ .*

*Proof.* Denote by  $\mathcal{E} = (\mathcal{E}_n)_{n \leq 0}$  the filtration generated by  $(\pi_n X)_{n \leq 0}$ . As  $\mathcal{E}$  is immersed in  $\mathcal{F}$ , one has  $\mathcal{L}[Y | \mathcal{F}_n] = \mathcal{L}[Y | \mathcal{E}_n]$  for any  $\mathcal{E}_0$ -measurable random variable  $Y$  valued in a compact metric space. So, under the assumption that  $\mathcal{E}$  is Kolmogorovian, by Lévy's theorem on reverse martingale convergence (see [Dud] or [Kall]), the conditional law  $\mathcal{L}[Y | \mathcal{F}_n]$  tends almost surely to a constant and  $\text{disp } \mathcal{L}[Y | \mathcal{F}_n]$  tends to 0 when  $n \rightarrow -\infty$ . Therefore, for any sequence  $(\delta_n)_{n \leq 0}$  of positive numbers, it is possible to construct a strictly increasing map  $\phi : -\mathbb{N} \rightarrow -\mathbb{N}$  such that  $\text{disp } \gamma_n \leq \delta_n$  for every  $n \leq 0$  where  $\gamma_n$  is recursively defined by  $\gamma_0 = X$  and  $\gamma_{n-1} = \mathcal{L}[\gamma_n | \mathcal{F}_{\phi(n-1)}]$ . Thus  $\gamma_n$  is the  $n$ -th progressive prediction  $\pi_n^{\mathcal{G}} X$  of  $X$  with respect to the filtration  $\mathcal{G} = (\mathcal{G}_n)_{n \leq 0}$  defined by  $\mathcal{G}_n = \mathcal{F}_{\phi(n)}$ . By letting  $\delta_n \rightarrow 0$ , we see that  $X$  is Vershikian with respect to  $\mathcal{G}$ .  $\square$

**Theorem 5.2.** *Let  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  be an essentially separable filtration. If  $\mathcal{F}$  is Kolmogorovian, there exists a strictly increasing map  $\phi : -\mathbb{N} \rightarrow -\mathbb{N}$  such that the extracted filtration  $(\mathcal{F}_{\phi(n)})_{n \leq 0}$  is Vershikian.*

*Proof.* With the help of lemma 3.8 and proposition 4.7, this is deduced from theorem 5.1 applied to a random variable  $X$  generating  $\mathcal{F}_0$ .  $\square$

**Remark 5.3.** There is no difficulty in checking that I-cosiness, as well as *standardness* and *weak standardness* (see [Lau10]), of a filtration  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  is inherited to each extracted filtration  $(\mathcal{F}_{\phi(n)})_{n \leq 0}$ . However this fact is not so obvious for the Vershik property.

## 6 Parametric extensions

The main results of this section involve *parametric extensions* (definition 2.12) in their statements or in their proof. We shall see (proposition 6.1) that the Vershik property holds for any parametric extension of a filtration as soon as it holds for this filtration (this justifies remark 3.26 in [Lau10]), and we shall see (corollary 6.5) that the Vershik property of a random variable  $X$  is equivalent to the Vershik property of the filtration  $(\pi_n X)_{n \leq 0}$ .

This section illustrates the benefits of the Vershik property as compared to the equivalent I-cosiness property: whereas the results we give are rather easily proved, we do not know such elementary proofs of the analogous results for I-cosiness, derived from the equivalence between the Vershik property and I-cosiness (assertion (1.1) and theorem 8.2).

**Proposition 6.1.** *Let  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  be a filtration and  $\mathbf{V} = (V_n)_{n \leq 0}$  a global superinnovation of  $\mathcal{F}$ ; call  $\mathcal{G}$  the extension of  $\mathcal{F}$  with  $\mathbf{V}$  (definition 2.12). Then  $\mathcal{F}$  is Vershikian if and only if  $\mathcal{G}$  is Vershikian.*

*Proof.* We know that  $\mathcal{F}$  is immersed in  $\mathcal{G}$  (lemma 2.11), so the ‘if’ part follows from the Vershik property being hereditary (lemma 4.3). To show the converse, assume that  $\mathcal{F}$  is Vershikian. For each integer  $k < 0$  call  $\mathcal{G}^k$  the extension of  $\mathcal{F}$  with the local innovation  $\mathbf{V}^k := (V_n)_{n \in ]k, 0] \cap \mathbb{Z}}$ . By proposition 4.11, each filtration  $\mathcal{G}^k$  is Vershikian. Thanks to lemma 2.8, we can easily see that all these filtrations are immersed in  $\mathcal{G}$ . Consequently, by lemma 4.3, every random variable  $X \in \cup_k L^1(\mathcal{G}_0^k; [0, 1])$  is Vershikian with respect to  $\mathcal{G}$ . Finally,  $\mathcal{G}$  is Vershikian by lemma 4.4.  $\square$

**Lemma 6.2.** *Let  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  be a filtration and  $\mathbf{V} = (V_n)_{n \leq 0}$  a global superinnovation of  $\mathcal{F}$ ; call  $\mathcal{G}$  the extension of  $\mathcal{F}$  with  $\mathbf{V}$ . Let  $N \leq 0$  be an integer and  $\mathcal{E}_N \subset \mathcal{F}_N$  a  $\sigma$ -field. The  $\mathcal{E}_N$  is Vershikian if and only if the  $\sigma$ -field  $\mathcal{E}_N \vee \sigma(V_n; N < n \leq 0)$  is Vershikian with respect to  $\mathcal{G}$ .*

We do not specify with respect to which filtration  $\mathcal{E}_N$  is Vershikian because we implicitly use lemma 4.3 asserting that the Vershik property of  $\mathcal{E}_N$  is the same when considered either with respect to  $\mathcal{F}$  or with respect to  $\mathcal{G}$ .

*Proof of lemma 6.2.* The ‘if’ part obviously follows from the Vershik property being inherited by immersion (lemma 4.3). We now prove the converse. Assume  $\mathcal{E}_N$  is Vershikian and let  $Y \in L^1(\mathcal{E}_N \vee \sigma(V_n; N < n \leq 0); [0, 1])$ . It is easy to convince oneself that  $\pi_N^{\mathcal{G}} Y$  is measurable with respect to  $\mathcal{E}_N$ . Therefore, owing to corollary 4.8,  $\pi_N^{\mathcal{G}} Y$  is Vershikian with respect to  $\mathcal{G}$ , which amounts to say that  $Y$  is Vershikian with respect to  $\mathcal{G}$  by virtue of lemma 4.10.  $\square$

**Proposition 6.3.** *Let  $(X_n)_{n \leq 0}$  be a process and  $\phi: -\mathbb{N} \rightarrow -\mathbb{N}$  a sequence of memory-loss times of type II (definition 2.14) for  $(X_n)_{n \leq 0}$ . Then the filtration  $\mathcal{F}$  generated by  $(X_n)_{n \leq 0}$  is Vershikian if and only if each  $\sigma$ -field  $\sigma(X_{\phi(n)})$  is Vershikian.*

Note that, in the case when  $(X_n)_{n \leq 0}$  is a Markov process, every strictly increasing map  $\phi: -\mathbb{N} \rightarrow -\mathbb{N}$  with  $\phi(0) = 0$  is a sequence of memory-loss times of type I for  $(X_n)_{n \leq 0}$  (definition 2.14), and hence of type II in view of lemma 2.15.

*Proof of proposition 6.3.* The ‘only if’ part is obvious. To show the converse, we take, up to isomorphism, a parameterization  $\mathbf{U} = (U_n)_{n \leq 0}$  of  $\mathcal{F}$  as in definition 2.14.2, and we introduce the extension  $\mathcal{G}$  of  $\mathcal{F}$  with  $\mathbf{U}$ . Since the Vershik property is inherited by immersion (lemma 4.3), it suffices to show that  $\mathcal{G}$  is Vershikian assuming that each  $\sigma$ -field  $\sigma(X_{\phi(n)})$  is Vershikian. But we have  $\sigma(X_{\phi(N)}) \vee \sigma(U_n; \phi(N) < n \leq 0) \nearrow \mathcal{G}_0$  as  $N \rightarrow -\infty$ , hence, in order for  $\mathcal{G}$  to be Vershikian, owing to lemma 4.4, it is sufficient that each  $\sigma$ -field  $\sigma(X_{\phi(N)}) \vee \sigma(U_n; \phi(N) < n \leq 0)$  is Vershikian. This stems from lemma 6.2.  $\square$

Below is a corollary of proposition 6.1 and proposition 6.3.

**Corollary 6.4.** *Let  $(X_n)_{n \leq 0}$  be a process and  $\phi: -\mathbb{N} \rightarrow -\mathbb{N}$  a sequence of memory-loss times of type II for  $(X_n)_{n \leq 0}$ . Let  $\mathcal{F}$  be the filtration generated by  $(X_n)_{n \leq 0}$ . Then the following assertions are equivalent:*

- (i) *the extracted filtration  $(\mathcal{F}_{\phi(n)})_{n \leq 0}$  is Vershikian;*
- (ii) *the filtration generated by the extracted process  $(X_{\phi(n)})_{n \leq 0}$  is Vershikian;*



(iii) each  $\sigma$ -field  $\sigma(X_{\phi(n)})$  is Vershikian with respect to the filtration  $(\mathcal{F}_{\phi(n)})_{n \leq 0}$ .

*Proof.* We firstly show that (i)  $\iff$  (ii). Under the assumption that  $\phi$  is a sequence of memory-loss times of type II, it is clear that the process  $(X_{\phi(n)})_{n \leq 0}$  is Markovian with respect to  $(\mathcal{F}_{\phi(n)})_{n \leq 0}$ , hence the filtration  $\mathcal{E} = (\mathcal{E}_n)_{n \leq 0}$  generated by  $(X_{\phi(n)})_{n \leq 0}$  is immersed in  $(\mathcal{F}_{\phi(n)})_{n \leq 0}$ . Consequently, (i)  $\implies$  (ii) follows from the Vershik property being inherited by immersion (lemma 4.3). Now, let  $\mathbf{U} = (U_n)_{n \leq 0}$  be a parameterization of  $\mathcal{F}$  as in definition 2.14.2 up to isomorphism; define  $V_n = (U_{\phi(n-1)+1}, \dots, U_{\phi(n)})$  for each  $n \leq 0$ . Then  $\mathbf{V} := (V_n)_{n \leq 0}$  is a superinnovation of the filtration  $\mathcal{E}$  as well as a superinnovation of the filtration  $(\mathcal{F}_{\phi(n)})_{n \leq 0}$ , and these two filtrations have the same extension  $\mathcal{G}$  with  $\mathbf{V}$ . Hence, proposition 6.1 shows that (ii) implies that  $\mathcal{G}$  is Vershikian, and hence (ii)  $\implies$  (i) follows from the Vershik property being inherited by immersion (lemma 4.3).

Since the Vershik property of  $\sigma(X_{\phi(n)})$  is the same when considered either with respect to  $(\mathcal{F}_{\phi(n)})_{n \leq 0}$  or with respect to  $\mathcal{E}$  (lemma 4.3), then (ii)  $\iff$  (iii) stems from proposition 6.3, because the identity map of  $-\mathbb{N}$  is a sequence of memory-loss times of type II for the Markov process  $(X_{\phi(n)})_{n \leq 0}$ .  $\square$

Corollary 6.4 obviously does not remain true without the memory-loss assumption. Indeed, consider a non-constant random variable  $Z$  and define the process  $(X_n)_{n \leq 0}$  by  $X_{2n-1} = Z$  and  $X_{2n} = 0$  for every  $n \leq 0$ . Then the extracted filtration  $(\mathcal{F}_{2n})_{n \leq 0}$  is not Kolmogorovian and therefore does not satisfy the Vershik property in view of lemma 4.13, whereas the trivial filtration generated by  $(X_{2n})_{n \leq 0}$  obviously satisfies the Vershik property.

Below is a corollary of proposition 6.3.

**Corollary 6.5.** *A random variable  $X$  is Vershikian with respect to a filtration  $\mathcal{F}$  if and only if the filtration  $\mathcal{F}^X$  generated by  $(\pi_n X)_{n \leq 0}$  is Vershikian.*

*Proof.* By lemma 4.10 and proposition 4.7,  $X$  is Vershikian if and only if all the  $\sigma$ -fields  $\sigma(\pi_n X)$  are Vershikian. We know that  $\mathcal{F}^X$  is immersed in  $\mathcal{F}$  (lemma 3.5), hence the Vershik property of  $\sigma(\pi_n X)$  is the same when considered either with respect to  $\mathcal{F}$  or with respect to  $\mathcal{F}^X$  (lemma 4.3). Since the process  $(\pi_n X)_{n \leq 0}$  obviously is Markovian, the corollary results from proposition 6.3.  $\square$

**Example 6.6** (*A fifth definition of the scale of an automorphism*). The notion of *scale of an automorphism* has been introduced by Vershik in [Ver73b].

We have checked that the first definition of the scale given by Vershik can equivalently be rephrased in terms of the Vershik property of the  $\sigma$ -field  $\sigma(Z_0)$  with respect to the filtration  $\mathcal{F}$  generated by the process  $(Z_n, \eta_n)_{n \leq 0}$  to be defined below.

Let  $T$  be an automorphism of a Lebesgue probability space  $(E, \mathfrak{B}, \nu)$ . Let  $(r_n)_{n \leq 0}$  be a sequence consisting of integers  $r_n \geq 2$ , called the *splitting sequence*, and define the sequence  $(\ell_n)_{n \leq 0}$  by  $\ell_n = \prod_{k=n+1}^0 r_k$ . Let  $\xi_0$  be a random variable whose law is  $\nu$  and define the random variable  $\xi_k$  for each  $k \geq 0$  by  $\xi_k = T^k(\xi_0)$ .

Define the Markov process  $(Z_n, \eta_n)_{n \leq 0}$  whose law is characterized by the following two conditions:

- ◇ for each  $n \leq 0$ ,  $Z_n$  is distributed on  $E^{\ell_n}$  according to the law of  $(\xi_0, \dots, \xi_{\ell_n})$ , and  $\eta_n$  is independent of  $Z_n$  and has the uniform law on the set  $\{1, \dots, r_n\}$ ;
- ◇  $\eta_n$  is independent of  $(Z_{n-1}, \eta_{n-1})$ , and  $Z_n = Z_{n-1}(\eta_n)$  where we write  $Z_{n-1} = (Z_{n-1}(1), \dots, Z_{n-1}(r_n))$  as an element of  $(E^{\ell_n})^{r_n}$ .

The *scale* of  $T$  is a set of such sequences of integers  $(r_n)_{n \leq 0}$ . We have checked that:

- the first definition given by Vershik is equivalent to say that  $(r_n)_{n \leq 0}$  *belongs to the scale of  $T$  if and only if the  $\sigma$ -field  $\sigma(Z_0)$  is Vershikian<sup>2</sup> with respect to  $\mathcal{F}$ .*
- *under the assumption that  $T$  is completely ergodic (i.e. all powers of  $T$  are ergodic), the filtration  $\mathcal{F}$  generated by  $(Z_n, \eta_n)_{n \leq 0}$  is also generated by the process of Vershik's progressive predictions  $(\pi_n Z_0)_{n \leq 0}$  of  $Z_0$ .*

Hence, by virtue of corollary 6.5,  $(r_n)_{n \leq 0}$  *belongs to the scale of a completely ergodic automorphism  $T$  if and only if  $\mathcal{F}$  is Vershikian.*

Of particular interest is the case of Bernoulli automorphisms. Let  $(A, \mathfrak{A}, \mu)$  be a Lebesgue probability space and  $T$  be the shift on  $A^{\mathbb{Z}}$  equipped with the product measure of  $\mu$ . Let  $c: A^{\mathbb{Z}} \rightarrow A$  be the map that takes the 0-th component. Letting  $W_n = c(Z_n)$ , it is easy to see that  $\mathcal{F}$  is generated by the process

---

<sup>2</sup> Actually we have checked, with the terminology of [Lau10], that the first definition is equivalent to say that  $\sigma(Z_0)$  satisfies *Vershik's self-joining criterion* (or the “combinatorial Vershik's standardness criterion” for  $\{r_n\}$ -adic filtrations; see [VG07]). Since  $\mathcal{F}$  is *homogeneous*, this is equivalent to the Vershik property of  $\sigma(Z_0)$  by the results in [Lau10] and because of the equivalence between the Vershik property and I-cosiness (assertion (1.1) and theorem 8.2).

$(W_n, \eta_n)_{n \leq 0}$ . With the terminology of [Lau10], this process is the *split-word process* on the alphabet  $(A, \mathfrak{A}, \mu)$ , and  $\mathcal{F}$  is an *homogeneous* filtration, hence we know from the pioneering works of Vershik that the Vershik property of  $\mathcal{F}$  is equivalent to productness of  $\mathcal{F}$  (this can be deduced from corollary 4.5 in [Lau10] and from the equivalence between the Vershik property and I-cosiness). Therefore, theorem 2.39 in [Lau10] (productness of the split-word filtrations) shows that *the scale of Bernoulli automorphisms* consists of those sequences  $(r_n)_{n \leq 0}$  for which the condition

$$\sum_{k=-\infty}^0 \frac{\log(r_k)}{\ell_k} < \infty \quad (\Delta)$$

is not satisfied. It was already asserted in [Ver73b] that  $\neg(\Delta)$  is a necessary condition for  $(r_n)_{n \leq 0}$  to belong to the scale of Bernoulli automorphisms. The converse stated in theorem 2.39 in [Lau10] is mainly derived from Ceillier's result in [Ceil09] (the particular case when  $A$  is finite and  $\mu$  is the uniform probability measure).

**Corollary 6.7.** *The filtration generated by a martingale  $(M_n)_{n \leq 0}$  is Vershikian if and only if the  $\sigma$ -field  $\sigma(M_0)$  is Vershikian.*

*Proof.* The  $\sigma$ -field  $\sigma(M_0)$  is always generated by a bounded random variable. Thus, with the help of proposition 4.7, the result stems from the previous corollary and lemma 3.9.  $\square$

## 7 The key lemma

Lemma 7.2 below will be used in section 8 to prove the equivalence between the Vershik property and I-cosiness (theorem 8.2), and in section 9 to prove the “coparameterized representation” of joinings (proposition 9.2). Roughly speaking, given two jointly immersed copies  $(\mathcal{F}'_k)_{k \leq n}$  and  $(\mathcal{F}''_k)_{k \leq n}$  of a filtration  $\mathcal{F}$  up to some time  $n$ , and given two  $\mathcal{F}_{n+1}$ -measurable random variables  $X_{n+1}, Y_{n+1}$ , this lemma will allow us to extend the joining until time  $n+1$  while imposing a joint conditional distribution  $\mathcal{L}[(X'_{n+1}, Y''_{n+1}) | \mathcal{F}'_n \vee \mathcal{F}''_n]$  of the respective copies  $X'_{n+1}$  and  $Y''_{n+1}$  of  $X_{n+1}$  and  $Y_{n+1}$ .

**Lemma 7.1.** *Let  $(\Omega, \mathcal{A}, \mathbb{P})$  be a probability space. Let  $\mathcal{B}_0, \mathcal{B}, \mathcal{C}$  and  $\mathcal{D}$  be four sub- $\sigma$ -fields of  $\mathcal{A}$ , with  $\mathcal{B}_0 \subset \mathcal{B}$ . If  $\mathcal{D}$  is conditionally independent of  $\mathcal{B} \vee \mathcal{C}$  given  $\mathcal{B}_0 \vee \mathcal{C}$ , and if  $\mathcal{C}$  is conditionally independent of  $\mathcal{B}$  given  $\mathcal{B}_0$ , then  $\mathcal{D}$  is conditionally independent of  $\mathcal{B}$  given  $\mathcal{B}_0$ .*

*Proof.* Left to the reader.  $\square$

Fact (iv) of lemma 7.2 below will only be used in proposition 9.2, and not in proposition 8.1 from which theorem 8.2 derives.

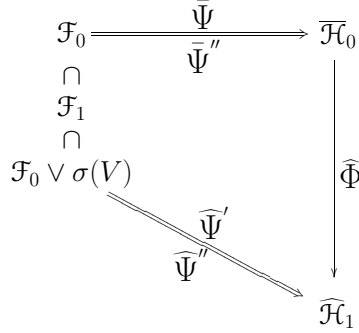


Figure 1: A diagram for lemma 7.2.

**Lemma 7.2.** *Let  $\mathcal{F}_0$  be a  $\sigma$ -field and  $V$  a random variable independent of  $\mathcal{F}_0$  and taking its values in some Polish space. Let  $X$  and  $Y$  be two  $\mathcal{F}_0 \vee \sigma(V)$ -measurable random variables taking their values in Polish spaces  $E$  and  $F$  respectively. Let  $(\bar{\Omega}, \bar{\mathcal{H}}_0, \bar{\mathbb{P}})$  be a probability space, and let  $\bar{\Psi}' : \mathcal{F}_0 \rightarrow \bar{\mathcal{H}}_0$  and  $\bar{\Psi}'' : \mathcal{F}_0 \rightarrow \bar{\mathcal{H}}_0$  be two embeddings. We put  $\bar{\mathcal{F}}'_0 = \bar{\Psi}'(\mathcal{F}_0)$  and  $\bar{\mathcal{F}}''_0 = \bar{\Psi}''(\mathcal{F}_0)$ . Let  $\bar{\Lambda}$  be a random probability on  $E \times F$  measurable with respect to  $\bar{\mathcal{F}}'_0 \vee \bar{\mathcal{F}}''_0$ , with first margin  $\bar{\Psi}'(\mathcal{L}[X | \mathcal{F}_0])$  and second margin  $\bar{\Psi}''(\mathcal{L}[Y | \mathcal{F}_0])$ .*

*There exist an embedding  $\hat{\Phi} : \bar{\mathcal{H}}_0 \rightarrow \hat{\mathcal{H}}_1$  from  $(\bar{\Omega}, \bar{\mathcal{H}}_0, \bar{\mathbb{P}})$  into a probability space  $(\hat{\Omega}, \hat{\mathcal{H}}_1, \hat{\mathbb{P}})$ , and two embeddings  $\hat{\Psi}', \hat{\Psi}'' : \mathcal{F}_0 \vee \sigma(V) \rightarrow \hat{\mathcal{H}}_1$  such that, putting  $\hat{\mathcal{H}}_0 = \hat{\Phi}(\bar{\mathcal{H}}_0)$ ,  $\hat{\mathcal{F}}'_0 = \hat{\Phi}(\bar{\mathcal{F}}'_0)$ ,  $\hat{\mathcal{F}}''_0 = \hat{\Phi}(\bar{\mathcal{F}}''_0)$ , and  $\hat{\Lambda} = \hat{\Phi}(\bar{\Lambda})$ , we have the following facts:*

- (i) *the restrictions of  $\hat{\Psi}'$  and  $\hat{\Psi}''$  to  $\mathcal{F}_0$  equal  $\hat{\Phi} \circ \bar{\Psi}'$  and  $\hat{\Phi} \circ \bar{\Psi}''$  respectively (the diagram on figure 1 is commutative);*
- (ii)  $\mathcal{L}[(\hat{X}', \hat{Y}'') | \hat{\mathcal{F}}'_0 \vee \hat{\mathcal{F}}''_0] = \hat{\Lambda}$  where  $\hat{X}' = \hat{\Psi}'(X)$  and  $\hat{Y}'' = \hat{\Psi}''(Y)$ ;
- (iii)  $\hat{V}' := \hat{\Psi}'(V)$  and  $\hat{V}'' := \hat{\Psi}''(V)$  are each independent of  $\hat{\mathcal{H}}_0$ , and consequently, by lemma 2.7, for any  $\sigma$ -field  $\mathcal{F}_1$  such that  $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_0 \vee \sigma(V)$ , the  $\sigma$ -fields  $\hat{\mathcal{F}}'_1 := \hat{\Psi}'(\mathcal{F}_1)$  and  $\hat{\mathcal{F}}''_1 := \hat{\Psi}''(\mathcal{F}_1)$  are conditionally independent of  $\hat{\mathcal{H}}_0$  given  $\hat{\mathcal{F}}'_0$  and  $\hat{\mathcal{F}}''_0$ , respectively;

(iv) the random pair  $(\widehat{V}', \widehat{V}'')$  is conditionally independent of  $\widehat{\mathcal{H}}_0$  given  $\widehat{\mathcal{F}}'_0 \vee \widehat{\mathcal{F}}''_0$ .

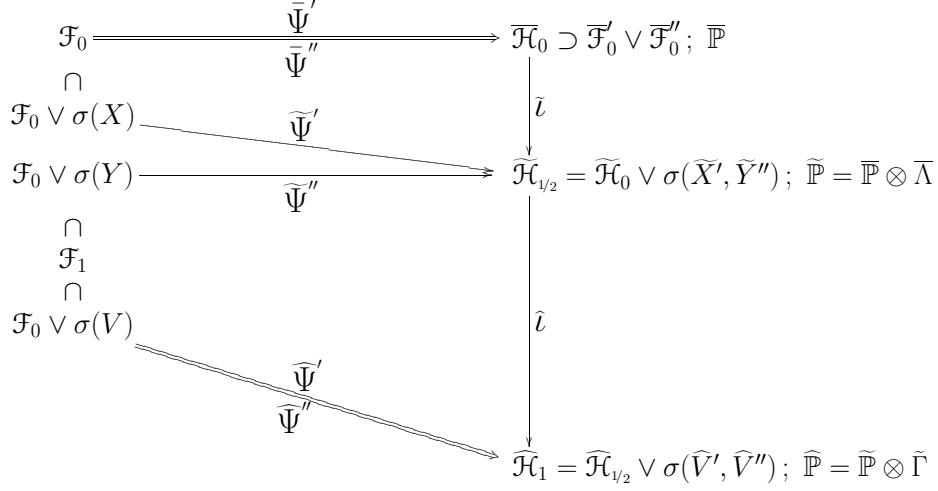


Figure 2: A helpful diagram.

*Proof.* The picture on figure 2 is helpful when reading the proof. As an intermediate step, we firstly define the measurable space

$$(\widetilde{\Omega}, \widetilde{\mathcal{H}}_{1/2}) = (\widetilde{\Omega}, \widetilde{\mathcal{H}}_0) \otimes (E \times F, \mathfrak{B}_{E \times F})$$

and equip it with the probability  $\widetilde{\mathbb{P}} = \widetilde{\mathbb{P}} \otimes \widetilde{\Lambda}$  defined in lemma 2.4. Thus the identification with the first factor  $\widetilde{\iota}: \widetilde{\mathcal{H}}_0 \rightarrow \widetilde{\mathcal{H}}_{1/2}$  is an embedding and one has  $\widetilde{\iota}(\widetilde{\Lambda}) = \mathcal{L}[(\widetilde{X}', \widetilde{Y}'') \mid \widetilde{\mathcal{H}}_0] = \mathcal{L}[(\widetilde{X}', \widetilde{Y}'') \mid \widetilde{\mathcal{F}}'_0 \vee \widetilde{\mathcal{F}}''_0]$  where  $\widetilde{\mathcal{H}}_0 = \widetilde{\iota}(\mathcal{H}_0)$ ,  $\widetilde{\mathcal{F}}'_0 = \widetilde{\iota}(\mathcal{F}'_0)$  and  $\widetilde{\mathcal{F}}''_0 = \widetilde{\iota}(\mathcal{F}''_0)$  and  $(\widetilde{X}', \widetilde{Y}'')$  is defined by  $(\widetilde{X}', \widetilde{Y}'')(\overline{\omega}, x', y'') = (x', y'')$ . As the first and second margins  $\mathcal{L}[\widetilde{X}' \mid \widetilde{\mathcal{H}}_0]$  and  $\mathcal{L}[\widetilde{Y}'' \mid \widetilde{\mathcal{H}}_0]$  of  $\widetilde{\Lambda} := \widetilde{\iota}(\widetilde{\Lambda})$  respectively equal  $\widetilde{\iota} \circ \widetilde{\Psi}'(\mathcal{L}[X \mid \mathcal{F}_0])$  and  $\widetilde{\iota} \circ \widetilde{\Psi}''(\mathcal{L}[Y \mid \mathcal{F}_0])$ , they respectively are measurable with respect to  $\widetilde{\mathcal{F}}'_0$  and  $\widetilde{\mathcal{F}}''_0$ , hence

$$\mathcal{L}[\widetilde{X}' \mid \widetilde{\mathcal{H}}_0] = \mathcal{L}[\widetilde{X}' \mid \widetilde{\mathcal{F}}'_0] = \widetilde{\iota} \circ \widetilde{\Psi}'(\mathcal{L}[X \mid \mathcal{F}_0])$$

and

$$\mathcal{L}[\widetilde{Y}'' \mid \widetilde{\mathcal{H}}_0] = \mathcal{L}[\widetilde{Y}'' \mid \widetilde{\mathcal{F}}''_0] = \widetilde{\iota} \circ \widetilde{\Psi}''(\mathcal{L}[Y \mid \mathcal{F}_0]).$$

Therefore, proposition 2.3 provides two isomorphisms  $\widetilde{\Psi}': \mathcal{F}_0 \vee \sigma(X) \rightarrow \widetilde{\mathcal{F}}'_0 \vee \sigma(\widetilde{X}')$  and  $\widetilde{\Psi}'': \mathcal{F}_0 \vee \sigma(Y) \rightarrow \widetilde{\mathcal{F}}''_0 \vee \sigma(\widetilde{Y}'')$  whose restrictions to  $\mathcal{F}_0$  respectively extend  $\widetilde{\iota} \circ \widetilde{\Psi}'$  and  $\widetilde{\iota} \circ \widetilde{\Psi}''$ , and which respectively send  $X$  to  $\widetilde{X}'$  and  $Y$  to  $\widetilde{Y}''$ .

For the sake of notational convenience, we assume that  $V$  takes its values in  $\mathbb{R}$ . Now, we extend  $(\tilde{\Omega}, \tilde{\mathcal{H}}_{1/2}, \tilde{\mathbb{P}})$  into  $(\hat{\Omega}, \hat{\mathcal{H}}_1, \hat{\mathbb{P}})$  where

$$(\hat{\Omega}, \hat{\mathcal{H}}_1) = (\tilde{\Omega}, \tilde{\mathcal{H}}_{1/2}) \otimes (\mathbb{R} \times \mathbb{R}, \mathfrak{B}_{\mathbb{R} \times \mathbb{R}})$$

and  $\hat{\mathbb{P}} = \tilde{\mathbb{P}} \otimes \tilde{\Gamma}$  (lemma 2.4) where

$$\tilde{\Gamma} = \tilde{\Psi}'(\mathcal{L}[V | \mathcal{F}_0 \vee \sigma(X)]) \otimes \tilde{\Psi}''(\mathcal{L}[V | \mathcal{F}_0 \vee \sigma(Y)])$$

is the  $\tilde{\mathcal{H}}_{1/2}$ -measurable random probability on  $(\mathbb{R} \times \mathbb{R}, \mathfrak{B}_{\mathbb{R} \times \mathbb{R}})$  defined as the product of the respective copies of  $\mathcal{L}[V | \mathcal{F}_0 \vee \sigma(X)]$  and  $\mathcal{L}[V | \mathcal{F}_0 \vee \sigma(Y)]$  by  $\tilde{\Psi}'$  and  $\tilde{\Psi}''$ . We introduce the canonical embedding  $\hat{\iota}: \tilde{\mathcal{H}}_{1/2} \rightarrow \hat{\mathcal{H}}_1$ , we put  $\hat{\mathcal{H}}_{1/2} = \hat{\iota}(\tilde{\mathcal{H}}_{1/2})$ ,  $\hat{X}' = \hat{\iota}(\tilde{X}')$  and  $\hat{Y}'' = \hat{\iota}(\tilde{Y}'')$ , and  $\hat{\Gamma} = \hat{\iota}(\tilde{\Gamma})$ , and we define the random variables  $\hat{V}'$  and  $\hat{V}''$  on  $(\hat{\Omega}, \hat{\mathcal{H}}_1, \hat{\mathbb{P}})$  by

$$\hat{V}'(\bar{\omega}, x', y'', t', t'') = t' \quad \text{and} \quad \hat{V}''(\bar{\omega}, x', y'', t', t'') = t'',$$

so that  $\hat{\Gamma} = \mathcal{L}[(\hat{V}', \hat{V}'') | \hat{\mathcal{H}}_{1/2}]$ .

The conditional law  $\mathcal{L}[\hat{V}' | \hat{\mathcal{H}}_{1/2}]$  is the first margin of  $\hat{\Gamma}$ , hence one has

$$\mathcal{L}[\hat{V}' | \hat{\mathcal{H}}_{1/2}] = \hat{\iota} \circ \tilde{\Psi}'(\mathcal{L}[V | \mathcal{F}_0 \vee \sigma(X)]) = \mathcal{L}[\hat{V}' | \hat{\mathcal{F}}'_0 \vee \sigma(\hat{X}')].$$

Therefore, proposition 2.3 provides an isomorphism  $\hat{\Psi}': \mathcal{F}_0 \vee \sigma(V) \rightarrow \hat{\mathcal{F}}'_0 \vee \sigma(\hat{V}')$  that extends  $\hat{\iota} \circ \tilde{\Psi}'$  on  $\mathcal{F}_0 \vee \sigma(X)$  and that sends  $V$  to  $\hat{V}'$ . This implies that  $\hat{V}'$  is independent of  $\hat{\mathcal{F}}'_0$ . Moreover, the equality  $\mathcal{L}[\hat{V}' | \hat{\mathcal{H}}_{1/2}] = \mathcal{L}[\hat{V}' | \hat{\mathcal{F}}'_0 \vee \sigma(\hat{X}')]$  implies  $\mathcal{L}[\hat{V}' | \hat{\mathcal{H}}_0 \vee \sigma(\hat{X}')] = \mathcal{L}[\hat{V}' | \hat{\mathcal{F}}'_0 \vee \sigma(\hat{X}')]$ , and since we have seen that  $\mathcal{L}[\hat{X}' | \hat{\mathcal{H}}_0] = \mathcal{L}[\hat{X}' | \hat{\mathcal{F}}'_0]$ , lemma 7.1 shows that  $\mathcal{L}[\hat{V}' | \hat{\mathcal{H}}_0] = \mathcal{L}[\hat{V}' | \hat{\mathcal{F}}'_0]$  and thus  $\hat{V}'$  is independent of  $\hat{\mathcal{H}}_0$ . In the same way we obtain an isomorphism  $\hat{\Psi}'': \mathcal{F}_0 \vee \sigma(V) \rightarrow \hat{\mathcal{F}}''_0 \vee \sigma(\hat{V}'')$  that extends  $\hat{\iota} \circ \tilde{\Psi}''$  on  $\mathcal{F}_0 \vee \sigma(Y)$  and that sends  $V$  to  $\hat{V}''$ , and we prove that  $\hat{V}''$  is independent of  $\hat{\mathcal{H}}_0$ . Assertions (i), (ii) and (iii) of the lemma are then fulfilled by putting  $\hat{\Phi} = \hat{\iota} \circ \tilde{\iota}$ .

It remains to show (iv). As  $\sigma(\hat{\Gamma}) \subset \hat{\mathcal{F}}'_0 \vee \hat{\mathcal{F}}''_0 \vee \sigma(\hat{X}', \hat{Y}'')$ , one has

$$\hat{\Gamma} = \mathcal{L}[(\hat{V}', \hat{V}'') | \hat{\mathcal{F}}'_0 \vee \hat{\mathcal{F}}''_0 \vee \sigma(\hat{X}', \hat{Y}'')],$$

hence lemma 7.1 shows that  $\mathcal{L}[(\hat{V}', \hat{V}'') | \hat{\mathcal{H}}_0] = \mathcal{L}[(\hat{V}', \hat{V}'') | \hat{\mathcal{F}}'_0 \vee \hat{\mathcal{F}}''_0]$  because we have seen that  $\mathcal{L}[(\hat{X}', \hat{Y}'') | \hat{\mathcal{H}}_0] = \mathcal{L}[(\hat{X}', \hat{Y}'') | \hat{\mathcal{F}}'_0 \vee \hat{\mathcal{F}}''_0]$ .  $\square$

## 8 Vershik property and I-cosiness

Using the technical lemma 7.2 of the previous section, we prove here the equivalence between the Vershik property and I-cosiness for a random variable with respect to a locally separable filtration (definition 2.9). As explained in the introduction, we state this result for a locally separable filtration because this is cost-free; there is no loss of interest when assuming the filtration to be essentially separable, and no simplifications in the proofs. This result will be a consequence of the following proposition. All no(ta)tions have been introduced in sections 3 and 4.

**Proposition 8.1.** *Let  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  be a locally separable filtration on  $(\Omega, \mathcal{A}, \mathbb{P})$ . Let  $(K, \rho)$  be a compact metric space and  $R \in L^1(\mathcal{F}_0; K)$ . For any  $n_0 < 0$ , there exists, on a probability space  $(\overline{\Omega}, \overline{\mathcal{A}}, \overline{\mathbb{P}})$ , a joining  $(\mathcal{F}', \mathcal{F}'')$  of  $\mathcal{F}$  independent up to  $n_0$  such that  $(\rho_n(\pi_n R', \pi_n R''))_{n \leq 0}$  is a martingale with respect to  $\mathcal{F}' \vee \mathcal{F}''$  on the time-interval  $\{n_0, \dots, 0\}$ , in the sense that*

$$\overline{\mathbb{E}}[\rho_n(\pi_n R', \pi_n R'') \mid \mathcal{F}'_{n-1} \vee \mathcal{F}''_{n-1}] = \rho_{n-1}(\pi_{n-1} R', \pi_{n-1} R'')$$

for every  $n \in \{n_0+1, \dots, 0\}$ . In particular, one has  $\overline{\mathbb{E}}[\rho(R', R'')] = \text{disp}(\pi_{n_0} R)$ .

*Proof.* As  $\mathcal{F}$  is assumed to be locally separable, we know from lemma 2.13 that there exists, up to isomorphism, a parameterization  $(U_n)_{n \leq 0}$  of  $\mathcal{F}$  (definition 2.10). Fix  $n_0 < 0$ . We show by induction on  $m \in \{n_0, \dots, 0\}$  that there exists, on a probability space  $(\overline{\Omega}, \overline{\mathcal{A}}, \overline{\mathbb{P}})$ , a joining  $((\mathcal{F}'_n)_{n \leq m}, (\mathcal{F}''_n)_{n \leq m})$  independent up to  $n_0$  of the truncated filtration  $(\mathcal{F}_n)_{n \leq m}$  such that

$$\overline{\mathbb{E}}[\rho_n(\gamma'_n, \gamma''_n) \mid \mathcal{F}'_{n-1} \vee \mathcal{F}''_{n-1}] = \rho_{n-1}(\gamma'_{n-1}, \gamma''_{n-1})$$

for every  $n \in \{n_0 + 1, \dots, m - 1, m\}$ , where  $\gamma'_n$  and  $\gamma''_n$  are the respective copies of  $\pi_n R$ . The proposition then follows from this fact for  $m = 0$ .

For  $m = n_0$ , it suffices to consider two independent copies of  $(\mathcal{F}_n)_{n \leq n_0}$  and there is nothing else to do. Suppose that the assertion holds true at rank  $m$ . To extend the construction to  $m + 1$ , it suffices to apply lemma 7.2 with  $V = U_{m+1}$ ,  $X = Y = \pi_{m+1} R$  and with  $\overline{\Lambda}$  a probability kernel such that  $\rho_m(\gamma'_m, \gamma''_m) = \iint \rho_{m+1}(\cdot, \cdot) d\overline{\Lambda}(\cdot, \cdot)$ , obtained thanks to a Borel selection theorem (see [Bog]) to guarantee measurability. The joint immersion follows from the last point of lemma 7.2 and from lemma 2.6.  $\square$

**Theorem 8.2.** *Let  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  be a locally separable filtration. Let  $(K, \rho)$  be a compact metric space and  $X \in L^1(\mathcal{F}_0; K)$ . The following statements are equivalent:*

- (i)  *$X$  is I-cosy (definition 2.16);*
- (ii)  *$X$  is Vershikian (definition 4.1).*

*Proof.* Proposition 4.2 shows that (i)  $\implies$  (ii), and (ii)  $\implies$  (i) is a consequence of proposition 8.1.  $\square$

**Corollary 8.3.** *Let  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  be a locally separable filtration and  $\mathcal{E}_0 \subset \mathcal{F}_0$  a  $\sigma$ -field. Then  $\mathcal{E}_0$  is Vershikian if and only if  $\mathcal{E}_0$  is I-cosy.*

*Proof.* This follows from theorem 8.2 and lemma 2.21.  $\square$

The following corollary is announced in remark 3.48 of [Lau10].

**Corollary 8.4.** *Let  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  and  $\mathcal{G} = (\mathcal{G}_n)_{n \leq 0}$  be two locally separable filtrations with  $\mathcal{F}$  immersed in  $\mathcal{G}$ , and  $\mathcal{E}_0 \subset \mathcal{F}_0$  a  $\sigma$ -field. Then  $\mathcal{E}_0$  is I-cosy with respect to  $\mathcal{F}$  if and only if  $\mathcal{E}_0$  is I-cosy with respect to  $\mathcal{G}$ . Consequently, given a Polish space  $E$ , the same result holds with  $\mathcal{E}_0$  replaced by a random variable  $X \in L^1(\mathcal{F}_0; E)$ .*

*Proof.* The analogous fact for the Vershik property holds true owing to lemma 4.3. Hence it also holds for I-cosiness in view of corollary 8.3. The consequence results from lemma 2.20.  $\square$

The equivalence between I-cosiness and the Vershik property for a  $\sigma$ -field, stated in corollary 8.3, immediately extends to filtrations:

**Theorem 8.5.** *Let  $\mathcal{F}$  be a locally separable filtration. The following assertions are equivalent:*

- (i)  *$\mathcal{F}$  is I-cosy;*
- (ii)  *$\mathcal{F}$  is Vershikian.*



## 9 Coparameterized joinings

The main result of this section is proposition 9.2 which is a by-product of lemma 7.2. Recall that a *joining* of a filtration  $\mathcal{F}$  is a pair of two jointly immersed copies of  $\mathcal{F}$ . Roughly speaking, this proposition firstly says that every joining  $(\mathcal{F}', \mathcal{F}'')$  of a locally separable  $\mathcal{F}$  can be written with the help of a local superinnovation  $\mathbf{V}'$  of  $\mathcal{F}'$  and a local superinnovation  $\mathbf{V}''$  of  $\mathcal{F}''$ , where  $\mathbf{V}'$  and  $\mathbf{V}''$  are the copies of a given local superinnovation  $\mathbf{V}$  of  $\mathcal{F}$ . It says in addition that this is also true for a global superinnovation under the assumption that the joining is independent in small time.

**Definition 9.1.** Let  $\mathcal{F}$  be a filtration.

- Two joinings  $(\mathcal{F}', \mathcal{F}'')$  and  $(\mathcal{F}^*, \mathcal{F}^{**})$  of  $\mathcal{F}$  are *isomorphic* if there exists an isomorphism from the filtration  $\mathcal{F}' \vee \mathcal{F}''$  to the filtration  $\mathcal{F}^* \vee \mathcal{F}^{**}$  that sends  $\mathcal{F}'$  to  $\mathcal{F}^*$  and  $\mathcal{F}''$  to  $\mathcal{F}^{**}$ .
- A joining  $(\mathcal{F}', \mathcal{F}'')$  is *immersed* in a joining  $(\mathcal{G}', \mathcal{G}'')$  of a filtration  $\tilde{\mathcal{G}}$  if  $\mathcal{F}' \vee \mathcal{F}''$  is immersed in  $\mathcal{G}' \vee \mathcal{G}''$  (this implies that  $\mathcal{F}'$  and  $\mathcal{F}''$ , respectively, are immersed in  $\mathcal{G}'$  and  $\mathcal{G}''$ , hence that  $\tilde{\mathcal{G}}$  is an extension of  $\mathcal{F}$ ).
- A joining  $(\mathcal{F}', \mathcal{F}'')$  of  $\mathcal{F}$  is *immersible* in a joining  $(\mathcal{G}^*, \mathcal{G}^{**})$  if  $(\mathcal{G}^*, \mathcal{G}^{**})$  is a joining of an extension  $\tilde{\mathcal{G}}$  of  $\mathcal{F}$  that induces a joining  $(\mathcal{F}^*, \mathcal{F}^{**})$  of  $\mathcal{F}$  isomorphic to  $(\mathcal{F}', \mathcal{F}'')$  and immersed in  $(\mathcal{G}^*, \mathcal{G}^{**})$ .

**Proposition 9.2.** Let  $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$  be a filtration and  $\mathbf{V} = (V_n)_{n \leq 0}$  a global superinnovation of  $\mathcal{F}$ . Let  $(\mathcal{F}', \mathcal{F}'')$  be a joining of  $\mathcal{F}$ .

For every integer  $n_0 < 0$ , the joining  $(\mathcal{F}', \mathcal{F}'')$  is immersible in a joining of the extension of  $\mathcal{F}$  with the local superinnovation  $\mathbf{V}^{\text{loc}} := (V_{n_0+1}, \dots, V_0)$ .

If  $\mathcal{F}'$  and  $\mathcal{F}''$  are independent in small time, the same result holds true with the global superinnovation  $\mathbf{V}$  instead of the local superinnovation  $\mathbf{V}^{\text{loc}}$ .

*Proof.* Let  $n_0 < 0$  be an integer. Let  $\mathcal{G}$  be the extension of  $\mathcal{F}$  with the local superinnovation  $\mathbf{V}^{\text{loc}}$  in the case when  $(\mathcal{F}', \mathcal{F}'')$  is an arbitrary joining of  $\mathcal{F}$ , or let  $\mathcal{G}$  be the extension of  $\mathcal{F}$  with the global superinnovation  $\mathbf{V}$  in the case when  $(\mathcal{F}', \mathcal{F}'')$  is a joining independent up to  $n_0$ . Let  $(\overline{\Omega}, \overline{\mathcal{A}}, \overline{\mathbb{P}})$  be the probability space on which  $(\mathcal{F}', \mathcal{F}'')$  is given and let  $\Psi', \Psi'' : \mathcal{F}_0 \rightarrow \overline{\mathcal{A}}$  be the two embeddings defining  $(\mathcal{F}', \mathcal{F}'')$ .

Now we recursively prove that for any  $m \in \{n_0, \dots, 0\}$  we have an embedding  $\hat{\Phi}$  from  $(\overline{\Omega}, \mathcal{F}'_m \vee \mathcal{F}''_m, \overline{\mathbb{P}})$  into a probability space  $(\hat{\Omega}, \hat{\mathcal{A}}, \hat{\mathbb{P}})$ , and two embeddings  $\hat{\Psi}', \hat{\Psi}'' : \mathcal{G}_m \rightarrow \hat{\mathcal{A}}$  such that

- the restrictions of  $\widehat{\Psi}'$  and  $\widehat{\Psi}''$  to  $\mathcal{F}_m$  equal  $\widehat{\Phi} \circ \Psi'$  and  $\widehat{\Phi} \circ \Psi''$  respectively;
- the respective copies  $(\widehat{\mathcal{G}}'_n)_{n \leq m}$  and  $(\widehat{\mathcal{G}}''_n)_{n \leq m}$  of  $(\mathcal{G}_n)_{n \leq m}$  by  $\widehat{\Psi}'$  and  $\widehat{\Psi}''$  are jointly immersed and  $(\widehat{\mathcal{F}}'_n \vee \widehat{\mathcal{F}}''_n)_{n \leq m}$  is immersed in  $(\widehat{\mathcal{G}}'_n \vee \widehat{\mathcal{G}}''_n)_{n \leq m}$ .

For  $m = n_0$ , in the first case, where we have  $\mathcal{G}_n = \mathcal{F}_n$  for  $n \leq n_0$ , it suffices to take  $\widehat{\Psi}' = \Psi'$ ,  $\widehat{\Psi}'' = \Psi''$ , and  $\widehat{\Phi} = \text{Id}$ . For  $m = n_0$ , in the second case, we consider, on some probability space  $(\Omega^*, \mathcal{A}^*, \mathbb{P}^*)$ , two independent copies  $(\mathcal{G}_n^*)_{n \leq n_0}$  and  $(\mathcal{G}_n^{**})_{n \leq n_0}$  of the filtration  $\mathcal{G}$  up to  $n_0$ , given by two isomorphisms  $\Psi^*: \mathcal{G}_{n_0} \rightarrow \mathcal{G}_{n_0}^*$  and  $\Psi^{**}: \mathcal{G}_{n_0} \rightarrow \mathcal{G}_{n_0}^{**}$ , and we define  $\Phi^*: \mathcal{F}'_{n_0} \vee \mathcal{F}''_{n_0} \rightarrow \Psi^*(\mathcal{F}_{n_0}) \vee \Psi^{**}(\mathcal{F}_{n_0})$  as the unique isomorphism that extends  $\Psi^* \circ (\Psi')^{-1}$  on  $\mathcal{F}'_{n_0}$  and  $\Psi^{**} \circ (\Psi'')^{-1}$  on  $\mathcal{F}''_{n_0}$ , whose existence is given by proposition 2.2.

Assuming the construction has been performed at rank  $m - 1$ , we extend it to rank  $m$  by taking a random variable  $W_m$  such that  $\mathcal{F}_m = \mathcal{F}_{m-1} \vee \sigma(W_m)$ , whose existence is guaranteed by lemma 2.13, and by applying lemma 7.2 with  $V = V_m$ ,  $X = Y = W_m$ , and  $\bar{\Lambda} = \widehat{\Phi}(\mathcal{L}[W'_m, W''_m | \mathcal{F}'_{m-1} \vee \mathcal{F}''_{m-1}])$ . Using lemma 2.6 and lemma 2.7, it is easy to check the joint immersion of  $(\widehat{\mathcal{G}}'_n)_{n \leq m}$  and  $(\widehat{\mathcal{G}}''_n)_{n \leq m}$  as well as the immersion of  $(\widehat{\mathcal{F}}'_n \vee \widehat{\mathcal{F}}''_n)_{n \leq m}$  in  $(\widehat{\mathcal{G}}'_n \vee \widehat{\mathcal{G}}''_n)_{n \leq m}$  with the help of the random variables  $\widehat{V}'_m := \widehat{\Psi}'(V_m)$  and  $\widehat{V}''_m := \widehat{\Psi}''(V_m)$  provided by lemma 7.2.  $\square$

Remark that the immersibility of  $(\mathcal{F}', \mathcal{F}'')$  in the joining of the extension of  $\mathcal{F}$  with  $\mathbf{V}^{\text{loc}}$  or  $\mathbf{V}$  is a consequence of assertion (iv) of lemma 7.2; however, if we denote by  $(\widehat{\mathcal{G}}', \widehat{\mathcal{G}}'')$  this joining, assertion (iv) actually says that each pair  $(\widehat{V}'_n, \widehat{V}''_n)$  is conditionally independent of  $\widehat{\mathcal{G}}'_{n-1} \vee \widehat{\mathcal{G}}''_{n-1}$  given  $\widehat{\mathcal{F}}'_{n-1} \vee \widehat{\mathcal{F}}''_{n-1}$ , and this property is not implied by the immersion of  $(\widehat{\mathcal{F}}', \widehat{\mathcal{F}}'')$  in  $(\widehat{\mathcal{G}}', \widehat{\mathcal{G}}'')$ .

## References

- [Bil] P. Billingsley, *Convergence of Probability Measures*, 1968. John Wiley & Sons, Inc., New York.
- [Bog] V.I. Bogachev, *Measure Theory*, Volume 2, 2007. Springer-Verlag, Berlin.
- [Ceil09] G. Ceillier, *The filtration of the split-word processes*. Preprint, 2009

- [DM] C. Dellacherie, P-A. Meyer, *Probabilités et potentiel*, Chapitres I à IV, 1975. Hermann, Paris.
- [Dud] R.M. Dudley, *Real Analysis and Probability*, 1989. Wadsworth and Brooks/Cole Math Series, Pacific Grove.
- [Em] M. Émery, *Espaces probabilisés filtrés: de la théorie de Vershik au mouvement brownien, via des idées de Tsirelson*. Séminaire Bourbaki 43, Exp. 882, 63–83, Vol. 2000/2001,
- [ES01] M. Émery, W. Schachermayer, *On Vershik’s standardness criterion and Tsirelson’s notion of cosiness*, Séminaire de probabilités XXXV, Springer Lectures Notes in Math. 1755, 265–305, 2001.
- [Kall] O. Kallenberg, *Foundations of Modern Probability*, 2nd ed, 2001. Springer, Berlin Heidelberg New York, 2001.
- [Lau10] S. Laurent, *On standardness and I-cosiness*. To appear in: Séminaire de Probabilités XLIII, 2010.
- [Par89] W. Parry, *Decoding with Two Independent Processes*. Contemporary Mathematics **94**, 207–209, 1989.
- [Tsi97] B. Tsirelson, *Triple points: from non-Brownian filtrations to harmonic measures*. Geom. Funct. Anal. (GAFA) **7**, 1096–1142, 1997.
- [Ver68] A.M. Vershik, *Theorem on lacunary isomorphisms of monotonic sequences of partitions*. Funktsional’nyi Analiz i Ego Prilozheniya **2:3**, 17–21, 1968. English translation: Functional Analysis and Its Applications **2:3**, 200–203, 1968.
- [Ver70] A.M. Vershik, *Decreasing sequences of measurable partitions, and their applications*. Dokl. Akad. Nauk SSSR, **193**, 748–751, 1970. English translation: Soviet Math. Dokl. **11**, 1007–1011, 1970.
- [Ver71] A.M. Vershik, *Continuum of pairwise nonisomorphic diadic sequences*. Funktsional’nyi Analiz i Ego Prilozheniya **5:3**, 16–18, 1971. English translation: Functional Analysis and Its Applications **5:3**, 182–184, 1971.

- [Ver73a] A.M. Vershik, *Approximation in measure theory* (in Russian). PhD Dissertation, Leningrad University, 1973. (expanded and updated version: [Ver94])
- [Ver73b] A.M. Vershik, *Four definitions of the scale of an automorphism*. Funktsional'nyi Analiz i Ego Prilozheniya **7**:3, 1–17, 1973. English translation: Functional Analysis and Its Applications **7**:3, 169–181, 1973.
- [Ver94] A.M. Vershik, *The theory of decreasing sequences of measurable partitions*. Algebra i Analiz **6**:4, 1–68, 1994 (in Russian). English translation: St. Petersburg Mathematical Journal **6**:4, 705–761, 1995.
- [Ver04] A.M. Vershik, *The Kantorovich metric: the initial history and little-known applications*. Representation theory, dynamical systems. Part XI, Special issue, Zap. Nauchn. Sem. POMI **312**, POMI, St. Petersburg, 69–85, 2004. English translation: J. Math. Sci. (New York) **133**:4, 1410–1417, 2006.
- [VG07] A.M. Vershik, A.D. Gorbulsky, *Scaled entropy of filtrations of  $\sigma$ -fields*. Teor. Veroyatnost. i Primenen. **52**:3, 446–467, 2007. English translation: Theory of Probability and its Applications **52**:3, 493–508, 2008.