

I N S T I T U T D E
S T A T I S T I Q U E

UNIVERSITÉ CATHOLIQUE DE LOUVAIN



D I S C U S S I O N
P A P E R

0914

ON STANDARDNESS AND I-COSINESS

LAURENT, S.

This file can be downloaded from
<http://www.stat.ucl.ac.be/ISpub>

On standardness and I-cosiness

Stéphane Laurent

Université catholique de Louvain, Louvain-la-Neuve, Belgium.
`stephane.laurent@uclouvain.be`

Summary. The object of study of this work is the invariant characteristics of filtrations in discrete, negative time, pioneered by Vershik. We prove the equivalence between I-cosiness and standardness without using Vershik's standardness criterion. The equivalence between I-cosiness and productness for homogeneous filtrations is further investigated by showing that the I-cosiness criterion is equivalent to Vershik's first level criterion separately for each random variable. We also aim to derive the elementary properties of both these criteria, and to give a survey and some complements on the published and unpublished literature.

Acknowledgments. Financial support from the IAP research network (grant nr. P6/03 of the Belgian government, Belgian Science Policy) is gratefully acknowledged. I am also indebted to M. Émery for helpful and encouraging comments and suggestions on earlier drafts of this paper.

1 Introduction

A filtration $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ in discrete, negative time, is said to be of *local product type* if there exists a sequence $(V_n)_{n \leq 0}$ of independent random variables such that for each $n \leq 0$, one has $\mathcal{F}_n = \mathcal{F}_{n-1} \vee \sigma(V_n)$ and V_n is independent of \mathcal{F}_{n-1} . Such random variables V_n are called *innovations* of \mathcal{F} . A typical example is the case of a filtration generated by a sequence of independent random variables, termed as *filtration of product type*.

Originally, the theory of decreasing sequences of measurable partitions investigated by Vershik ([36, 37, 38, 39, 41]) was mainly oriented towards characterizing productness for *homogeneous* filtrations of local product type, that is, those for which each innovation V_n has either a uniform distribution on a finite set or a diffuse law. The *standardness criterion* introduced by Vershik provides such a characterization under the assumption that the final σ -field \mathcal{F}_0 of the filtration is essentially separable (in other words, it is countably

generated up to negligible sets, and we also say that the filtration is *essentially separable*).

Vershik's standardness criterion makes sense not only in the context of filtrations of local product type, and it characterizes essentially separable filtrations $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ having an *extension* of product type, hereafter called *standard filtrations*.

Vershik's theory of filtrations in discrete, negative time remained unknown to the western probabilistic culture for about 25 years, until Dubins, Feldman, Smorodinsky and Tsirelson used *Vershik's standardness criterion* in [10]. Later, in [14], Émery and Schachermayer partially translated Vershik's theory into the language of stochastic processes, and introduced the *I-cosiness criterion*, inspired by the notion of cosiness which Tsirelson devised in [35] and by Smorodinsky's proof in [32] that the filtration of a *split-word process* is not standard. In the context of essentially separable filtrations, the results of Vershik's theory of filtrations are summarized in figure 1.

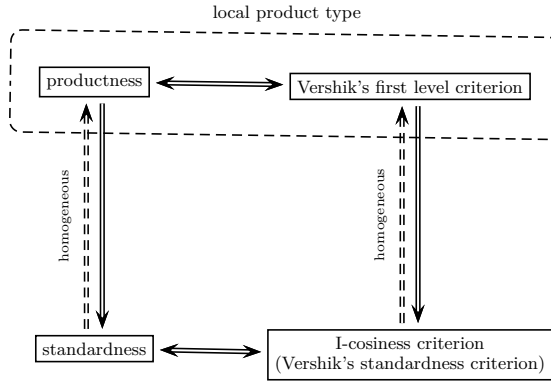


Fig. 1. Theorems for essentially separable filtrations

Among the contents of this paper is a proof of the following theorem.

Theorem A. *A homogeneous filtration $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ with an essentially separable final σ -field \mathcal{F}_0 is I-cosy if and only if it is generated by a sequence of independent random variables.*

The proof of this theorem is incomplete in [14], for only the case of homogeneous filtrations with diffuse innovations is considered there. Moreover, Vershik's standardness criterion is used to establish this result, whereas we give a more direct proof without using this criterion, which actually is not even stated in the present paper.

The proofs given in the literature ([41], [14], [15]) of theorem A, or of the equivalence between Vershik's standardness criterion and productness for a homogeneous filtration, use *Vershik's first level criterion* as a key step,

without naming it. Vershik's first level criterion is known to be equivalent to productness when the filtration is essentially separable. Roughly speaking, this criterion says that any random variable measurable with respect to the final σ -field can be approximated by a function of finitely many innovations.

Thus, theorem A derives from the equivalence between I-cosiness and Vershik's first level criterion for an essentially separable homogeneous filtration. We will see that this equivalence is still valid without assuming the filtration to be essentially separable, and thus we will deduce theorem A from theorem A' below.

Theorem A'. *A homogeneous filtration $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ is I-cosy if and only if it satisfies Vershik's first level criterion.*

Actually all our results will be stated under a weaker assumption than essential separability of the filtration. Namely, in this paper, the standing assumption on filtrations is *local separability*; we say that a filtration is *locally separable* if it admits essentially separable increments, with a final σ -field which is not necessarily essentially separable. All these results are summarized in figure 2.

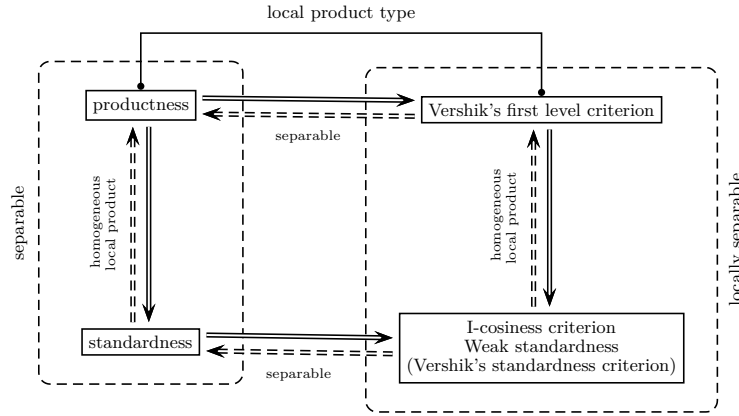


Fig. 2. Theorems for locally separable filtrations

For example, theorem B below admits theorem B' as its analogue for locally separable filtration. Theorems B and B' are elementarily deduced from theorems A and A' respectively.

Theorem B. *An essentially separable filtration $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ is I-cosy if and only if it is standard.*

Theorem B'. *A locally separable filtration $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ is I-cosy if and only if it is weakly standard.*

The definition of *weak standardness* is analogous to that of standardness, with productness replaced by Vershik's first level criterion: whereas a filtration is standard if it admits an extension of product type, a filtration is weakly standard if it admits an extension satisfying Vershik's first level criterion.

Our hypothesis of locally separability, less stringent than requiring the final σ -field to be essentially separable, has no practical interest; but it requires no additional efforts, and it sometimes provides a better understanding of the results.

Actually, our additional efforts are oriented towards investigating the I-cosiness criterion *for a random variable with respect to a filtration*, and not only for the whole filtration, in the following sense. The definition of the I-cosiness criterion for a filtration $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ requires a certain property, say $I(X)$, to hold for each 'test' random variable X measurable with respect to the final σ -field \mathcal{F}_0 . This property $I(X)$ will be called *I-cosiness of the random variable X (with respect to \mathcal{F})*. Shortly:

$$\underbrace{\forall X, \overbrace{I(X)}}_{\text{I-cosiness of } \mathcal{F}}.$$

Vershik's first level criterion has the same structure, and we will similarly define *Vershik's first level criterion for a random variable*. Then theorem A' will be an immediate consequence of theorem A" below.

Theorem A". *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a homogeneous filtration. Then a random variable is I-cosy with respect to \mathcal{F} if and only if it satisfies Vershik's first level criterion with respect to \mathcal{F} .*

This theorem is more interesting than theorem A, and its proof is not simplified when \mathcal{F} is essentially separable. Thus, our generalization to locally separable filtrations is only a by-product of our investigations of the I-cosiness criterion and Vershik's first level criterion at the level of random variables. We will also obtain the following characterization of I-cosiness for a random variable with respect to a general locally separable filtration.

Theorem. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a locally separable filtration. Then a random variable is I-cosy with respect to \mathcal{F} if and only if it satisfies Vershik's first level criterion with respect to a homogeneous extension of \mathcal{F} with diffuse innovations.*

In the same spirit, Vershik's standardness criterion, which is not stated in this paper, is investigated "random variable by random variable" in [21], where we show it to be equivalent to the I-cosiness criterion under the local separability assumption. The proof is self-contained and no familiarity with the subject is needed. Throughout this paper, we will sometimes announce results from [21].

Main notations and conventions

By a probability space, we always mean a triple $(\Omega, \mathcal{A}, \mathbb{P})$ where the σ -field \mathcal{A} is \mathbb{P} -complete. By a σ -field $\mathcal{C} \subset \mathcal{A}$ we always mean an $(\mathcal{A}, \mathbb{P})$ -complete σ -field. By a random variable on $(\Omega, \mathcal{A}, \mathbb{P})$, we mean a \mathbb{P} -equivalence class of measurable maps from Ω to a separable metric space. By convention, the σ -field generated by an empty family of random variables equals the trivial σ -field $\{\emptyset, \Omega\}$ up to negligible sets. A σ -field \mathcal{C} is *essentially separable* if it is countably generated up to negligible sets. Thus a random variable X generates an essentially separable σ -field $\sigma(X)$; equivalently, an essentially separable σ -field is a σ -field generated by a real-valued random variable. We will extensively use the following elementary lemma, which is often implicit in the probabilistic literature.

Lemma 1.1. *On $(\Omega, \mathcal{A}, \mathbb{P})$, let \mathcal{B} and \mathcal{C} be two σ -fields. For any random variable X measurable with respect to $\mathcal{B} \vee \mathcal{C}$, there exist a \mathcal{B} -measurable random variable B and a \mathcal{C} -measurable random variable C such that $\sigma(X) \subset \sigma(B, C)$.*

This lemma derives from the equality $\mathcal{B} \vee \mathcal{C} = \bigcup_{B, C} \sigma(B, C)$, where B and C range over all \mathcal{B} -measurable r.v. and all \mathcal{C} -measurable r.v. respectively. Of course we can also take bounded random variables B and C in this lemma.

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, ρ) ; the space $L^0(\mathcal{C}; (E, \rho))$ is endowed with the topology of convergence in probability; when $E = \mathbb{R}$ we just write $L^0(\mathcal{C})$. Similarly, the space $L^1(\mathcal{C}; E)$ is the set of all \mathcal{C} -measurable random variables X taking their values in E 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 $(X, Y) \mapsto \mathbb{E}[\rho(X, Y)]$. It is well-known that $L^0(\mathcal{C}; (E, \rho)) = L^1(\mathcal{C}; (E, \rho \wedge 1))$. The set of all simple, E -valued, \mathcal{C} -measurable random variables is a dense subset of $L^1(\mathcal{C}; E)$. If F is a finite set, we denote by $L(\mathcal{C}; F)$ the set of all \mathcal{C} -measurable random variables taking their values in F , considered as a metric space with the metric $(S, T) \mapsto \mathbb{P}[S \neq T]$. Thus $L(\mathcal{C}; F) = L^1(\mathcal{C}; F)$ where F is equipped with the 0 – 1 distance. The Borel σ -field on a separable metric space E is denoted by \mathfrak{B}_E .

A *Polish metric space* is a complete separable metric space. A *Polish space* is a topological space that admits a separable and complete metrization. A *Polish probability space* is (the completion of) a probability space on a Polish space with its Borel σ -field. Any Polish space F has the *Doob property*: for any measurable space (Ω, \mathcal{A}) , if $X: \Omega \rightarrow T$ is a measurable function taking its values in a measurable space T and $Y: \Omega \rightarrow E$ is a $\sigma(X)$ -measurable function taking its values in a Polish space E , then there exists a measurable function $f: T \rightarrow E$ such that $Y = f(X)$ (see for instance [8]). We will sometimes use the Doob property without invoking its name, or we will also call it *Doob's functional representation theorem*.

When X is a random variable taking values in a Polish space, the existence of the *conditional law* of X given any σ -field \mathcal{C} is guaranteed (see [11]); we denote it by $\mathcal{L}(X | \mathcal{C})$. It is itself a random variable in a Polish space. Some details on conditional laws are provided in Annex A.

Lebesgue isomorphisms

A *Lebesgue isomorphism* between two probability spaces (E, \mathfrak{B}, μ) and (F, \mathfrak{C}, ν) is a bimeasurable bijection T from a set $E_0 \in \mathfrak{B}$ of full μ -measure into a set $F_0 \in \mathfrak{C}$ of full ν -measure, and satisfying $T(\mu) = \nu$. Any Polish probability space is Lebesgue isomorphic to the completion of the Borel space \mathbb{R} equipped with some probability measure (see [6], [25], [29]).

Filtrations in discrete, negative time

On an underlying probability space $(\Omega, \mathcal{A}, \mathbb{P})$, a filtration is an increasing sequence of sub- σ -fields of \mathcal{A} indexed by a time-axis. Most filtrations considered 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 σ -field \mathcal{F}_0 is essentially separable; equivalently, each σ -field \mathcal{F}_n is essentially separable. We say that a filtration \mathcal{F} is *Kolmogorovian* if the tail σ -field $\mathcal{F}_{-\infty} := \bigcap_{n \leq 0} \mathcal{F}_n$ equals the trivial σ -field $\{\emptyset, \Omega\}$ up to negligible sets. A filtration \mathcal{F} is *included* in a filtration \mathcal{G} , and this is denoted by $\mathcal{F} \subset \mathcal{G}$, 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 $(\mathcal{F} \vee \mathcal{G})_n = \mathcal{F}_n \vee \mathcal{G}_n$. The independent product of two filtrations $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ and $\mathcal{G} = (\mathcal{G}_n)_{n \leq 0}$ respectively defined on two probability spaces $(\Omega, \mathcal{A}, \mathbb{P})$ and $(\Omega^*, \mathcal{A}^*, \mathbb{P}^*)$ is the filtration $\mathcal{F} \otimes \mathcal{G}$ defined on the product probability space $(\Omega, \mathcal{A}, \mathbb{P}) \otimes (\Omega^*, \mathcal{A}^*, \mathbb{P}^*)$ by $(\mathcal{F} \otimes \mathcal{G})_n = \mathcal{F}_n \otimes \mathcal{G}_n$.

Random variables, processes

It is understood, if not otherwise specified, that a random variable takes its values in a separable metric space or in \mathbb{R} if this is clear from the context. By a process, we mean a sequence of random variables (each taking its values in a separable metric space if nothing else is specified). Most processes considered in this paper are indexed by the time-axis $-\mathbb{N}$. Such a process $(X_n)_{n \leq 0}$ generates a filtration $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ defined by $\mathcal{F}_n = \sigma(X_m; m \leq n)$. The process $(X_n)_{n \leq 0}$ is *Markovian* if for each $n \leq 0$ the σ -field $\sigma(X_n)$ is conditionally independent of \mathcal{F}_{n-1} given $\sigma(X_{n-1})$. Given a filtration $\mathcal{G} \supset \mathcal{F}$, the process $(X_n)_{n \leq 0}$ is *Markovian with respect to* \mathcal{G} if for each $n \leq 0$ the σ -field $\sigma(X_n)$ is conditionally independent of \mathcal{G}_{n-1} given $\sigma(X_{n-1})$. Equivalently, the process is Markovian and its filtration \mathcal{F} is *immersed* in \mathcal{G} , as we shall see below.

Preliminary notion: immersion

The notion of immersion will be used throughout all this paper. 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 $\mathcal{F} \overset{m}{\subset} \mathcal{G}$ means that \mathcal{F} is immersed in \mathcal{G} . Obviously, the binary relation $\overset{m}{\subset}$ defines a partial order on the set of

filtrations on $(\Omega, \mathcal{A}, \mathbb{P})$. A typical example is provided by lemma 1.2 and some usual characterizations of immersion are given in lemma 1.3.

Lemma 1.2. *Let \mathcal{F} and \mathcal{G} be two independent filtrations. Then both \mathcal{F} and \mathcal{G} are immersed in $\mathcal{F} \vee \mathcal{G}$.*

Two filtrations \mathcal{F} and \mathcal{G} both immersed in $\mathcal{F} \vee \mathcal{G}$ are said to be *jointly immersed*; it suffices that $\mathcal{F} \stackrel{m}{\subset} \mathcal{H}$ and $\mathcal{G} \stackrel{m}{\subset} \mathcal{H}$ for some filtration \mathcal{H} .

Lemma 1.3. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ and $\mathcal{G} = (\mathcal{G}_n)_{n \leq 0}$ 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 the σ -field \mathcal{F}_0 is conditionally independent of \mathcal{G}_n given \mathcal{F}_n for each $n \leq 0$;
- (iii) 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 $n \leq 0$;
- (iv) 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 $n \leq 0$.

Note also that immersion of \mathcal{F} in \mathcal{G} implies $\mathcal{F}_n = \mathcal{F}_0 \cap \mathcal{G}_n$ for all $n \leq 0$. Proofs of the preceding two lemmas are left as an exercise to the reader, as well as those of the next three lemmas, which will frequently be used in this paper. The third one is a straightforward consequence of the first two ones.

Lemma 1.4. *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 σ -field \mathcal{F}_{n+1} is conditionally independent of \mathcal{G}_n given \mathcal{F}_n .*

Lemma 1.5. *If \mathcal{B} , \mathcal{C} and \mathcal{D} are three σ -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 1.6. *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} .*

Here are two straightforward applications of lemma 1.6. First, the filtration \mathcal{F} generated by an independent sequence $(V_n)_{n \leq 0}$ of random variables is immersed in a filtration \mathcal{G} if and only if $\mathcal{F} \subset \mathcal{G}$ and V_n is independent of \mathcal{G}_{n-1} for every $n \leq 0$. Second, a Markov process $(V_n)_{n \leq 0}$ is Markovian with respect to a filtration \mathcal{G} if and only if its generated filtration \mathcal{F} is immersed in \mathcal{G} .

Preliminary notion: isomorphic σ -fields and filtrations

An *embedding* Ψ between two probability spaces $(\Omega, \mathcal{B}, \mathbb{P})$ and $(\Omega', \mathcal{A}', \mathbb{P}')$ is (necessarily injective) map from the quotient σ -field \mathcal{B}/\mathbb{P} to the quotient σ -field \mathcal{A}'/\mathbb{P}' that preserves the σ -field structures and the probabilities. We

shortly write $\Psi: \mathcal{B} \rightarrow \mathcal{A}'$. It is called an *isomorphism* if moreover it is onto. Up to isomorphism, an essentially separable σ -field is characterized by the descending sequence (possibly empty, finite, or denumerable) of the masses of its atoms. An embedding Ψ extends uniquely to random variables taking their values in a Polish space, and we call $\Psi(X)$ the *copy* of such a random variable X . Details are provided in Annex A. However this Annex can be skipped since there is no risk when naively using isomorphisms: any expected property such as $\Psi(f(X)) = f(\Psi(X))$, $\Psi(X, Y) = (\Psi(X), \Psi(Y))$, $\sigma(\Psi(X)) = \Psi(\sigma(X))$, $\Psi(\mathbb{E}[X | \mathcal{C}]) = \mathbb{E}'[\Psi(X) | \Psi(\mathcal{C})]$, is true.

The definition of isomorphic σ -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 *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 say that $\Psi: \mathcal{F} \rightarrow \mathcal{F}'$ is an isomorphism. We denote by $\Psi(\mathcal{F})$ the filtration $(\mathcal{F}'_n)_{n \leq 0} = (\Psi(\mathcal{F}_n))_{n \leq 0}$ and we call it the *copy* of the filtration \mathcal{F} by the isomorphism Ψ .

A typical example of isomorphic filtrations is the case 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. In the case when the X_n (hence the X'_n) take their values in Polish spaces, there exists a unique isomorphism $\Psi: \mathcal{F} \rightarrow \mathcal{F}'$ that sends X_n to X'_n for each $n \leq 0$. This stems from lemma A.7. Another typical example of isomorphic filtrations is provided by the following lemma.

Lemma 1.7. *Let \mathcal{F} and \mathcal{G} be two independent filtrations. Then $\mathcal{F} \vee \mathcal{G}$ is isomorphic to the independent product $\mathcal{F} \otimes \mathcal{G}$ of \mathcal{F} and \mathcal{G} .*

Proof. By proposition A.11, there exists a unique isomorphism extending the canonical embeddings $\iota_1: \mathcal{F}_0 \rightarrow \mathcal{F}_0 \otimes \mathcal{G}_0$ and $\iota_2: \mathcal{G}_0 \rightarrow \mathcal{F}_0 \otimes \mathcal{G}_0$ (defined in example A.2). \square

2 Vershik's first level criterion

This section deals with filtrations of *product type* and *Vershik's first level criterion*.

Definition 2.1. *A filtration is of product type if it is generated by a sequence of independent random variables.*

As we shall see, productness is equivalent to Vershik's first level criterion for an essentially separable filtration of *local product type* (theorem 2.25). This result is far from new: Vershik's first level criterion appears, but without a name, in [41], [14], [15]. Corollary 2.46 shows that the assumption of essential separability cannot be waived: there exist some filtrations of local product type satisfying Vershik's first level criterion but which are not essentially separable, hence not of product type.

As said in the introduction, the important theorem A is deduced from theorem A' and from the equivalence between Vershik's first level criterion and productness for essentially separable filtrations of local product type. Vershik's first level criterion will also be used in section 3 to extend the notion of standardness to the notion of weak standardness (definition 3.21).

Theorem A' stated in the introduction is directly deduced from theorem A"; actually the latter will be proved (in section 4) with the help of the equivalent "self-joining version" of Vershik's first level criterion, which we study in subsection 2.2 and call *Vershik's self-joining criterion*.

In subsection 2.3 we introduce the filtrations of *split-word processes*. We state the theorems on productness for these filtrations which are found in the literature, and we initiate the proofs of these theorems assuming some intermediate key results. At this stage, we will not have at our disposal the tools for finishing these proofs; they will be pursued at the end of each following section, illustrating the new tools we shall acquire.

2.1 Productness and Vershik's first level criterion

In this subsection, we define Vershik's first level criterion and prove its equivalence (theorem 2.25) with productness for an essentially separable filtration of local product type (definition 2.3). We will derive this theorem from theorem 2.23 which gives a characterization of Vershik's first level criterion for filtrations of local product type that are not necessarily essentially separable.

With the terminology of definition 2.2 below, a filtration of product type is a filtration for which there exists a generating innovation.

Definition 2.2. *Given two σ -fields \mathcal{B} and \mathcal{C} such that $\mathcal{C} \subset \mathcal{B}$, an independent complement of \mathcal{C} in \mathcal{B} is a random variable V taking its values in a Polish space, independent of \mathcal{C} and such that $\mathcal{B} = \mathcal{C} \vee \sigma(V)$. An innovation, or a global innovation, of a filtration $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ is a process $(V_n)_{n \leq 0}$ such that for each $n \leq 0$, the random variable V_n is an independent complement of \mathcal{F}_{n-1} in \mathcal{F}_n . An innovation $(V_n)_{n \leq 0}$ is called generating if \mathcal{F} is generated by the process $(V_n)_{n \leq 0}$. For two given integers n_0 and m_0 such that $n_0 < m_0 \leq 0$, an innovation, or a local innovation of \mathcal{F} from n_0 to m_0 is a sequence of random variables $(V_{n_0+1}, \dots, V_{m_0})$ such that V_n is an independent complement of \mathcal{F}_{n-1} in \mathcal{F}_n for each $n \in \{n_0 + 1, \dots, m_0\}$.*

The random variables V_n appearing in a global or a local innovation of a filtration \mathcal{F} are themselves called innovations of \mathcal{F} .

Definition 2.3. *A filtration $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ is of local product type if there exists a global innovation of \mathcal{F} .*

Innovations are not unique in general. They are described by lemma 2.4 below. The notion of *Lebesgue isomorphism* has been recalled in section 1.

Lemma 2.4. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, \mathcal{C} and \mathcal{B} two sub- σ -fields of \mathcal{A} , and V an independent complement of \mathcal{C} in \mathcal{B} taking values in a Polish space E . Let V' be a random variable taking values in a Polish space E' . Then V' is an independent complement of \mathcal{C} in \mathcal{B} if and only if there exist a \mathcal{C} -measurable random variable C and a measurable function $\phi: \mathbb{R} \times E \rightarrow E'$ such that $V' = \phi(C, V)$ and, almost surely, the random map $T_C: v \mapsto \phi(C, v)$ is a Lebesgue isomorphism from the probability space induced by V into the probability space induced by V' . In particular, the σ -fields $\sigma(V)$ and $\sigma(V')$ are isomorphic.*

Proof. The ‘if’ part is easy to verify. To show the ‘only if’ part, assume V' to be an independent complement of \mathcal{C} in \mathcal{B} . There exist (lemma 1.1) two \mathcal{C} -measurable random variables C_1 and C_2 such that $\sigma(V') \subset \sigma(C_1, V)$ and $\sigma(V) \subset \sigma(C_2, V')$. We introduce a real-valued \mathcal{C} -measurable random variable C such that $\sigma(C_1, C_2) \subset \sigma(C)$ and we denote its law by \mathbb{P}_C . By Doob’s functional representation theorem, there exist two measurable functions ϕ and ψ such that $V' = \phi(C, V)$ and $V = \psi(C, V')$. Considering the conditional laws given \mathcal{C} , we see that $\phi(c, V)$ has the same law as V' and $\psi(c, V')$ has the same law as V for \mathbb{P}_C -almost every c . Moreover, since

$$1 = \mathbb{P}[V' = \phi(C, \psi(C, V'))] = \int \mathbb{P}[V' = \phi(c, \psi(c, V'))] d\mathbb{P}_C(c),$$

one has $V' = \phi(c, \psi(c, V'))$ almost surely for \mathbb{P}_C -almost every c . In the same way one has $V = \psi(c, \phi(c, V))$ almost surely for \mathbb{P}_C -almost every c . Hence, for \mathbb{P}_C -almost every c , the Borel subset $E_0 := \{v \in E \mid v = \psi(c, \phi(c, v))\}$ of E has full \mathbb{P}_V -measure, the Borel subset $E'_0 := \{v' \in E' \mid v' = \phi(c, \psi(c, v'))\}$ of E' has full $\mathbb{P}_{V'}$ -measure, and the maps $v \mapsto \phi(c, v)$ and $v' \mapsto \psi(c, v')$ define mutual inverse bijections between E_0 and E'_0 . Finally $v \mapsto \phi(c, v)$ defines a Lebesgue isomorphism from the probability space induced by V into the probability space induced by V' . Consequently, the σ -fields $\sigma(V)$ and $\sigma(V')$ have the same descending sequences of masses of their atoms, and hence are isomorphic. \square

A straightforward application of lemma 1.6 gives the following lemma:

Lemma 2.5. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration of local product type and $(V_n)_{n \leq 0}$ an innovation of \mathcal{F} . Then the filtration generated by $(V_n)_{n \leq 0}$ is immersed in \mathcal{F} . Consequently one has $\sigma(V_m; m \leq n) = \mathcal{F}_n \cap \sigma(V_m; m \leq 0)$ for each $n \leq 0$.*

Now we turn to Vershik’s first level criterion. We shall see at the end of this subsection that this criterion is equivalent to productness for an essentially separable filtration of local product type (theorem 2.25).

Definition 2.6. *On $(\Omega, \mathcal{A}, \mathbb{P})$, let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration of local product type.*

- Let (E, ρ) be a separable metric space and $X \in L^1(\mathcal{F}_0; E)$. The random variable X satisfies Vershik's first level criterion (with respect to \mathcal{F}) if for every $\delta > 0$, there exist an integer $n_0 < 0$, an innovation (V_{n_0+1}, \dots, V_0) of \mathcal{F} from n_0 to 0, and a random variable $S \in L^1(\sigma(V_{n_0+1}, \dots, V_0); E)$ such that $\mathbb{E}[\rho(X, S)] < \delta$.
- A σ -field $\mathcal{E}_0 \subset \mathcal{F}_0$ satisfies Vershik's first level criterion (with respect to \mathcal{F}) if every random variable $X \in L^1(\mathcal{E}_0; \mathbb{R})$ satisfies Vershik's first level criterion with respect to \mathcal{F} .
- The filtration $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ satisfies Vershik's first level criterion if the σ -field \mathcal{F}_0 satisfies Vershik's first level criterion with respect to \mathcal{F} .

When there is no ambiguity, we will omit the specification *with respect to* \mathcal{F} in this definition. We will see (proposition 2.17) that Vershik's first level criterion for a random variable X is equivalent to Vershik's first level criterion for the σ -field $\sigma(X)$. It is clear that Vershik's first level criterion is preserved by isomorphism. The following proposition is easily established from the definition; its proof is left to the reader.

Proposition 2.7. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration of local product type and let (E, ρ) be a separable metric space. The set of random variables $X \in L^1(\mathcal{F}_0; E)$ satisfying Vershik's first level criterion is closed in $L^1(\mathcal{F}_0; E)$.*

To establish other properties of Vershik's first level criterion, it will be convenient to rephrase it with the help of the following notion:

Definition 2.8. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $\mathcal{B} \subset \mathcal{A}$ be a σ -field. A family \mathcal{C} of sub- σ -fields of \mathcal{A} is substantial in \mathcal{B} if the L^1 -closure of $\bigcup_{\mathcal{C} \in \mathcal{C}} L^1(\Omega, \mathcal{C}, \mathbb{P})$ contains $L^1(\Omega, \mathcal{B}, \mathbb{P})$.*

Thus, we can restate definition 2.6 of Vershik's first level criterion for a σ -field as follows.

Definition 2.9. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration of local product type. Call \mathcal{C}^{loc} the family of all σ -fields $\sigma(V_{n_0+1}, \dots, V_0)$ generated by local innovations (V_{n_0+1}, \dots, V_0) from n_0 to 0, for all $n_0 < 0$. A σ -field $\mathcal{E}_0 \subset \mathcal{F}_0$ satisfies Vershik's first level criterion if \mathcal{C}^{loc} is substantial in \mathcal{E}_0 .*

The notion of substantial family of σ -fields appears in [13] in a slightly different form. Lemma below is a duplicate of Lemma 2 in [13], which could be proved identically in spite of this difference between the two notions of substantialness.

Lemma 2.10. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $\mathcal{B} \subset \mathcal{A}$ a σ -field, and \mathcal{C} a family of sub- σ -fields of \mathcal{A} . The following three conditions are equivalent:*

- (i) \mathcal{C} is substantial in \mathcal{B} ;
- (ii) for each finite set F , the closure of $\bigcup_{\mathcal{C} \in \mathcal{C}} L(\mathcal{C}; F)$ in $L(\mathcal{A}; F)$ contains $L(\mathcal{B}; F)$;

(iii) for each separable metric space E , the closure of $\bigcup_{\mathcal{C} \in \mathcal{C}} L^1(\mathcal{C}; E)$ in $L^1(\mathcal{A}; E)$ contains $L^1(\mathcal{B}; E)$.

Proposition 2.11. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration of local product type. If a σ -field $\mathcal{E}_0 \subset \mathcal{F}_0$ satisfies Vershik's first level criterion, then for any separable metric space E , every random variable $X \in L^1(\mathcal{E}_0; E)$ satisfies Vershik's first level criterion.*

Proof. Left to the reader as an easy application of the definitions and the previous lemma. \square

The following lemma provides a typical example of substantialness.

Lemma 2.12. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(\mathcal{B}_m)_{m \in \mathbb{N}}$ an increasing sequence of sub- σ -fields of \mathcal{A} . Then the family of σ -fields $\{\mathcal{B}_m; m \in \mathbb{N}\}$ is substantial in $\bigvee_m \mathcal{B}_m$.*

Proof. A classical result says that for any set $B \in \bigvee_m \mathcal{B}_m$, there exist some $B_1, B_2, \dots \in \bigcup_m \mathcal{B}_m$ such that $\mathbb{P}[B \Delta B_m] \rightarrow 0$ (this is easily established by a monotone class argument). With the help of this, the lemma follows from lemma 2.10.(ii). \square

Proposition 2.13. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration of local product type and $(\mathcal{B}_m)_{m \geq 0}$ an increasing sequence of sub- σ -fields of \mathcal{F}_0 . If each \mathcal{B}_m satisfies Vershik's first level criterion, then so does $\bigvee_m \mathcal{B}_m$.*

Proof. Straightforward from the previous lemma and proposition 2.7. \square

Corollary 2.14. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration of local product type and $(V_n)_{n \leq 0}$ an innovation of \mathcal{F} . The σ -field $\sigma(V_n; n \leq 0)$ satisfies Vershik's first level criterion. Consequently, a filtration of product type satisfies Vershik's first level criterion.*

Proof. Obviously, the σ -field $\mathcal{B}_m := \sigma(V_n; -m \leq n \leq 0)$ satisfies Vershik's first level criterion for every $m \in \mathbb{N}$. Hence, the result derives from proposition 2.13. \square

This corollary contains the easy implications of the equivalences stated in theorems 2.23 and 2.25 towards which we orient the rest of this subsection.

Lemma 2.16 is the key lemma to prove the equivalence between Vershik's first level criterion for a random variable X and Vershik's first level criterion for the σ -field $\sigma(X)$. It characterizes substantialness of a family of σ -fields in an essentially separable σ -field \mathcal{B} by a property on a random variable X generating \mathcal{B} . It will be proved with the help of the following lemma, which we shall also use several times in the next sections.

Lemma 2.15. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $X \in L^1(\mathcal{A}; E)$ where (E, ρ) 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. Let us denote by $\ell(X)$ this set of random variables. For every open set $O \subset \mathbb{R}$, the sequence of random variables $X_m := (m\rho(X, O^c)) \wedge 1$ converges almost surely to $\mathbb{1}_{X \in O}$, and $x \mapsto (m\rho(x, O^c)) \wedge 1$ is a Lipschitz function. It follows that the L^1 -closure of $\ell(X)$ contains all linear combinations of indicator random variables $\mathbb{1}_{\{X \in O_i\}}$ where O_i is an open set, and therefore is dense in $L^1(\sigma(X))$. \square

Lemma 2.16. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, (E, ρ) a separable metric space, $X \in L^1(\mathcal{A}; E)$, and \mathbf{C} a family of sub- σ -fields of \mathcal{A} . Then \mathbf{C} is substantial in $\sigma(X)$ if and only if for every $\delta > 0$, there exist $\mathcal{C} \in \mathbf{C}$ and a random variable $C \in L^1(\mathcal{C}; E)$ such that $\mathbb{E}[\rho(X, C)] < \delta$.*

Proof. This easily results from lemma 2.10.(iii) and lemma 2.15. \square

Proposition 2.17. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration of local product type. Let (E, ρ) be a separable metric space and $X \in L^0(\mathcal{F}_0; E)$. The following conditions are equivalent.*

- (i) *the σ -field $\sigma(X)$ satisfies Vershik's first level criterion;*
- (ii) *for every $\delta > 0$, there exist an integer $n_0 < 0$, an innovation (V_{n_0+1}, \dots, V_0) of \mathcal{F} from n_0 to 0, and a random variable $S \in L^0(\sigma(V_{n_0+1}, \dots, V_0); E)$ such that $\mathbb{P}[\rho(X, S) > \delta] < \delta$.*

If $X \in L^1(\mathcal{F}_0; E)$, these conditions are also equivalent to

- (iii) *X satisfies Vershik's first level criterion.*

Proof. We know (definition 2.9) that (i) is equivalent to \mathbf{C}^{loc} being substantial in $\sigma(X)$. Thus, lemma 2.16 directly shows (i) \iff (iii), and (i) \iff (ii) derives from the same lemma by replacing ρ with $\rho \wedge 1$. \square

Most of the results in the sequel of this subsection will be established with the help of the following elementary lemmas 2.18 and 2.19. Lemma 2.18 is a duplicate of Lemma 3 in [13], which could be proved identically in spite of the difference between our notion of substantialness and the one given in [13].

Lemma 2.18. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, \mathcal{B} and \mathcal{D} two sub- σ -fields of \mathcal{A} , and \mathbf{C} a family of sub- σ -fields of \mathcal{A} . If \mathbf{C} is substantial in \mathcal{B} , then the family of σ -fields $\{\mathcal{C} \vee \mathcal{D} \mid \mathcal{C} \in \mathbf{C}\}$ is substantial in $\mathcal{B} \vee \mathcal{D}$.*

Lemma 2.19. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $\mathcal{E} \subset \mathcal{A}$ a σ -field, \mathbf{C} a family of sub- σ -fields of \mathcal{E} , and \mathbf{D} a family of sub- σ -fields of \mathcal{A} such that each $\mathcal{D} \in \mathbf{D}$ is independent of \mathcal{E} . If the family of σ -fields $\{\mathcal{C} \vee \mathcal{D} \mid \mathcal{C} \in \mathbf{C}, \mathcal{D} \in \mathbf{D}\}$ is substantial in a σ -field $\mathcal{B} \subset \mathcal{E}$, then \mathbf{C} is substantial in \mathcal{B} .*

Proof. Let $X \in L^1(\mathcal{B})$ and $\delta > 0$. Assuming that $\{\mathcal{C} \vee \mathcal{D} \mid \mathcal{C} \in \mathbf{C}, \mathcal{D} \in \mathbf{D}\}$ is substantial in \mathcal{B} , there exist $\mathcal{C} \in \mathbf{C}$, $\mathcal{D} \in \mathbf{D}$, and a $\mathcal{C} \vee \mathcal{D}$ -measurable random variable S such that $\mathbb{E}[|X - S|] < \delta$. One can write (lemma 1.1) $S = f(C, D)$ where C and D are random variables measurable with respect to \mathcal{C} and \mathcal{D}

respectively and f is measurable, thus $\mathbb{E}[S | \mathcal{E}]$ is measurable with respect to \mathcal{C} because $\mathbb{E}[S | \mathcal{E}] = h(C)$ where $h(c) = \mathbb{E}[f(c, D)]$. Using the L^1 -contractivity of conditional expectations, we get $\mathbb{E}[|X - \mathbb{E}[S | \mathcal{C}]|] = \mathbb{E}[|\mathbb{E}(X - S | \mathcal{E})|] < \mathbb{E}[|X - S|] < \delta$, which shows that \mathcal{C} is substantial in \mathcal{B} . \square

Proposition 2.21 highlights the asymptotic nature of Vershik's first level criterion. It is proved with the help of the following lemma.

Lemma 2.20. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration of local product type and $N \leq 0$ an integer. If a σ -field $\mathcal{E}_N \subset \mathcal{F}_N$ satisfies Vershik's first level criterion with respect to \mathcal{F} , then it satisfies Vershik's first level criterion with respect to the truncated filtration $(\mathcal{F}_{N+n})_{n \leq 0}$.*

Proof. This results from lemma 2.19 by taking $\mathcal{E} = \mathcal{F}_N$, $\mathcal{B} = \mathcal{E}_N$, and by considering the family \mathcal{C} consisting of all the σ -fields $\sigma(V_{n_0+1}, \dots, V_N)$ generated by local innovations (V_{n_0+1}, \dots, V_N) of \mathcal{F} from n_0 to N for all $n_0 < N$, and the family \mathcal{D} consisting of all the σ -fields $\sigma(V_{N+1}, \dots, V_0)$ generated by local innovations (V_{N+1}, \dots, V_0) of \mathcal{F} from N to 0. \square

Proposition 2.21. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration of local product type. The following conditions are equivalent:*

- (i) \mathcal{F} satisfies Vershik's first level criterion;
- (ii) for every $N \in -\mathbb{N}$, the truncated filtration $(\mathcal{F}_{N+n})_{n \leq 0}$ satisfies Vershik's first level criterion;
- (iii) there exists $N \in -\mathbb{N}$ such that the truncated filtration $(\mathcal{F}_{N+n})_{n \leq 0}$ satisfies Vershik's first level criterion.

Proof. Let \mathcal{C} be the family of σ -fields $\sigma(V_{n_0+1}, \dots, V_N)$ generated by all local innovations (V_{n_0+1}, \dots, V_N) from n_0 to N for all $n_0 < N$, let $\mathcal{D} = \sigma(V_{N+1}, \dots, V_0)$ where (V_{N+1}, \dots, V_0) is an innovation of \mathcal{F} from N to 0, and let $\mathcal{B} = \mathcal{F}_N$. Lemma 2.18 applied with these notations shows that (iii) \implies (i); lemma 2.20 shows that (i) \implies (ii); finally, (ii) \implies (iii) is trivially true. \square

The following proposition will help in subsection 2.3.

Proposition 2.22. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration of local product type, $(V_n)_{n \leq 0}$ an innovation of \mathcal{F} , and $(\mathcal{C}_n)_{n \leq 0}$ a sequence of σ -fields such that $\mathcal{C}_n \subset \mathcal{F}_n$ for each $n \leq 0$ and such that $(\mathcal{C}_n \vee \sigma(V_{n+1}, \dots, V_0))_{n \leq 0}$ is an increasing sequence of σ -fields. Define $\mathcal{C}_\infty = \bigvee_n \mathcal{C}_n$. If \mathcal{C}_n satisfies Vershik's first level criterion for every $n \leq 0$, then $\mathcal{C}_\infty \vee \sigma(V_n; n \leq 0)$ satisfies Vershik's first level criterion.*

Proof. Thanks to proposition 2.13, it suffices to show that each σ -field $\mathcal{C}_n \vee \sigma(V_{n+1}, \dots, V_0)$ satisfies Vershik's first level criterion. Let \mathcal{C} be the family consisting of all the σ -fields $\sigma(V'_{n_0+1}, \dots, V'_n)$ generated by some local innovation $(V'_{n_0+1}, \dots, V'_n)$ of \mathcal{F} from n_0 to n for some $n_0 < n$. By lemma 2.20, \mathcal{C} is substantial in \mathcal{C}_n . Let $\mathcal{D} = \sigma(V_{n+1}, \dots, V_0)$. Then apply lemma 2.18. \square

Theorem 2.23. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration of local product type on $(\Omega, \mathcal{A}, \mathbb{P})$. Then \mathcal{F} satisfies Vershik's first level criterion if and only if for every separable metric space E and every random variable $X \in L^1(\mathcal{F}_0; E)$, there exists an innovation $(V_n)_{n \leq 0}$ of \mathcal{F} such that X is measurable with respect to $\sigma(V_n; n \leq 0)$.*

Proof. The ‘if’ part follows from corollary 2.14 and proposition 2.11. The ‘only if’ part is proved as follows. Assume that \mathcal{F} satisfies Vershik's first level criterion and let $X \in L^1(\mathcal{F}_0; E)$. Consider a random variable $Y \in L^1(\mathcal{F}_0; \mathbb{R})$ such that $\sigma(X) = \sigma(Y)$. Let $(\delta_k)_{k \leq 0}$ be a sequence of positive numbers such that $\delta_k \rightarrow 0$.

Consider the following construction at rank k : we have an integer $n_k < 0$, an innovation $(V_n; n_k < n \leq 0)$ of \mathcal{F} from n_k to 0 and a random variable $X_k \in L^1(\sigma(V_n; n_k < n \leq 0))$ such that $\mathbb{E}[|Y - X_k|] < \delta_k$. We firstly apply the Vershik first level property of Y with respect to \mathcal{F} to obtain this construction for $k = 0$. When the construction is performed at rank k , we perform it at rank $k - 1$ by exhibiting an innovation $(V_{n_{k-1}+1}, \dots, V_{n_k})$ from an integer $n_{k-1} < n_k$ to n_k , and a random variable $X_{k-1} \in L^1(\sigma(V_n; n_{k-1} < n \leq 0))$ such that $\mathbb{E}[|Y - X_{k-1}|] < \delta_{k-1}$. To do so, we apply proposition 2.21 to get Vershik's first level criterion of the truncated filtration $(\mathcal{F}_{n_k+n})_{n \leq 0}$ and then we use the fact, due to lemma 2.18, that the family of σ -fields of the form $\sigma(V_{m+1}, \dots, V_{n_k}, V_{n_k+1}, \dots, V_0)$ where $(V_{m+1}, \dots, V_{n_k})$ is an innovation of \mathcal{F} from some $m < n_k$ to n_k , is substantial in \mathcal{F}_0 .

Continuing so, we obtain a global innovation $(V_n)_{n \leq 0}$ of \mathcal{F} and a sequence of random variables $(X_k)_{k \leq 0}$ in $L^1(\sigma(V_n; n \leq 0))$ converging in L^1 to Y . \square

Remark 2.24. We do not know if theorem 2.23 is true “random variable by random variable”. More precisely, we do not know if each random variable satisfying Vershik's first level criterion is measurable with respect to the σ -field generated by some global innovation of the filtration.

Theorem 2.25 (Vershik's first level criterion). *Let \mathcal{F} be an essentially separable filtration of local product type. Then \mathcal{F} satisfies Vershik's first level criterion if and only if \mathcal{F} is of product type.*

Proof. The ‘if’ part is given in corollary 2.14. To show the converse, apply theorem 2.23 with a random variable X generating \mathcal{F}_0 . This yields a global innovation $(V_n)_{n \leq 0}$ such that $\mathcal{F}_0 = \sigma(V_n; n \leq 0)$. Therefore $\mathcal{F}_n \cap \sigma(V_n; n \leq 0) = \mathcal{F}_n$ for every $n \leq 0$. Consequently $\mathcal{F}_n = \sigma(V_m; m \leq n)$ for every $n \leq 0$ because the filtration generated by $(V_n)_{n \leq 0}$ is immersed in \mathcal{F} (lemma 2.5). \square

Obviously, a filtration of local product type which is not essentially separable cannot be of product type. However we will see in section 3 that it is possible that such a filtration satisfies Vershik's first level criterion (corollary 2.46). It is then interesting to notice that a filtration satisfying Vershik's first level criterion is Kolmogorovian, even if it is not of product type.

Corollary 2.26. *A filtration of local product type satisfying Vershik's first level criterion is Kolmogorovian.*

Proof. Let \mathcal{F} be such a filtration and $A \in \mathcal{F}_{-\infty}$. Thanks to theorem 2.23, there exists a global innovation $(V_n)_{n \leq 0}$ such that $\mathbb{1}_A$ is measurable with respect to $\sigma(V_n; n \leq 0)$. As the filtration generated by $(V_n)_{n \leq 0}$ is immersed in \mathcal{F} (lemma 2.5), it follows that $\mathbb{1}_A$ is measurable with respect to the trivial σ -field $\cap_n \sigma(V_m; m \leq n)$. \square

Below is another corollary of theorem 2.23 which will be used in section 3 to prove a result on weak standardness (proposition 3.23).

Corollary 2.27. *Let \mathcal{F} and \mathcal{G} be two independent filtrations of local product type satisfying Vershik's first level criterion. Then $\mathcal{F} \vee \mathcal{G}$ is a filtration of local product type satisfying Vershik's first level criterion.*

Proof. Let $R \in L^1(\mathcal{F}_0 \vee \mathcal{G}_0; [0, 1])$. We can write $R = f(X, Y)$ where X and Y are random variables measurable with respect to \mathcal{F}_0 and \mathcal{G}_0 respectively and f is a Borelian function. By theorem 2.23, there exist a global innovation $(V_n)_{n \leq 0}$ of \mathcal{F} and a global innovation $(W_n)_{n \leq 0}$ of \mathcal{G} such that X is measurable with respect to $\sigma(V_n; n \leq 0)$ and Y is measurable with respect to $\sigma(W_n; n \leq 0)$. Setting $Z_n = (V_n, W_n)$, then $(Z_n)_{n \leq 0}$ obviously is an innovation of $\mathcal{F} \vee \mathcal{G}$ and R is measurable with respect to $\sigma(Z_n; n \leq 0)$. Thus $\mathcal{F} \vee \mathcal{G}$ satisfies Vershik's first level criterion due to theorem 2.23. \square

The converse of corollary 2.27 holds true; actually we could prove the following stronger result, but we will not need it: *Let \mathcal{F} and \mathcal{G} be two independent filtrations of local product type. Let E_1 and E_2 be Polish spaces, $X \in L^1(\mathcal{F}_0; E_1)$ and $Y \in L^1(\mathcal{G}_0; E_2)$. Then X and Y satisfy Vershik's first level criterion with respect to \mathcal{F} and \mathcal{G} respectively, if and only if (X, Y) satisfies Vershik's first level criterion with respect to $\mathcal{F} \vee \mathcal{G}$.*

Remark 2.28. If \mathcal{F} is a filtration of local product type generated by a martingale $(M_n)_{n \leq 0}$, then it is possible to show that \mathcal{F} satisfies Vershik's first level criterion if and only if the random variable M_0 satisfies Vershik's first level criterion. We will not show this fact as we shall see in the next subsection that Vershik's first level criterion is equivalent to *Vershik's self-joining criterion*, and a result in [21] says that the same fact holds for Vershik's self-joining criterion. Actually this result says that the same fact holds more generally for any *self-joining criterion*, a notion defined in [21] that includes Vershik's self-joining criterion and the I-cosiness criterion as particular cases (see also remark 3.44 and the first paragraph of section 3.4).

2.2 Rosenblatt's and Vershik's self-joining criteria

Given a filtration \mathcal{F} of local product type, proposition 2.33 gives a “self-joining criterion” for a global innovation of \mathcal{F} to be generating (*Rosenblatt's self-joining criterion*), and theorem 2.38 (*Vershik's self-joining criterion*) gives

a “self-joining criterion” for \mathcal{F} to satisfy Vershik’s first level criterion. More precisely, these criteria are stated “random variable by random variable”. The terminologies are discussed at the end of the subsection. Both these criteria are a particular form of the I-cosiness criterion (definition 3.29). Rosenblatt’s self-joining criterion will be illustrated by the example given in section 3.1. Vershik’s self-joining criterion will be used to establish theorem A” stated in the introduction (and restated in theorem 4.4).

Joinings

As also does the I-cosiness criterion, Rosenblatt’s self-joining criterion and Vershik’s self-joining criterion involve *joinings* of filtrations, defined below.

Definition 2.29. *Let \mathcal{F} be a filtration.*

1. *A joining of \mathcal{F} is a pair $(\mathcal{F}', \mathcal{F}'')$ of two filtrations \mathcal{F}' and \mathcal{F}'' defined on the same probability space which are both isomorphic to \mathcal{F} and jointly immersed, that is, \mathcal{F}' and \mathcal{F}'' are both immersed in $\mathcal{F}' \vee \mathcal{F}''$ (or, equivalently, in a same filtration).*
2. *A joining $(\mathcal{F}', \mathcal{F}'')$ of \mathcal{F} is independent in small time if the σ -fields \mathcal{F}'_{n_0} and \mathcal{F}''_{n_0} are independent for some integer $n_0 \leq 0$. We also say that $(\mathcal{F}', \mathcal{F}'')$ is a joining of \mathcal{F} independent up to n_0 to specify this integer.*

A typical example of joining $(\mathcal{F}', \mathcal{F}'')$ is the case where \mathcal{F}' and \mathcal{F}'' are two independent copies of \mathcal{F} (lemma 1.2).

Rigorously, a joining is the 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. Considering a joining $(\mathcal{F}', \mathcal{F}'')$ of a filtration $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$, and given a \mathcal{F}_0 -measurable random variable X valued in a Polish space, we will traditionally denote by X' and X'' the respective copies of X given by the two underlying embeddings Ψ' and Ψ'' . Of course, Y' and Y'' will denote the copies of a \mathcal{F}_0 -measurable random variable Y , and so on. In the same way, the two copies of a σ -field $\mathcal{B} \subset \mathcal{F}_0$ will be respectively denoted by \mathcal{B}' and \mathcal{B}'' , and the two copies of a filtration $\mathcal{E} \subset \mathcal{F}$ will be respectively denoted by \mathcal{E}' and \mathcal{E}'' . Note that, 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} .

We shall need the following lemma in the proof of theorem 2.38. Its easy proof is left to the reader.

Lemma 2.30. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration and $n_0 \leq 0$ an integer. Let $(\mathcal{F}', \mathcal{F}'')$ be a joining of \mathcal{F} independent up to n_0 . Then the σ -fields \mathcal{F}'_0 and \mathcal{F}''_{n_0} are independent.*

The next lemma is obvious from the definitions; it will be used to construct joinings of a filtration of local product type with the help of innovations.

Lemma 2.31. *On $(\Omega, \mathcal{A}, \mathbb{P})$, let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration and V_0 be an independent complement of \mathcal{F}_{-1} in \mathcal{F}_0 . On $(\overline{\Omega}, \overline{\mathcal{A}}, \overline{\mathbb{P}})$, let $(\mathcal{F}'_n)_{n \leq -1}$ and $(\mathcal{F}''_n)_{n \leq -1}$ be two jointly immersed isomorphic copies of $(\mathcal{F}_n)_{n \leq -1}$, given by two isomorphisms $\Psi'_{-1}: \mathcal{F}_{-1} \rightarrow \mathcal{F}'_{-1}$ and $\Psi''_{-1}: \mathcal{F}_{-1} \rightarrow \mathcal{F}''_{-1}$, and let V'_0 and V''_0 be two random variables each having the same law as V_0 and independent of $\mathcal{F}'_{-1} \vee \mathcal{F}''_{-1}$. We put $\mathcal{F}'_0 = \mathcal{F}'_{-1} \vee \sigma(V'_0)$ and $\mathcal{F}''_0 = \mathcal{F}''_{-1} \vee \sigma(V''_0)$. Then $(\mathcal{F}', \mathcal{F}'')$ is a joining of \mathcal{F} , given by two unique isomorphisms Ψ' and Ψ'' that respectively extend Ψ'_{-1} and Ψ''_{-1} and respectively send V_0 to V'_0 and V''_0 .*

Proof. The two isomorphisms Ψ' and Ψ'' are given by corollary A.12. If $(\mathcal{F}'_n)_{n \leq -1}$ is immersed in $(\mathcal{F}'_n \vee \mathcal{F}''_n)_{n \leq -1}$, then we can see by lemma 1.4 that \mathcal{F}' is immersed in $\mathcal{F}' \vee \mathcal{F}''$ if and only if \mathcal{F}'_0 and \mathcal{F}'_{-1} are conditionally independent given $\mathcal{F}'_{-1} \vee \mathcal{F}''_{-1}$. By lemma 1.5, this conditional independence holds if (and only if) V'_0 is independent of $\mathcal{F}'_{-1} \vee \mathcal{F}''_{-1}$. Thus \mathcal{F}' is immersed in $\mathcal{F}' \vee \mathcal{F}''$ and the same fact obviously holds for \mathcal{F}'' . \square

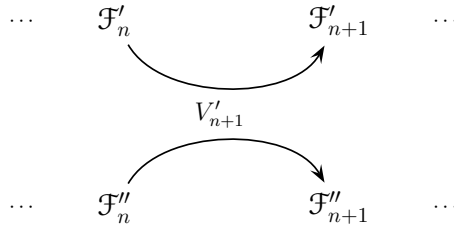
Rosenblatt's self-joining criterion (for generatingness)

Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration of local product type and $(V_n)_{n \leq 0}$ a global innovation of \mathcal{F} . Proposition 2.33 below gives a necessary and sufficient condition for a random variable to be measurable with respect to the σ -field generated by $(V_n)_{n \leq 0}$. The proposition involves joinings $(\mathcal{F}', \mathcal{F}'')$ constructed as follows.

Given an integer $n_0 < 0$, we consider, on some probability space $(\overline{\Omega}, \overline{\mathcal{A}}, \overline{\mathbb{P}})$, two copies $(\mathcal{F}'_n)_{n \leq n_0}$ and $(\mathcal{F}''_n)_{n \leq n_0}$ of the filtration $(\mathcal{F}_n)_{n \leq n_0}$ and a random vector $(V'_{n_0+1}, \dots, V'_0)$ having the same law as (V_{n_0+1}, \dots, V_0) and independent of $\mathcal{F}'_{n_0} \vee \mathcal{F}''_{n_0}$. We complete the filtrations $(\mathcal{F}'_n)_{n \leq n_0}$ and $(\mathcal{F}''_n)_{n \leq n_0}$ up to time 0 by putting

$$\mathcal{F}'_n = \mathcal{F}'_{n_0} \vee \sigma(V'_{n_0+1}, \dots, V'_n) \quad \text{and} \quad \mathcal{F}''_n = \mathcal{F}''_{n_0} \vee \sigma(V'_{n_0+1}, \dots, V'_n)$$

for each $n \in \{n_0+1, \dots, 0\}$. Assuming that $(\mathcal{F}'_n)_{n \leq n_0}$ and $(\mathcal{F}''_n)_{n \leq n_0}$ are jointly immersed, one easily checks with the help of lemma 2.31 that $(\mathcal{F}', \mathcal{F}'')$ is a joining of \mathcal{F} , given by two isomorphisms $\Psi': \mathcal{F} \rightarrow \mathcal{F}'$ and $\Psi'': \mathcal{F} \rightarrow \mathcal{F}''$, both of them sending V_n to V'_n for each $n \in \{n_0+1, \dots, 0\}$. For a time $n \geq n_0$, the joining can be pictured as follows:



Such a joining $(\mathcal{F}', \mathcal{F}'')$ is characterized by the fact that $V'_n = V''_n$ for every $n \in \{n_0+1, \dots, 0\}$.

Definition 2.32 (Rosenblatt's self-joining criterion). Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration of local product type and $(V_n)_{n \leq 0}$ an innovation of \mathcal{F} . Let (E, ρ) be a Polish metric space and $X \in L^1(\mathcal{F}_0; E)$. We say that X satisfies Rosenblatt's self-joining criterion with $(V_n)_{n \leq 0}$ if for each real number $\delta > 0$, there exists, on some probability space $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$, a joining $(\mathcal{F}', \mathcal{F}'')$ of \mathcal{F} independent up to an integer $n_0 \leq 0$, such that $V'_n = V''_n$ for every $n \in \{n_0 + 1, \dots, 0\}$, and for which one has $\bar{\mathbb{E}}[\rho(X', X'')] < \delta$, where X' and X'' are the respective copies of X in \mathcal{F}' and in \mathcal{F}'' .

Proposition 2.33 (Rosenblatt's self-joining criterion). Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration of local product type and $(V_n)_{n \leq 0}$ an innovation of \mathcal{F} . Let (E, ρ) be a Polish metric space and $X \in L^1(\mathcal{F}_0; E)$. Then X is measurable with respect to $\sigma(V_n; n \leq 0)$ if and only if X satisfies Rosenblatt's self-joining criterion with $(V_n)_{n \leq 0}$.

Proof. If X is measurable with respect to $\sigma(V_n; n \leq 0)$ then, by lemma 2.12 and lemma 2.10, for any $\delta > 0$ there exist an integer $n_0 \leq 0$ and a random variable $S \in L^1(\sigma(V_{n_0+1}, \dots, V_0); E)$ such that $\mathbb{E}[\rho(X, S)] < \delta/2$. Considering a joining $(\mathcal{F}', \mathcal{F}'')$ of \mathcal{F} independent up to n_0 as defined in the proposition, one has $\bar{\mathbb{E}}[\rho(X', S')] = \bar{\mathbb{E}}[\rho(X'', S'')] = \mathbb{E}[\rho(X, S)]$ due to isomorphisms, and $S' = S''$ because S is measurable with respect to $\sigma(V_{n_0+1}, \dots, V_0)$. This gives $\bar{\mathbb{E}}[\rho(X', X'')] < \delta$ by the triangular inequality, thereby showing that X satisfies Rosenblatt's self-joining criterion.

Conversely, assume that Rosenblatt's self-joining criterion holds for X . Then, considering two independent copies \mathcal{F}' and \mathcal{F}^* of \mathcal{F} , we see that the family of σ -fields $\{\mathcal{F}_0^* \vee \sigma(V'_m; n < m \leq 0) \mid n < 0\}$ is substantial in $\sigma(X')$ (definition 2.8 and lemma 2.16). As \mathcal{F}_0^* is independent of \mathcal{F}'_0 , lemma 2.19 shows that the family of σ -fields $\{\sigma(V'_m; n < m \leq 0) \mid n < 0\}$ is substantial in $\sigma(X')$. Consequently X' is measurable with respect to $\sigma(V'_n; n \leq 0)$, and therefore, due to isomorphism, X is measurable with respect to $\sigma(V_n; n \leq 0)$. \square

Corollary 2.34. Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be an essentially separable filtration of local product type and $(V_n)_{n \leq 0}$ an innovation of \mathcal{F} . Then the following conditions are equivalent:

- (i) \mathcal{F} is the filtration of product type generated by $(V_n)_{n \leq 0}$;
- (ii) for every Polish space E , every random variable $X \in L^1(\mathcal{F}_0; E)$ satisfies Rosenblatt's self-joining criterion with $(V_n)_{n \leq 0}$;
- (iii) for some Polish space E , there exists a random variable $X \in L^1(\mathcal{F}_0; E)$ generating \mathcal{F}_0 and satisfying Rosenblatt's self-joining criterion with $(V_n)_{n \leq 0}$;
- (iv) every random variable $X \in L^1(\mathcal{F}_0; \mathbb{R})$ satisfy Rosenblatt's self-joining criterion with $(V_n)_{n \leq 0}$.

Proof. The previous proposition shows that (i) \implies (ii), (iii), (iv), and that each of (ii), (iii) and (iv) implies that $\mathcal{F}_0 = \sigma(V_n; n \leq 0)$. As the filtration generated by $(V_n)_{n \leq 0}$ is immersed in \mathcal{F} (lemma 2.5), this yields $\sigma(V_m; m \leq n) = \mathcal{F}_n$ for each $n \leq 0$. \square

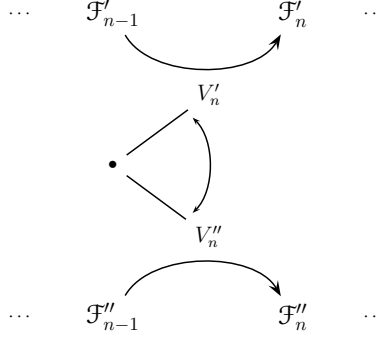
Vershik's self-joining criterion

Vershik's self-joining criterion is the “self-joining version” of Vershik's first level criterion (definition 2.6). Given a filtration \mathcal{F} of local product type and an innovation $(V_n)_{n \leq 0}$ of \mathcal{F} , its statement involves the joinings $(\mathcal{F}', \mathcal{F}'')$ of \mathcal{F} defined as follows and including as particular cases the joinings involved in Rosenblatt's self-joining criterion.

Given an integer $n_0 \leq 0$, we consider, on some probability space $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$, two copies $(\mathcal{F}'_n)_{n \leq n_0}$ and $(\mathcal{F}''_n)_{n \leq n_0}$ of the filtration $(\mathcal{F}_n)_{n \leq n_0}$ and a random vector $(V'_{n_0+1}, \dots, V'_0)$ having the same law as (V_{n_0+1}, \dots, V_0) and independent of $\mathcal{F}'_{n_0} \vee \mathcal{F}''_{n_0}$. We complete the filtrations $(\mathcal{F}'_n)_{n \leq n_0}$ and $(\mathcal{F}''_n)_{n \leq n_0}$ up to time 0 by defining \mathcal{F}'_n and \mathcal{F}''_n for each $n \in \{n_0 + 1, \dots, 0\}$ by

$$\mathcal{F}'_n = \mathcal{F}'_{n_0} \vee \sigma(V'_{n_0+1}, \dots, V'_n) \quad \text{and} \quad \mathcal{F}''_n = \mathcal{F}''_{n_0} \vee \sigma(V''_{n_0+1}, \dots, V''_n),$$

where the random vector $(V''_{n_0+1}, \dots, V''_0)$ is constructed as follows. At each step $n \in \{n_0 + 1, \dots, 0\}$, calling E_n the Polish state space of V_n , we consider a measurable function $\phi^n: \mathbb{R} \times E_n \rightarrow E_n$ such that for each fixed $x \in \mathbb{R}$, the function $v \mapsto \phi^n(x, v)$ is a Lebesgue automorphism of the probability space induced by V_n , and then we put $V''_n = \phi^n(\bar{H}_{n-1}, V'_n)$ where \bar{H}_{n-1} is some random variable measurable with respect to $\mathcal{F}'_{n-1} \vee \mathcal{F}''_{n-1}$. This construction can be pictured as follows:



It is clear from this construction that V''_n has the same law as V'_n and that each of V'_n and V''_n is independent of $\mathcal{F}'_{n-1} \vee \mathcal{F}''_{n-1}$. Therefore, assuming that $(\mathcal{F}'_n)_{n \leq n_0}$ and $(\mathcal{F}''_n)_{n \leq n_0}$ are jointly immersed, lemma 2.31 ensures that $(\mathcal{F}', \mathcal{F}'')$ is a joining of \mathcal{F} , given by two isomorphisms $\Psi': \mathcal{F} \rightarrow \mathcal{F}'$ and $\Psi'': \mathcal{F} \rightarrow \mathcal{F}''$ that respectively send V_n to V'_n and V''_n for each $n \in \{n_0 + 1, \dots, 0\}$. Such joinings will be given a name in the following definition.

Definition 2.35. Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration of local product type and $(V_n)_{n \leq 0}$ an innovation of \mathcal{F} . A joining $(\mathcal{F}', \mathcal{F}'')$ of \mathcal{F} is permutational after n_0 for an integer $n_0 \leq 0$ if for each $n \in \{n_0 + 1, \dots, 0\}$ we have $V''_n = T_n(V'_n)$ where T_n is a random Lebesgue automorphism of the probability space induced by V_n , defined by $T_n(\cdot) = \phi^n(\bar{H}_{n-1}, \cdot)$ where ϕ^n is a measurable function and \bar{H}_{n-1} is some random variable measurable with respect to $\mathcal{F}'_{n-1} \vee \mathcal{F}''_{n-1}$.

The joinings featuring in Rosenblatt's self-joining criterion appear as the particular case when the T_n are almost surely equal to identity. Note that the definition does not depend on the choice of the innovation $(V_n)_{n \leq 0}$ in view of lemma 2.4. Actually we can see by this lemma that a joining $(\mathcal{F}', \mathcal{F}'')$ permutational after n_0 is characterized by the fact that $\mathcal{F}'_n \vee \mathcal{F}''_n = \mathcal{F}'_n \vee \mathcal{F}''_{n_0}$ for each $n \in \{n_0 + 1, \dots, 0\}$, and this characterization does not involve any innovation of \mathcal{F} .

The following easy lemma will be used in the proof of theorem 2.38 and lemma 4.2

Lemma 2.36. *In the context of the above definition, and, in addition, given an \mathcal{F}_0 -measurable random variable X , there exist two \mathcal{F}_{n_0} -measurable random variables C_{n_0} and D_{n_0} such that $\sigma(X) \subset \sigma(D_{n_0}, V_{n_0+1}, \dots, V_0)$ and it is possible to write*

$$T_n(\cdot) = \psi^n(C'_{n_0}, D''_{n_0}, V'_{n_0+1}, \dots, V'_{n-1}, \cdot)$$

for every $n \in \{n_0 + 1, \dots, 0\}$, where ψ^n is a measurable function.

Proof. The proof is a successive application of lemma 1.1. We use the notations of definition 2.35, thus one has $V''_n = T_n(V'_n)$ where the random transformations T_n are written in form $T_n(\cdot) = \phi^n(\bar{H}_{n-1}, \cdot)$. To show the lemma, it suffices to find C_{n_0} and D_{n_0} such that $\sigma(X'') \subset \sigma(D''_{n_0}, V''_{n_0+1}, \dots, V''_0)$ and $\sigma(\bar{H}_{n-1}) \subset \sigma(C'_{n_0}, D''_{n_0}, V'_{n_0+1}, \dots, V'_{n-1})$ for each $n \in \{n_0 + 1, \dots, 0\}$. Note that $(V'_{n_0+1}, \dots, V'_0)$ is a local innovation of $\mathcal{F}' \vee \mathcal{F}''$ from n_0 to 0.

For each $n \in \{n_0 + 1, \dots, 0\}$, take (lemma 1.1) an $(\mathcal{F}'_{n_0} \vee \mathcal{F}''_{n_0})$ -measurable r.v. $\bar{S}^n_{n_0}$ such that $\sigma(\bar{H}_{n-1}) \subset \sigma(\bar{S}^n_{n_0}, V'_{n_0+1}, \dots, V'_{n-1})$. Then, take a r.v. \bar{R}_{n_0} such that $\sigma(\bar{R}_{n_0}) = \sigma(\bar{S}^n_{n_0}; n \in \{n_0 + 1, \dots, 0\})$, so that one has $\sigma(\bar{H}_{n-1}) \subset \sigma(\bar{R}_{n_0}, V'_{n_0+1}, \dots, V'_{n-1})$ for each $n \in \{n_0 + 1, \dots, 0\}$. Then, take (lemma 1.1) an \mathcal{F}'_{n_0} -measurable r.v. C'_{n_0} and an \mathcal{F}''_{n_0} -measurable r.v. B''_{n_0} such that $\sigma(\bar{R}_{n_0}) \subset \sigma(C'_{n_0}, B''_{n_0})$, and, finally, take (lemma 1.1) an \mathcal{F}''_{n_0} -measurable r.v. X''_{n_0} such that $\sigma(X'') \subset \sigma(X''_{n_0}, V''_{n_0+1}, \dots, V''_0)$ and take a r.v. D''_{n_0} such that $\sigma(D''_{n_0}) = \sigma(B''_{n_0}, X''_{n_0})$. \square

Definition 2.37. Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration of local product type, (E, ρ) a Polish metric space, and let $X \in L^1(\mathcal{F}_0; E)$. We say that X satisfies Vershik's self-joining criterion if for each real number $\delta > 0$, there exist an integer $n_0 \leq 0$ and, on some probability space $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$, a joining $(\mathcal{F}', \mathcal{F}'')$ of \mathcal{F} independent up to n_0 and permutational after n_0 such that one has $\bar{\mathbb{E}}[\rho(X', X'')] < \delta$, where X' and X'' are the respective copies of X in \mathcal{F}' and in \mathcal{F}'' .

Theorem 2.38. Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration of local product type. Let (E, ρ) be a Polish metric space and $X \in L^1(\mathcal{F}_0; E)$. Then X satisfies Vershik's self-joining criterion if and only if X satisfies Vershik's first level criterion.

Proof. The proof of the ‘only if’ part is similar the proof of the ‘only if’ part of proposition 2.33. Indeed, assuming that X satisfies Vershik’s first level criterion, then for any $\delta > 0$ there exist $n_0 \leq 0$, a local innovation $(\tilde{V}_{n_0+1}, \dots, \tilde{V}_0)$ from n_0 to 0, and a random variable $S \in L^1(\sigma(\tilde{V}_{n_0+1}, \dots, \tilde{V}_0); E)$ such that $\mathbb{E}[\rho(X, \tilde{S})] < \delta/2$. Thus Vershik’s self-joining criterion for X can be proved in the same way as we have proved Rosenblatt’s self-joining criterion for X in the first part of the proof of proposition 2.33, by replacing (V_{n_0+1}, \dots, V_0) with $(\tilde{V}_{n_0+1}, \dots, \tilde{V}_0)$.

Conversely, assume Vershik’s self-joining criterion holds for X . Fix $\delta > 0$, $n_0 \leq 0$ and $(\mathcal{F}', \mathcal{F}'')$ as in definition 2.37. By lemma 2.36, there are two \mathcal{F}_{n_0} -measurable random variables C_{n_0} and D_{n_0} such that $\sigma(X) \subset \sigma(D_{n_0}, V_{n_0+1}, \dots, V_0)$ and

$$T_n(\cdot) = \psi^n(C'_{n_0}, D''_{n_0}, V'_{n_0+1}, \dots, V'_{n-1}, \cdot)$$

for every $n \in \{n_0 + 1, \dots, 0\}$. By taking a Borelian function f such that $X = f(D_{n_0}, V_{n_0+1}, \dots, V_0)$, one has $X'' = h_{D''_{n_0}}(C'_{n_0}, V'_{n_0+1}, \dots, V'_0)$ where, for a given value y of D''_{n_0} ,

$$h_y(C'_{n_0}, V'_{n_0+1}, \dots, V'_0) = f(y, \tilde{V}_{n_0+1}, \dots, \tilde{V}_0),$$

with

$$\tilde{V}_n = \psi^n(C'_{n_0}, y, V'_{n_0+1}, \dots, V'_{n-1}, V'_n),$$

for each $n \in \{n_0 + 1, \dots, 0\}$, and thus \tilde{V}_n is an independent complement of \mathcal{F}'_{n-1} in \mathcal{F}'_n by lemma 2.4. In view of lemma 2.30, the random variable D''_{n_0} is independent of \mathcal{F}'_0 . The assumption that $\overline{\mathbb{E}}[\rho(X', X'')] < \delta$ therefore implies that for some y , we have $\overline{\mathbb{E}}[\rho(X', \tilde{S})] < \delta$ where $\tilde{S} = h_y(C'_{n_0}, V'_{n_0+1}, \dots, V'_0)$. Thus X' satisfies Vershik’s first level criterion with respect to \mathcal{F}' because the random variable \tilde{S} is measurable with respect to $\sigma(\tilde{V}_{n_0+1}, \dots, \tilde{V}_0)$. Obviously, Vershik’s first level criterion with respect to \mathcal{F} is satisfied for X due to isomorphism. \square

On the terminology

We have called the self-joining criterion of proposition 2.33 *Rosenblatt’s self-joining criterion* because this result was often used in the works of Rosenblatt ([26, 27, 28]) and their further developments ([17], [5]). The self-joining criterion in definition 2.37 is called *Vershik’s self-joining criterion* because it is close to the “combinatorial” standardness criterion stated in Vershik’s works in the particular case of homogeneous filtrations (defined in the introduction and in definition 4.1) with atomic innovations.

2.3 Example: split-word processes

We shall define the split-word processes. Their filtrations are known to be of product type or not according to some conditions on the parameters defining these processes. Admitting the key results found in the literature, we will start proving the main theorem on productness for these filtrations (theorem 2.39). At this stage we do not yet have at our disposal the tools needed to entirely prove this theorem; its proof will be continued at the end of each section.

To define a *split-word process*, the first ingredient is a probability space (A, \mathfrak{A}, μ) called the *alphabet*. A *word* on A is an element $w \in A^\ell$ for some integer $\ell \geq 1$, called the *length* of w , and $w(1), \dots, w(\ell)$ are the *letters* of w . Given any function f from A to a set B we naturally define $f(w) \in B^\ell$ as the word on B with letters $f(w(1)), \dots, f(w(\ell))$ in this order.

The second ingredient is the *splitting sequence* $(r_n)_{n \leq 0}$, consisting of integers $r_n \geq 2$. Given this sequence, we define the sequence $(\ell_n)_{n \leq 0}$ of *lengths* by $\ell_n = \prod_{k=n+1}^0 r_k$ for all $n \leq 0$; in other words, the sequence $(\ell_n)_{n \leq 0}$ is recursively defined by $\ell_0 = 1$ and $\ell_{n-1} = r_n \ell_n$.

Then we define the *split-word process* with alphabet (A, \mathfrak{A}, μ) and splitting sequence $(r_n)_{n \leq 0}$ to be the (non time-homogeneous) Markov process $(W_n, \eta_n)_{n \leq 0}$ whose law is characterized by the following two conditions:

- ◊ for each $n \leq 0$, W_n is a random word on A of length ℓ_n , whose letters are i.i.d. random variables with law μ , and η_n is independent of W_n and has the uniform law on the set $\{1, \dots, r_n\}$;
- ◊ the transition from $n-1$ to n is obtained by taking η_n independent of (W_{n-1}, η_{n-1}) , and then by choosing W_n as the η_n -th subword of W_{n-1} considered as a concatenation of r_n subwords of equal length ℓ_n .

We denote by \mathcal{F} the filtration generated by $(W_n, \eta_n)_{n \leq 0}$. Of course the process $(\eta_n)_{n \leq 0}$ is an innovation of \mathcal{F} . It can be shown that $\mathcal{F}_{-\infty}$ is degenerate whatever the alphabet (A, \mathfrak{A}, μ) and the splitting sequence $(r_n)_{n \leq 0}$; the proof is the same as in [14] where the particular case that $r_n \equiv 2$ and μ is uniform on a finite alphabet A is treated. We will derive the following theorem using results from the literature.

Theorem 2.39. *Call (Δ) the condition on the splitting sequence $(r_n)_{n \leq 0}$:*

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

Then:

- (a) *If (Δ) holds, then \mathcal{F} is not of product type unless (A, \mathfrak{A}, μ) is degenerate.*
- (b) *If (Δ) does not hold and if (A, \mathfrak{A}, μ) is Polish, then \mathcal{F} is of product type.*

Remark 2.40. As we have seen in this section, part (a) of theorem 2.39 means that under condition (Δ) , Vershik's first level criterion, or Vershik's self-joining criterion, fails to be true for some random variables. It would be interesting to have more information on those random variables. For example, does the final letter W_0 of the split-word process $(W_n, \eta_n)_{n \leq 0}$ never satisfy Vershik's first level criterion under condition (Δ) on the splitting sequence, whatever the non-trivial alphabet ?

Remark 2.41 (Vershik's example 1). The similar theorem holds for the so-called *example 1* in [41]. Part (a) of theorem 2.39 for Vershik's example 1 is shown in [41] with the help of its notion of entropy of filtrations. Part (b) is shown by Hecklen in [18] in the particular case when A is finite and μ is the uniform probability on A (thus Hecklen shows the analogue of Ceilier's result 2.44 for Vershik's example 1). This example deals with filtrations defined as sequences of the invariant σ -fields of the actions of a decreasing sequence of groups on a Lebesgue probability space. Unfortunately we have failed to check whether or not Vershik's example 1 is exactly the same as the filtrations of the split-word processes, but similar mathematics appear in the proof of the two theorems. In fact, we know that the following coincidence holds: Vershik's self-joining criterion of the random variable W_0 with respect to the filtration of the split-word process $(W_n, \eta_n)_{n \leq 0}$ can be expressed in a problem of purely combinatorial nature, and this problem is also equivalent to Vershik's self-joining criterion of a certain random variable W_0^* with respect to the corresponding filtration of Vershik's example 1¹. Other examples of filtrations for which Vershik's self-joining criterion of a certain random variable is equivalent to Vershik's self-joining criterion of W_0 , are given in [41].

Remark 2.42 (The scale of an automorphism). Theorem 2.39 is closely related to the *scale* of Bernoulli automorphisms. The notion of *scale of an automorphism* has been introduced by Vershik in [40]. The scale is a set of sequences of integers $(r_n)_{n \leq 0}$. Vershik asserts in [40], without giving a proof, that $(r_n)_{n \leq 0}$ does not belong to the scale of Bernoulli automorphisms under condition (Δ) , and he proves that it belongs to the scale of Bernoulli automorphisms under a stronger condition than $\neg(\Delta)$. We have checked, by using the first definition of the scale given by Vershik, that a sequence $(r_n)_{n \leq 0}$ belongs to the scale of a Bernoulli automorphism if and only if Vershik's self-joining criterion holds for a certain random variable with respect to the filtration \mathcal{F} of a split-word process with splitting sequence $(r_n)_{n \leq 0}$. Thus, part (b) of theorem 2.39 shows that $(r_n)_{n \leq 0}$ belongs to the scale of Bernoulli automorphisms under condition $\neg(\Delta)$, thereby improving the proposition in [40]. In [21], we describe this random variable and we show that Vershik's self-joining criterion of this random variable is actually equivalent to productness of \mathcal{F} . Finally theorem 2.39

¹ With the terminology of [21], the respective sequences of *Vershik's progressive predictions* $(\pi_n W_0)_{n \leq 0}$ and $(\pi_n W_0^*)_{n \leq 0}$ of W_0 and W_0^* are two processes with the same law; that shows that the Vershik property, or the I-cosiness, or Vershik's self-joining criterion, are the same for W_0 and W_0^* .

then shows that *the scale of a Bernoulli automorphism is the set of sequences $(r_n)_{n \leq 0}$ satisfying $\neg(\Delta)$* .

We shall derive theorem 2.39 from the following two facts, which will be admitted:

Result 2.43 *If (Δ) holds and A is finite, then \mathcal{F} is not of product type, unless μ is degenerate.*

If $r_n \equiv 2$, condition (Δ) holds; in this case, Smorodinsky ([32]) has shown that \mathcal{F} is not of product type. His proof is copied in [14] with the language of I-cosiness. Result 2.43 is shown in [20] with the help of Vershik's self-joining criterion (theorem 2.38) and by generalizing some lemmas given in [14] for the particular case $r_n \equiv 2$. Result 2.43 is also shown in [7] by means of the I-cosiness criterion, by a more direct generalization of the proof given in [14].

Result 2.44 *If A is finite and μ is the uniform probability on A , then \mathcal{F} is of product type if (Δ) does not hold.*

Result 2.44 is due to Ceillier ([7]). This is the most recent and from some point of view the most difficult part of theorem 2.39.

We do not yet have the material needed to prove theorem 2.39 from these two results. We will only give, in result 2.45 below, the first step towards the derivation of part (b) of the theorem from result 2.44. Result 2.43 will be discussed in section 3 where we shall demonstrate how to derive part (a) of theorem 2.39 from this result. Finally we shall demonstrate in section 4 how to derive part (b) of theorem 2.39 from result 2.45.

Result 2.45 *If $A = [0, 1]$ and μ is the Lebesgue measure, then \mathcal{F} is of product type if (Δ) does not hold.*

Proof. Consider the split-word process $(W_n, \eta_n)_{n \leq 0}$ on the alphabet $A = [0, 1]$ equipped with the Lebesgue measure, and its generated filtration \mathcal{F} . Assuming result 2.43, we shall see that \mathcal{F} satisfies Vershik's first level criterion if (Δ) does not hold. This will prove that \mathcal{F} is of product type thanks to Vershik's first level theorem (theorem 2.25). By proposition 2.22, for \mathcal{F} to satisfy Vershik's first level criterion, it suffices that each σ -field $\sigma(W_n)$ satisfies Vershik's first level criterion. In turn, this is proved as follows. For each $k \in \mathbb{N}$, we define the approximation of identity $f^k: [0, 1] \rightarrow [0, 1]$ by

$$f^k(u) = \sum_{i=0}^{2^k-1} \frac{i}{2^k} \mathbb{1}_{\{\frac{i}{2^k} < u \leq \frac{i+1}{2^k}\}}.$$

By result 2.44, we can see that $(f^k(W_n), \eta_n)_{n \leq 0}$ is a split-word process which generates a filtration \mathcal{F}^k of product type. Consequently, each filtration \mathcal{F}^k satisfies Vershik's first level criterion (corollary 2.14). Lemma 2.4 shows that

any innovation of a filtration \mathcal{F}^k is also an innovation of \mathcal{F} . Therefore, all random variables $f^k(W_n)$ satisfy Vershik's first level criterion with respect to \mathcal{F} . Hence, due to proposition 2.7, the random variable W_n satisfies Vershik's first level criterion with respect to \mathcal{F} . This amounts to saying that the σ -field $\sigma(W_n)$ satisfies Vershik's first level criterion (proposition 2.17). \square

Before turning to the next section, we give a corollary of part (b) of theorem 2.39 showing that, as announced above corollary 2.26, there exist some filtrations of local product type satisfying Vershik's first level criterion although they are not of product type. Note that the existence of split-word processes with a non-Polish alphabet is guaranteed by Ionescu-Tulcea's theorem.

Corollary 2.46. *In the context of theorem 2.39, if (Δ) does not hold, then \mathcal{F} satisfies Vershik's first level criterion, whatever (A, \mathfrak{A}, μ) . However, \mathcal{F} is not of product type if (A, \mathfrak{A}, μ) is not essentially separable.*

Proof. By proposition 2.22, it suffices to show that each σ -field $\sigma(W_n)$ satisfies Vershik's first level criterion. Thus, considering a measurable function $f: A^{\ell_n} \rightarrow \mathbb{R}$ such that the random variable $f(W_n)$ is integrable, we have to show that this random variable satisfies Vershik's first level criterion. As f is measurable with respect to $\mathfrak{A}^{\otimes \ell_n}$, there exist (lemma 1.1) some essentially separable σ -fields $\mathfrak{L}_1 \subset \mathfrak{A}, \dots, \mathfrak{L}_{\ell_n} \subset \mathfrak{A}$ such that f is measurable with respect to $\mathfrak{L}_1 \otimes \dots \otimes \mathfrak{L}_{\ell_n}$. Introduce the essentially separable σ -field $\mathfrak{B} = \mathfrak{L}_1 \vee \dots \vee \mathfrak{L}_{\ell_n}$; then f is measurable with respect to $\mathfrak{B}^{\otimes \ell_n}$. Thus, considering a measurable function $G: A \rightarrow \mathbb{R}$ such that $\sigma(G) = \mathfrak{B}$, the random variable $f(W_n)$ is measurable with respect to the σ -field generated by the random word $X_n := G(W_n(1)) \dots G(W_n(\ell_n))$. The process $(X_n, \eta_n)_{n \leq 0}$ is a split-word process on a Polish alphabet, and thus we know from theorem 2.39 that it generates a filtration of product type, so the σ -field $\sigma(X_n)$ satisfies Vershik's first level criterion with respect to the filtration generated by $(X_n, \eta_n)_{n \leq 0}$. By lemma 2.4, any innovation of this filtration is also an innovation of \mathcal{F} , hence the σ -field $\sigma(X_n)$ satisfies Vershik's first level criterion with respect to \mathcal{F} , and so does also the random variable $f(W_n)$. \square

3 Standardness and I-cosiness

Subsection 3.1 deals with *standard filtrations*, defined as filtrations that are *immersible* in a filtration (of product type) generated by independent random variables having a diffuse law. We shall see that any filtration of product type is standard and that being standard is equivalent to being immersible in a filtration of product type.

As an example of a sufficient condition for standardness, we provide an unpublished result (theorem 3.15) from Tsirelson on the existence of a *generating parameterization* for a filtration under a certain ergodicity condition

on a Markov process generating this filtration. The proof we give makes use of Rosenblatt's self-joining criterion (proposition 2.33). As we shall see, the existence of a generating parameterization for a filtration obviously implies standardness, but we disagree with a result in a literature asserting that the converse is true. This point is discussed in subsection 3.2.

In subsection 3.3, the notion of standardness is extended to its analogue for *locally separable filtrations* (definition 3.17), namely *weak standardness*, defined similarly to standardness but with productness replaced by Vershik's first level criterion.

Next, the I-cosiness criterion is defined in subsection 3.4 and its basic properties are given. The implications productness \Rightarrow I-cosiness and Vershik's first level criterion \Rightarrow I-cosiness will be directly deduced from Rosenblatt's self-joining criterion and Vershik's self-joining criterion, defined in the previous section. I-cosiness of standard or weakly standard filtrations will follow as an obvious consequence of I-cosiness being inherited by immersion. As an illustration of the I-cosiness criterion, we give a sufficient condition for a stationary Markov process to generate an I-cosy filtration.

Finally, in subsection 3.5, we pursue the proof of theorem 2.39 (productness of the filtrations of split-words processes).

3.1 Standardness, superinnovations, parameterizations

Two notions must preliminarily be defined before standardness: the notion of a *standard conditionally non-atomic filtration* and the notion of an *extension* of a filtration.

Definition 3.1. A conditionally non-atomic filtration is a filtration of local product type admitting a global innovation $(U_n)_{n \leq 0}$ such that each U_n is uniformly distributed on the interval $[0, 1]$. A standard conditionally non-atomic filtration is a filtration generated by a sequence $(U_n)_{n \leq 0}$ of independent random variables uniformly distributed on the interval $[0, 1]$.

Observe that all standard conditionally non-atomic filtrations are isomorphic to each other. Remark also that “uniformly distributed on the interval $[0, 1]$ ” in this definition can equivalently be replaced with “having a diffuse law”. The two lemmas below respectively characterize standard conditionally non-atomic filtrations and conditionally non-atomic filtrations.

Lemma 3.2. A filtration is standard conditionally non-atomic if and only if it is of product type and conditionally non-atomic.

Proof. The ‘only if’ part is obvious. The ‘if’ part follows from the fact that for a conditionally non-atomic filtration \mathcal{F} , any independent complement of \mathcal{F}_{n-1} in \mathcal{F}_n necessarily has a diffuse law, by virtue of lemma 2.4. \square

Lemma 3.3. *A filtration $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ is conditionally non-atomic if and only if for every $n \leq 0$, there exists an \mathcal{F}_n -measurable random variable V_n such that $\mathcal{F}_{n-1} \vee \sigma(V_n) = \mathcal{F}_n$ and such that the conditional law $\mathcal{L}(V_n | \mathcal{F}_{n-1})$ is almost surely diffuse.*

Proof. The ‘only if’ part is trivially true. For the ‘if’ part, consider V_n as in the lemma and let $F(\cdot | \mathcal{F}_{n-1})$ be the conditional cumulative distribution function of V_n given \mathcal{F}_{n-1} . One easily checks that the random variable $F(V_n | \mathcal{F}_{n-1})$ is an innovation from \mathcal{F}_{n-1} into \mathcal{F}_n having uniform law on $[0, 1]$. \square

To define standardness, we need one more notion, an *extension* of a filtration. Roughly speaking, an extension of \mathcal{F} is a filtration in which \mathcal{F} “can be immersed”.

Definition 3.4. *Let \mathcal{F} and \mathcal{G}' be two filtrations defined on possibly different probability spaces. The filtration \mathcal{F} is immersible in the filtration \mathcal{G}' , and \mathcal{G}' is an extension of \mathcal{F} , if \mathcal{F} is isomorphic to some filtration \mathcal{F}' immersed in \mathcal{G}' .*

Lemma 3.5. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ and $\mathcal{G} = (\mathcal{G}_n)_{n \leq 0}$ be two filtrations defined on possibly different probability spaces. Then both \mathcal{F} and \mathcal{G} are immersible in the product filtration $\mathcal{F} \otimes \mathcal{G}$.*

Proof. Let $\iota_1: \mathcal{F} \rightarrow \mathcal{F} \otimes \mathcal{G}$ and $\iota_2: \mathcal{G} \rightarrow \mathcal{F} \otimes \mathcal{G}$ be the identifications with the first factor and the second factor respectively (see example A.2). Then we know from lemma 1.2 that the two independent filtrations $\iota_1(\mathcal{F})$ and $\iota_2(\mathcal{G})$ are both immersed in $\iota_1(\mathcal{F}) \vee \iota_2(\mathcal{G}) = \mathcal{F} \otimes \mathcal{G}$. \square

Now we turn on to the notion of standard filtrations.

Definition 3.6. *A standard filtration is a filtration immersible in a standard conditionally non-atomic filtration.*

As obvious facts on standardness, we note:

- Standardness is preserved by isomorphism.
- A standard conditionally non-atomic filtration in the sense of definition 3.1 is standard and is conditionally non-atomic. But at this stage we are not yet able to prove the converse; this will be done in section 4 (corollary 4.7).
- The standardness property for a filtration is inherited by immersion, i.e., any filtration immersible in a standard filtration is itself standard.
- A standard filtration is essentially separable.

Proposition 3.7. *Any filtration of product type is standard, and a filtration is standard if and only if it is immersible in a filtration of product type.*

Proof. The independent product of a filtration \mathcal{F} with a standard non-atomic filtration is an extension of \mathcal{F} (lemma 3.5). Obviously, this product filtration is itself standard non-atomic if \mathcal{F} is of product type, and hence \mathcal{F} is standard. Consequently, a filtration is standard if it is immersible in a filtration of product type because standardness is inherited by immersion. The converse is obvious from the definition of standardness. \square

As we know from the pioneering works of Vershik, there exist some Kolmogorovian essentially separable filtrations that are not standard, i.e., which cannot be immersed in a standard conditionally non-atomic filtration. However theorem 3.9 below shows that *any essentially separable filtration can be immersed in the supremum of two jointly immersed standard conditionally non-atomic filtrations*. This theorem firstly says that an essentially separable conditionally non-atomic filtration *equals* the supremum of two jointly immersed standard conditionally non-atomic filtrations. This is the main assertion and it is due to Parry ([24]). The second assertion of this theorem follows from lemma 3.8 below. An interesting consequence of theorem 3.9 is given in [21] and is invoked in our remark 3.34. Except for this remark and for the remark below proposition 3.46, we will never use this theorem.

Lemma 3.8. *Any essentially separable filtration admits an essentially separable conditionally non-atomic extension.*

Proof. Using lemma 3.3, it is easy to see that the independent product of an essentially separable filtration and a standard conditionally non-atomic filtration is an essentially separable conditionally non-atomic filtration. Then the lemma immediately follows from lemma 3.5. \square

Theorem 3.9. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be an essentially separable conditionally non-atomic filtration. Then $\mathcal{F} = \mathcal{H}_1 \vee \mathcal{H}_2$ where \mathcal{H}_1 and \mathcal{H}_2 are two jointly immersed standard conditionally non-atomic filtrations. Consequently, any essentially separable filtration is immersible in such a filtration $\mathcal{H}_1 \vee \mathcal{H}_2$.*

Proof. The consequence follows from lemma 3.8. To prove the first assertion, we strictly follow [24]. Consider an essentially separable conditionally non-atomic filtration $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$. Let $(U_n)_{n \leq 0}$ be an innovation of \mathcal{F} such that each U_n is uniformly distributed on the interval $[0, 1]$. For every $n \leq 0$, let X_n be a random variable generating \mathcal{F}_n . Then X_n necessarily has a diffuse law, and we assume without loss of generality that this is the uniform law on $[0, 1]$. Define $V_n = X_{n-1} + U_n \pmod{1}$ for every $n \leq 0$. By lemma 2.4, $(V_n)_{n \leq 0}$ is an innovation of \mathcal{F} . It suffices to define \mathcal{H}_1 as the filtration generated by $(U_n)_{n \leq 0}$ and \mathcal{H}_2 as the filtration generated by $(V_n)_{n \leq 0}$. Then we easily see that $\mathcal{F} = \mathcal{H}_1 \vee \mathcal{H}_2$, and the joint immersion of \mathcal{H}_1 and \mathcal{H}_2 is a consequence of lemma 1.6. \square

A typical example of a standard conditionally non-atomic extension of a filtration \mathcal{F} is the filtration generated by a generating parameterization of \mathcal{F} , defined below.

Definition 3.10. A (global) superinnovation of a filtration $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ is a process $(V_n)_{n \leq 0}$ such that for each $n \leq 0$, the random variable V_n takes its values in a Polish space, is independent of $\mathcal{F}_{n-1} \vee \sigma(V_m; m \leq n-1)$, and satisfies $\mathcal{F}_n \subset \mathcal{F}_{n-1} \vee \sigma(V_n)$. The superinnovation $(V_n)_{n \leq 0}$ is a generating superinnovation if moreover \mathcal{F} is contained in the filtration generated by $(V_n)_{n \leq 0}$. A filtration \mathcal{F} admits a superinnovation if there exists a superinnovation of a filtration isomorphic to \mathcal{F} . A superinnovation $(V_n)_{n \leq 0}$ of \mathcal{F} is called a parameterization if V_n has the uniform law on $[0, 1]$ for every $n \leq 0$.

Obviously, a filtration admits a parameterization if and only if it admits a superinnovation $(V_n)_{n \leq 0}$ such that each V_n has a diffuse law. Proposition below shows that a filtration admitting a generating superinnovation is actually immersed in the filtration generated by this superinnovation.

Proposition 3.11. On $(\Omega, \mathcal{A}, \mathbb{P})$, let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration and $(V_n)_{n \leq 0}$ a process such that each V_n takes its values in a Polish space. Let $\mathcal{V} = (\mathcal{V}_n)_{n \leq 0}$ be the filtration generated by $(V_n)_{n \leq 0}$. The following conditions are equivalent.

- (i) $(V_n)_{n \leq 0}$ is a superinnovation of \mathcal{F} ;
- (ii) $(V_n)_{n \leq 0}$ is a sequence of independent random variables, one has $\mathcal{F}_n \subset \mathcal{F}_{n-1} \vee \sigma(V_n)$ for each $n \leq 0$, and \mathcal{V} is immersed in $\mathcal{F} \vee \mathcal{V}$;
- (iii) $(V_n)_{n \leq 0}$ is a sequence of independent random variables, one has $\mathcal{F}_n \subset \mathcal{F}_{n-1} \vee \sigma(V_n)$ for each $n \leq 0$, and \mathcal{F} and \mathcal{V} are jointly immersed.

Consequently, if $(V_n)_{n \leq 0}$ is a generating superinnovation for \mathcal{F} , then \mathcal{F} is immersed in \mathcal{V} , and hence \mathcal{F} is standard.

Proof. If $(V_n)_{n \leq 0}$ is a sequence of independent random variables, then lemma 1.6 shows that \mathcal{V} is immersed in $\mathcal{F} \vee \mathcal{V}$ if and only if V_n is independent of $\mathcal{F}_{n-1} \vee \sigma(V_m; m \leq n-1)$ for each $n \leq 0$. It follows that (i) \iff (ii). If $(V_n)_{n \leq 0}$ is a superinnovation of \mathcal{F} , then \mathcal{F} is immersed in $\mathcal{F} \vee \mathcal{V}$ as an easy consequence of lemma 1.6. That finally shows that (i) \iff (ii) \iff (iii). Consequently, if $(V_n)_{n \leq 0}$ is a generating superinnovation, then \mathcal{F} is standard by proposition 3.7 since it is immersed in the product type filtration $\mathcal{V} = \mathcal{F} \vee \mathcal{V}$. \square

Remark 3.12. It is claimed in the literature that any standard filtration admits a generating parameterization. Actually some confusion occurred, and to our knowledge there exists no valid proof of this assertion. This will be discussed in subsection 3.2.

An example from Tsirelson

Theorem 3.15 below gives a sufficient condition on a Markov process for its generated filtration to admit a generating parameterization, and hence to be standard (proposition 3.11). This result is borrowed from [34]. Lemma 3.14 is the key point of the proof. The last part of the proof we give illustrates Rosenblatt's self-joining criterion (proposition 2.33).

The following lemma will be used in the proof of lemma 3.14 (and in the statement of lemma 3.24 and the proof of lemma 3.27). This lemma is a verbatim copy of lemma 6.4.6 in [6].

Lemma 3.13. *Let (Ω, \mathcal{A}) be a measurable space and $f: \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ a function satisfying the following conditions: for every fixed $t \in \mathbb{R}$, the function $\omega \mapsto f(\omega, t)$ is \mathcal{A} -measurable, and for every fixed $\omega \in \Omega$, the function $t \mapsto f(\omega, t)$ is right-continuous. Then the function f is measurable with respect to $\mathcal{A} \otimes \mathfrak{B}_{\mathbb{R}}$.*

The statement and the proof of lemma 3.14 and theorem 3.15 involve infimum of measures. Given two measures ν_1 and ν_2 on a measurable space, we denote by $\nu_1 \wedge \nu_2$ the infimum of ν_1 and ν_2 . The existence of this measure is guaranteed; more generally, the infimum of an infinite family of measures always exists (see [33], Theorem 7.1, or [9], Appendices III & IV.) When $\nu_1 = f_1 \cdot \mu$ and $\nu_2 = f_2 \cdot \mu$ (Radon-Nikodým derivatives), $\nu_1 \wedge \nu_2 = (f_1 \wedge f_2) \cdot \mu$.

Lemma 3.14. *Let X and Y be two random variables taking values in some Polish spaces E and F respectively, and $\{\nu_x\}_{x \in E}$ a regular version of the conditional law of Y given X . Denote by μ_1 and μ_2 the respective laws of X and Y and by Leb the Lebesgue probability measure on $[0, 1]$. There exists a measurable function $p: E \times F \rightarrow [0, 1]$ such that for μ_1 -almost all $x \in E$, the function $y \mapsto p(x, y)$ is a Radon-Nikodým derivative of $\nu_x \wedge \mu_2$ with respect to μ_2 , and there exists a measurable function $\alpha: E \times F \times [0, 1] \rightarrow \mathbb{R}$ such that one has $\alpha(x, \mu_2 \times \text{Leb}) = \nu_x$ for every x and $\alpha(x, y, u) = y$ for every x, y, u satisfying $u \leq p(x, y)$.*

Proof. Since any Polish probability space is Lebesgue isomorphic to a probability space on \mathbb{R} , it suffices to do the proof for $F = \mathbb{R}$. Let f be a Radon-Nikodým derivative of the absolutely continuous part in the Lebesgue decomposition of the joint distribution of (X, Y) with respect to $\mu_1 \otimes \mu_2$. The set of values of $x \in E$ such that $\int f(x, y) d\mu_2(y) = 0$, is μ_1 -negligible, and for those values of x for which $\int f(x, y) d\mu_2(y) \neq 0$, it is not difficult to check that the function $p_x: y \mapsto \min \{f(x, y) / \int f(x, z) d\mu_2(z), 1\}$ is a Radon-Nikodým derivative of the measure $\nu_x \wedge \mu_2$ with respect to μ_2 . We define p by $p(x, y) = p_x(y)$. Now we are going to construct α . The function $t \mapsto \nu_x([-\infty, t]) - (\nu_x \wedge \mu_2)([-\infty, t])$ is right-continuous and increasing, and takes its values in $[0, 1 - m_x]$ where m_x is the total mass of $\nu_x \wedge \mu_2$. Call g_x the right-continuous inverse of this function. The function $(x, v) \mapsto g_x(v)$ is measurable by virtue of lemma 3.13. Then put $\alpha(x, y, u) = g_x(\frac{1-m_x}{1-p_x(y)}(1-u))$ for $u \in [p_x(y), 1]$. One checks without difficulty that $\alpha(x, \mu_2 \times \text{Leb}) = \nu_x$. \square

Theorem 3.15. *Consider a Markov process $(X_n)_{n \leq 0}$ where each X_n takes its values in a Polish space. Denote by μ_n the law of X_n and by $\nu_{x_{n-1}}^n$ the conditional law $\mathcal{L}(X_n | X_{n-1} = x_{n-1})$. Let m_n be the $\mu_{n-1} \otimes \mu_{n-1}$ -essential infimum over x'_{n-1}, x''_{n-1} of the total masses of the measures $\nu_{x'_{n-1}}^n \wedge \nu_{x''_{n-1}}^n \wedge \mu_n$. If $\sum m_n = +\infty$, then the filtration generated by $(X_n)_{n \leq 0}$ admits a generating parameterization.*

As an application of this theorem, we can see that the filtration generated by a stationary random walk on the vertices of a triangle admits a generating parameterization. Note that the number m_n defined in theorem 3.15 satisfies $m_n \geq \beta(\nu^n)$ where $\beta(\nu^n)$ is the total mass of the essential infimum over all x of the measures ν_x^n . Thus the condition $\sum \beta(\nu^n) = +\infty$ guarantees the existence of a generating parameterization. In the stationary case, this condition is equivalent to the existence of a positive non-null measure that minorizes the probability measures $\mathcal{L}(X_n | X_{n-1} = x)$ for almost all x . A weaker minorization condition given in [17] guarantees the existence of a generating parameterization in the stationary case.

Proof (Proof of theorem 3.15). Let p_n and α_n be the functions p and α obtained from lemma 3.14 applied with $X = X_{n-1}$ and $Y = X_n$. We write $p_x^n(y) = p_n(x, y)$. Now consider the Markov process $(X'_n, (Y'_n, U'_n))_{n \leq 0}$ defined as follows:

- for each $n \leq 0$, (Y'_n, U'_n) is independent of the past up to $n - 1$, Y'_n has law μ_n , U'_n has the uniform law on $[0, 1]$, Y'_n and U'_n are independent;
- $X'_n = \alpha_n(X'_{n-1}, Y'_n, U'_n)$ for each $n \leq 0$.

Then we know from lemma 3.14 that $(X'_n)_{n \leq 0}$ has the same distribution as the Markov process $(X_n)_{n \leq 0}$, and that $X'_n = Y'_n$ if $U'_n \leq p_{X'_{n-1}}^n(Y'_n)$. Let \mathcal{F}' be the filtration generated by $(X'_n)_{n \leq 0}$. Obviously, the process $(Y'_n, U'_n)_{n \leq 0}$ is a superinnovation of \mathcal{F}' (definition 3.10), and we shall show with the help of Rosenblatt's self-joining criterion (proposition 2.33) that this superinnovation is generating if $\sum m_n = +\infty$.

Let $(X_n^*)_{n \leq 0}$ be a copy of $(X_n)_{n \leq 0}$ lying on the same probability space $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$ as $(X'_n, (Y'_n, U'_n))_{n \leq 0}$ and independent of $(X'_n, (Y'_n, U'_n))_{n \leq 0}$. For a given integer $n_0 < 0$, we define another copy $(X''_n)_{n \leq 0}$ of $(X_n)_{n \leq 0}$ by setting $X''_n = X_n^*$ for $n \leq n_0$ and $X''_{n+1} = \alpha_n(X''_n, Y'_{n+1}, U'_{n+1})$ for n going from n_0 to -1 . By proposition 2.33, it suffices to show that $\bar{\mathbb{P}}[X'_n \neq X''_n] \rightarrow 0$ as n_0 goes to $-\infty$ for each $n \leq 0$. For $n \geq n_0$ one has $(X'_n = X''_n) \implies (X'_{n+1} = X''_{n+1})$, so it suffices to show that the inequality $\bar{\mathbb{P}}[X'_n \neq X''_n | X'_{n-1}, X''_{n-1}] \leq 1 - m_n$ almost surely holds for every $n \in \{n_0 + 1, \dots, 0\}$, because this yields the inequality $\bar{\mathbb{P}}[X'_n \neq X''_n] \leq \prod_{k=n_0+1}^n (1 - m_k)$. Now one has

$$\bar{\mathbb{P}}[\alpha_n(x'_{n-1}, Y'_n, U'_n) = \alpha_n(x''_{n-1}, Y'_n, U'_n)] \leq \bar{\mathbb{P}}[U'_n \leq p_{x'_{n-1}}^n(Y'_n) \wedge p_{x''_{n-1}}^n(Y'_n)].$$

But we can see that

$$\bar{\mathbb{P}}[U'_n \leq p_{x'_{n-1}}^n(Y'_n) \wedge p_{x''_{n-1}}^n(Y'_n)] = \int d\mu_n(y_n) (p_{x'_{n-1}}^n(y_n) \wedge p_{x''_{n-1}}^n(y_n) \wedge 1)$$

is nothing but the total mass of $\nu_{x'_{n-1}}^n \wedge \nu_{x''_{n-1}}^n \wedge \mu_n$. The proof is over. \square

Remark 3.16. It is known that the total variation $\|\nu_1 - \nu_2\| \in [0, 2]$ between two probability measures ν_1 and ν_2 on an arbitrary measurable space E satisfies

$\|\nu_1 - \nu_2\| = 2(1 - (\nu_1 \wedge \nu_2)(E))$ (see [33]). Therefore the number m_n defined in theorem 3.15 satisfies $2m_n \leq 2 - \alpha(\nu^n)$ where $\alpha(\nu^n) = \text{ess sup}_{x', x''} \|\nu_{x'}^n - \nu_{x''}^n\|$. It would be interesting to know if the condition $\sum_n (2 - \alpha(\nu^n)) = +\infty$ guarantees standardness of the filtration.

3.2 Erratum on generating parameterizations

Some confusion occurred in the articles [15], [32], [30]. It was erroneously claimed that, for a discrete negative-time filtration, every standard conditionally non-atomic extension is *obviously* induced by a generating parameterization (definition 3.10). Schachermayer gave a counter-example in [31]. As this (false) claim was considered as a proof that every standard filtration admits a generating parameterization, a new proof of the latter fact was needed. Feldman and Smorodinsky claimed in [16] that this fact is nonetheless true, and gave a proof using results from the literature. Unfortunately, a confusion of the same kind occurred again in the proof proposed in [16]. Thus, to our knowledge, *there does not exist any proof of this assertion*. The proof given in [20] contains an error too.

The reiterated confusion in the proof proposed in [16] lies on page 1086 of [15], where it is claimed that if the independent product of a filtration \mathcal{F} with a standard non-atomic filtration is itself standard non-atomic, then any generating innovation of the latter is a generating parameterization of \mathcal{F} . This is false, as we shall see, and this is unfortunately an ingredient in the proof proposed in [16]. This confusion is of the same kind as the one which was pointed out in [31]: a filtration is immersed in its independent product with another filtration, but a sequence that generates the product may not be a parameterization of \mathcal{F} , even though it generates a standard conditionally non-atomic extension of \mathcal{F} .

Here is a counter-example. It is of the same spirit as the counter-example given in [31]. On a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, consider a sequence $(U_n)_{n \leq 0}$ of independent random variables uniformly distributed in $[0, 1]$, and a random variable U_0^* uniformly distributed in $[0, 1]$ and independent of $(U_n)_{n \leq 0}$. Define $X_0 = U_{-1} + U_0 \pmod{1}$. Let \mathcal{F} be the filtration defined by $\mathcal{F}_n = \{\emptyset, \Omega\}$ for $n \leq -1$ and $\mathcal{F}_0 = \sigma(X_0)$. Let \mathcal{G} be the filtration defined by $\mathcal{G}_n = \sigma(U_m, m \leq n)$ for $n \leq -1$ and $\mathcal{G}_0 = \mathcal{G}_{-1} \vee \sigma(U_0^*)$. Then \mathcal{G} is a standard conditionally non-atomic filtration independent of \mathcal{F} . Consider any random variable V_0 uniformly distributed on $[0, 1]$ and such that $\sigma(V_0) = \sigma(U_0, U_0^*)$. One easily verifies that $\mathcal{F} \vee \mathcal{G}$ is generated by the sequence of independent random variables $(\dots, U_{-2}, U_{-1}, V_0)$. However, $(\dots, U_{-2}, U_{-1}, V_0)$ is not a parameterization for \mathcal{F} , because the inclusion $\mathcal{F}_0 \subset \mathcal{F}_{-1} \vee \sigma(V_0)$ does not hold.

So we consider the statement S_1 : “Standardness is equivalent to the existence of a generating parametrization” as an open question. Note that S_1 is equivalent to S_2 : “A filtration immersed in a filtration which admits a generating parameterization, admits itself a generating parameterization”. Indeed, if every standard filtration admits a generating parameterization, then S_2 is true

owing to the fact that standardness is inherited by immersion. Conversely, if S_2 is true, then every standard filtration admits a generating parameterization because a standard conditionally non-atomic filtration obviously admits a generating parameterization.

3.3 Weak standardness. Locally separable filtrations

Obviously, standard filtrations must be essentially separable. For essentially separable filtrations, it is already known ([14]) that standardness is equivalent to I-cosiness (definition 3.29), and also to Vershik's standardness criterion (which is not stated in this paper; see [14] and [21]). However the I-cosiness criterion and Vershik's standardness criterion could a priori be satisfied for a filtration which is not essentially separable. We shall soon define *weak standardness*, and we shall see in section 4 that weak standardness is equivalent to I-cosiness for a *locally separable filtration*², defined as follows.

Definition 3.17. 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 $\mathcal{F}_n = \mathcal{F}_{n-1} \vee \sigma(V_n)$.

Thus, any essentially separable filtration is locally separable, and filtrations of local product type (definition 2.3), obviously are locally separable filtrations. We take the opportunity of definition 3.17 to state a conjecture about local separability.

Conjecture 3.18. A filtration immersible in a locally separable filtration is itself locally separable.

We will later see (corollary 3.28) that locally separable filtrations are precisely filtrations that admit a superinnovation or, equivalently, a parameterization (definition 3.10).

The notion of *weak standardness* defined below invokes Vershik's first level criterion (definition 2.6, definition 2.9, proposition 2.17). The generalization from standardness to weak standardness is based on the fact that Vershik's first level criterion is equivalent to productness for an essentially separable filtration (theorem 2.25). Inspired by lemma 3.2, we first define a *weakly standard conditionally non-atomic filtration* as the following generalization of a standard conditionally non-atomic filtration.

Definition 3.19. A weakly standard conditionally non-atomic filtration is a conditionally non-atomic filtration satisfying Vershik's first level criterion.

Lemma 3.20. A filtration is standard non-atomic if and only if it is weakly standard conditionally non-atomic and essentially separable.

² It is also proved in [21] that the equivalence between I-cosiness and Vershik's standardness criterion remains true for locally separable filtrations.

Proof. This follows from Vershik's first level criterion (theorem 2.25) and lemma 3.2. \square

Then, the notion of weak standardness is defined analogously to the notion of standardness (definition 3.6).

Definition 3.21. *A filtration is weakly standard if it is immersible in a weakly standard conditionally non-atomic filtration.*

As obvious remarks, we note:

- Weak standardness is preserved by isomorphism.
- If conjecture 3.18 is true, then every weakly standard filtration is locally separable.
- Weak standardness is hereditary for immersion: a filtration immersible in a weakly standard filtration is itself weakly standard.
- As a product type filtration satisfies Vershik's first level criterion (theorem 2.25), standardness implies weak standardness. Thus a standard filtration is weakly standard and essentially separable. However we are not yet able to show the converse; it will be proved in section 4 (corollary 4.10).
- Obviously, a weakly standard conditionally non-atomic filtration is weakly standard and conditionally non-atomic. We will see in section 4 that the converse is true (corollary 4.6).

Proposition 3.23 below is analogous to proposition 3.7. Its proof invokes the following lemma.

Lemma 3.22. *The independent product of a conditionally non-atomic filtration and a locally separable filtration is conditionally non-atomic.*

Proof. Left to the reader as an easy application of lemma 3.3. \square

Proposition 3.23. *Any filtration satisfying Vershik's first level criterion is weakly standard; and a filtration is weakly standard if and only if it is immersible in a filtration satisfying Vershik's first level criterion.*

Proof. The independent product of a filtration \mathcal{F} with a standard conditionally non-atomic filtration is an extension of \mathcal{F} (lemma 3.5), and this product filtration is conditionally non-atomic by lemma 3.22. Moreover, a standard conditionally non-atomic filtration satisfies Vershik's first level criterion (corollary 2.14), hence if \mathcal{F} satisfies Vershik's first level criterion, then so does this product filtration in view of corollary 2.27. Thus we have proved that every filtration satisfying Vershik's first level criterion is weakly standard. Consequently, a filtration is weakly standard if it is immersible in a filtration satisfying Vershik's first level criterion because weak standardness is hereditary for immersion. The converse is obviously true from the definition of weak standardness. \square

Before turning to the next subsection devoted to I-cosiness, we are going to prove (corollary 3.28) that a filtration is locally separable if and only if it admits a global parameterization (definition 3.10). We will make use of the following lemma. The construction appearing in this lemma is the general *conditional quantile transformation* (see [23]). In the statement of this lemma, we implicitly use lemma 3.13 to justify the measurability of $F_{\mathcal{C}}$ and $F_{\mathcal{C}}^-$ with respect to $\mathcal{C} \otimes \mathfrak{B}_{\mathbb{R}}$ and the measurability of $G_{\mathcal{C}}$ with respect to $\mathcal{C} \otimes \mathfrak{B}_{[0,1]}$.

Lemma 3.24. *Let X be a real random variable on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and $\mathcal{C} \subset \mathcal{A}$ be a σ -field. Let $F_{\mathcal{C}}$ be the cumulative distribution function of the conditional law of X given \mathcal{C} , and let $F_{\mathcal{C}}^-(x) = \lim_{x' \rightarrow x^-} F_{\mathcal{C}}(x')$ be the left limit of $F_{\mathcal{C}}(x')$ as x' approaches x . Let ξ be a random variable with uniform law on $[0, 1]$ and independent of $\mathcal{C} \vee \sigma(X)$. We put*

$$U = F_{\mathcal{C}}^-(X) + \xi (F_{\mathcal{C}}(X) - F_{\mathcal{C}}^-(X)).$$

Then U is a random variable independent of \mathcal{C} , uniformly distributed on $[0, 1]$, and one has $X = G_{\mathcal{C}}(U)$ where $G_{\mathcal{C}}$ is the right-continuous inverse function of $F_{\mathcal{C}}$, defined by $G_{\mathcal{C}}(u) = \inf \{x \mid F_{\mathcal{C}}(x) > u\}$.

Proof. It suffices to show the lemma in the case when \mathcal{C} is degenerate. We write F , F^- and G instead of $F_{\mathcal{C}}$, $F_{\mathcal{C}}^-$ and $G_{\mathcal{C}}$ respectively. Denote by $S = \{x_1, x_2, \dots\}$ the denumerable set of atoms of X . Conditionally on the event $X = x_i$, the random variable U has the uniform law on $(F^-(x_i), F(x_i))$ and one has $X = G(U)$; and conditionally on the event $X \notin S$, one has $U = F(X)$, so the distribution of U is uniform on $[0, 1] \setminus \bigcup_i [F^-(x_i), F(x_i)]$, and one has $X = G(U)$. Finally the distribution of U is the uniform law on $[0, 1]$ and one has $X = G(U)$. \square

Proposition 3.25. *Any locally separable filtration admits a global parameterization.*

Proof. Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a locally separable filtration on $(\Omega, \mathcal{A}, \mathbb{P})$. For each $n \leq 0$, let V_n be a random variable such that $\mathcal{F}_n = \mathcal{F}_{n-1} \vee \sigma(V_n)$. On some probability space $(\Omega^*, \mathcal{C}^*, \mathbb{P}^*)$, consider a sequence $(\xi_n)_{n \leq 0}$ of independent random variables having uniform law on $[0, 1]$. We work on the product probability space $(\widehat{\Omega}, \widehat{\mathcal{A}}, \widehat{\mathbb{P}}) := (\Omega, \mathcal{F}_0, \mathbb{P}) \otimes (\Omega^*, \mathcal{C}^*, \mathbb{P}^*)$ and we identify \mathcal{F} and $(\xi_n)_{n \leq 0}$ with their image under the canonical embedding (see example A.2) from $(\Omega, \mathcal{F}_0, \mathbb{P})$ to $(\widehat{\Omega}, \widehat{\mathcal{A}}, \widehat{\mathbb{P}})$ and the canonical embedding from $(\Omega^*, \mathcal{C}^*, \mathbb{P}^*)$ to $(\widehat{\Omega}, \widehat{\mathcal{A}}, \widehat{\mathbb{P}})$ respectively, so the sequence $(\xi_n)_{n \leq 0}$ is independent of \mathcal{F}_0 . We denote by $\mathcal{D} = (\mathcal{D}_n)_{n \leq 0}$ the filtration generated by $(\xi_n)_{n \leq 0}$. For each $n \leq 0$, the random variable ξ_n is independent of $\mathcal{F}_0 \vee \mathcal{D}_{n-1}$ and, since \mathcal{F} and \mathcal{D} are jointly immersed (lemma 1.2), one has $\mathcal{L}(V_n \mid \mathcal{F}_{n-1} \vee \mathcal{D}_{n-1}) = \mathcal{L}(V_n \mid \mathcal{F}_{n-1})$. For each $n \leq 0$, let U_n be the random variable called U in lemma 3.24 when this lemma is applied with $X = V_n$, $\mathcal{C} = \mathcal{F}_{n-1} \vee \mathcal{D}_{n-1}$ and $\xi = \xi_n$. One easily checks that $(U_n)_{n \leq 0}$ is a parameterization of \mathcal{F} . \square

Remark 3.26. Let $\mathcal{U} = (\mathcal{U}_n)_{n \leq 0}$ be the filtration generated by the parameterization $(U_n)_{n \leq 0}$ of \mathcal{F} in the preceding proof. It follows from lemma 1.6 that $\mathcal{F} \vee \mathcal{U}$ is immersed in $\mathcal{F} \vee \mathcal{D}$. Therefore, if \mathcal{F} is standard, then $\mathcal{F} \vee \mathcal{U}$ is also standard because $\mathcal{F} \vee \mathcal{D}$ is standard (proposition 3.7). It is shown in [21], with the help of Vershik's standardness criterion, that this is actually true for an arbitrary parameterization $(U_n)_{n \leq 0}$ of \mathcal{F} .

The converse of proposition 3.25 is an easy consequence of (ii) \implies (i) in the following lemma.

Lemma 3.27. *On $(\Omega, \mathcal{A}, \mathbb{P})$, let \mathcal{C} and \mathcal{B} be two σ -fields such that $\mathcal{C} \subset \mathcal{B}$. The following conditions are equivalent:*

- (i) *there exists a random variable V such that $\mathcal{B} = \mathcal{C} \vee \sigma(V)$;*
- (ii) *there exists a random variable W such that $\mathcal{B} \subset \mathcal{C} \vee \sigma(W)$;*
- (iii) *there exist a probability space $(\Omega', \mathcal{A}', \mathbb{P}')$ and an embedding $\Psi: \mathcal{B} \rightarrow \mathcal{A}'$ such that $\Psi(\mathcal{B}) \subset \Psi(\mathcal{C}) \vee \sigma(U')$ where U' is a random variable uniformly distributed on $[0, 1]$ and independent of $\Psi(\mathcal{C})$.*

Proof. Obviously, (i) \implies (ii) is true. To prove (ii) \implies (i) we make the non restrictive assumption that W is valued in a Polish space. Then (ii) \implies (i) is obtained by putting $V = \mathcal{L}(W | \mathcal{B})$. Indeed, any bounded random variable X measurable with respect to $\mathcal{C} \vee \sigma(W)$ can be written $X = f(C, W)$ where C is a \mathcal{C} -measurable random variable and f is a bounded measurable function (lemma 1.1); thus $\mathbb{E}[X | \mathcal{B}] = \int f(C, w) \mathcal{L}(W | \mathcal{B})(dw)$ is measurable with respect to $\mathcal{C} \vee \sigma(V)$, and hence $\mathcal{B} \subset \mathcal{C} \vee \sigma(V)$. To prove (i) \implies (iii) we make the non restrictive assumption that V is real-valued. Consider the product probability space of $(\Omega, \mathcal{A}, \mathbb{P})$ with a probability space on which is defined a random variable U^* having the uniform law on $[0, 1]$. Let $\mathcal{C}' = \iota_1(\mathcal{C})$ and $U' = \iota_2(U^*)$ respectively be the copies of \mathcal{C} and U^* on the product probability space with the canonical embeddings $\iota_1: \mathcal{C} \rightarrow \mathcal{C} \otimes \sigma(U^*)$ and $\iota_2: \sigma(U^*) \rightarrow \mathcal{C} \otimes \sigma(U^*)$ (see example A.2). We introduce the right-continuous inverse $G(\cdot | \mathcal{C}')$ of the cumulative distribution function of the copy of the conditional law $\mathcal{L}(V | \mathcal{C})$ with the first embedding ι_1 , and then we put $V' = G(U' | \mathcal{C}')$. Then V' is a well-defined random variable by virtue of lemma 3.13, and, with the help of lemma A.8, it is easy to check that the conditional law $\mathcal{L}(V' | \mathcal{C}')$ is the copy of $\mathcal{L}(V | \mathcal{C})$ with ι_1 , therefore the embedding Ψ is given by corollary A.12. Finally, (iii) \implies (i) is a consequence of (ii) \implies (i). \square

Corollary 3.28. *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;*
- (iv) *\mathcal{F} admits a global superinnovation.*

Proof. This stems from lemma 3.27 and proposition 3.25. \square

3.4 I-cosiness

The *I-cosiness criterion* introduced in [14] was inspired from two sources: it is a variant of the notion of *cosiness* introduced by Tsirelson in [35] in the framework of continuous time, and the authors of [14] noticed that I-cosiness is used (but not named) in [32] to prove result 2.43 about the split-word filtrations. In fact, as was pointed out in [3], “there is a whole range of possible variations” on the definition of cosiness introduced in [35], and the main underlying idea, due to Tsirelson, is what [Lau09] calls a self-joining criterion, which comprises these possible variants of cosiness and in particular the I-cosiness criterion, as well as Rosenblatt’s self-joining criterion and Vershik’s self-joining criterion introduced in section 2.2. Actually many elementary results we give on I-cosiness remain valid for any self-joining criterion, as defined in [21]. However we prefer not to introduce this notion here: there are already too many definitions.

It is shown in [14] that standardness and I-cosiness are equivalent properties for an essentially separable filtration. In the next section, we will give another proof of this fact and extend it to locally separable filtrations by showing that I-cosiness is equivalent to weak standardness.

In this subsection, we define I-cosiness and give more or less elementary results concerning it. We shall see that, for a filtration of product type and a filtration satisfying Vershik’s first level criterion, I-cosiness is straightforward from Rosenblatt’s self-joining criterion and Vershik’s self-joining criterion respectively (subsection 2.2). That a standard or a weakly standard filtration is I-cosy follows from the fact that I-cosiness, as standardness and weak standardness, is inherited by immersion. We end this subsection by giving an example of stationary Markov processes whose filtrations are I-cosy.

Definition 3.29. Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration.

1. Let (E, ρ) be a Polish metric space and $X \in L^1(\mathcal{F}_0; E)$. We say that the random variable X is I-cosy (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 $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$ such that:
 - (i) $(\mathcal{F}', \mathcal{F}'')$ is a joining of \mathcal{F} independent in small time (definition 2.29);
 - (ii) $\bar{\mathbb{E}}[\rho(X', X'')] < \delta$, where X' and X'' are the respective copies of X in \mathcal{F}' and in \mathcal{F}'' .
2. We say that a σ -field $\mathcal{E}_0 \subset \mathcal{F}_0$ is I-cosy (with respect to \mathcal{F}) if every random variable $X \in L^1(\mathcal{E}_0)$ is I-cosy with respect to \mathcal{F} .
3. We say that the filtration \mathcal{F} is I-cosy if the final σ -field \mathcal{F}_0 is I-cosy with respect to \mathcal{F} .

As with Vershik’s first level criterion, we will sometimes omit to specify *with respect to \mathcal{F}* when no ambiguity is possible. We will see in proposition 3.36 that I-cosiness of a random variable X is equivalent to the σ -field $\sigma(X)$ being I-cosy. It is clear that I-cosiness is preserved by isomorphism. Another

obvious property of I-cosiness is hereditability by immersion, stated in the next lemma.

Lemma 3.30. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration, E a Polish space, $X \in L^1(\mathcal{F}_0; E)$, and \mathcal{E} a filtration immersed in \mathcal{F} . If X is I-cosy with respect to \mathcal{F} and is \mathcal{E}_0 -measurable, then X is I-cosy with respect to \mathcal{E} . Consequently, if \mathcal{F} is I-cosy, so is also \mathcal{E} .*

Proof. This is a straightforward consequence of the definition of I-cosiness and the transitivity property of immersion. \square

Proposition 3.31. *A filtration of product type is I-cosy. A filtration satisfying Vershik's first level criterion is I-cosy. More precisely, a random variable, or a σ -field, satisfying Vershik's first level criterion with respect to a filtration of local product type, is I-cosy with respect to this filtration.*

Proof. This follows from Rosenblatt's self-joining criterion (proposition 2.33) and Vershik's self-joining criterion (theorem 2.38). Indeed, each of these criteria is a particular case of the I-cosiness criterion. \square

Corollary 3.32. *Any standard (definition 3.6) or weakly standard (definition 3.21) filtration is I-cosy.*

Proof. This follows from lemma 3.30 and proposition 3.31. \square

Below we shall list some elementary properties of I-cosiness.

Lemma 3.33. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration and E a Polish metric space. The random variables $X \in L^1(\mathcal{F}_0; E)$ which are I-cosy form a closed subset of $L^1(\mathcal{F}_0; E)$.*

Proof. Take $R \in L^1(\mathcal{F}_0; E)$ in the L^1 -closure of the set of I-cosy random variables $X \in L^1(\mathcal{F}_0; E)$. Given $\delta > 0$, there exists an I-cosy $X \in L^1(\mathcal{F}_0; E)$ such that $\mathbb{E}[\rho(R, X)] < \delta/3$. By I-cosiness, there exists a joining $(\mathcal{F}', \mathcal{F}'')$ such that $\mathbb{E}[\rho(X', X'')] < \delta/3$. By isomorphisms, we have $\mathbb{E}[\rho(X', R')] = \mathbb{E}[\rho(X'', R'')] < \delta/3$; so the triangular inequality gives $\mathbb{E}[\rho(R', R'')] < \delta$. \square

Remark 3.34. If X and Y are two I-cosy random variables, it is not true in general that (X, Y) is also I-cosy. This is shown in [21] with the help of theorem 3.9 and Vershik's standardness criterion.

The following lemma plays the same role for I-cosiness as lemma 2.10 for Vershik's first level criterion. In the second condition, we consider $L(\mathcal{E}_0; F)$ as the space $L^1(\mathcal{E}_0; (F, \rho))$ where the metric ρ is the 0 – 1 distance; so $\mathbb{E}[\rho(X', X'')] = \mathbb{P}[X' \neq X'']$.

Lemma 3.35. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration on $(\Omega, \mathcal{A}, \mathbb{P})$ and $\mathcal{E}_0 \subset \mathcal{F}_0$ a σ -field. The following conditions are equivalent:*

- (i) *the σ -field \mathcal{E}_0 is I-cosy with respect to \mathcal{F} ;*
- (ii) *for any finite set $F \subset \mathbb{R}$, every random variable $X \in L(\mathcal{E}_0; F)$ is I-cosy with respect to \mathcal{F} ;*
- (iii) *for any Polish space E , every random variable $X \in L^1(\mathcal{E}_0; E)$ is I-cosy with respect to \mathcal{F} .*

Proof. (iii) \implies (i) is trivial.

(i) \implies (ii): Fix F finite, $R \in L(\mathcal{E}_0; F)$, and $\delta > 0$. Let a be the minimum distance $|s - t|$ between two distinct elements s, t of F . Applying hypothesis (i), one obtains a joining $(\mathcal{F}', \mathcal{F}'')$ such that $\mathbb{E}[|R' - R''|] < \epsilon\delta$ where $\epsilon = \min\{a, \delta\}$; hence we have $\mathbb{P}(|R' - R''| \geq \epsilon) < \delta$ and therefore $\mathbb{P}[R' \neq R''] < \delta$.

(ii) \implies (iii): Fix $X \in L^1(\mathcal{E}_0; (E, \rho))$ and $\delta > 0$. There exist some finite subset F of E and some $R \in L(\mathcal{E}_0; F)$ such that $\mathbb{E}[\rho(X, R)] < \delta/3$. Call d the diameter of F . Given a measurable injection $\phi: F \rightarrow \mathbb{R}$ and applying hypothesis (ii) to the random variable $\phi(R)$, one obtains a joining $(\mathcal{F}', \mathcal{F}'')$ such that $\mathbb{P}[R' \neq R''] < \delta/(3d)$; so $\mathbb{E}[\rho(R', R'')] < \delta/3$. Now the isomorphisms give $\mathbb{E}[\rho(X', R')] = \mathbb{E}[\rho(X'', R'')] < \delta/3$, wherefrom $\mathbb{E}[\rho(X', X'')] < \delta$ by the triangular inequality. \square

Proposition 3.36. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration. Let (E, ρ) be a Polish metric space and $X \in L^0(\mathcal{F}_0; E)$. The following conditions are equivalent.*

- (i) *the σ -field $\sigma(X)$ is I-cosy;*
- (ii) *for every $\delta > 0$, there exists, on some probability space $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$, a joining $(\mathcal{F}', \mathcal{F}'')$ of \mathcal{F} independent in small time such that $\mathbb{P}[\rho(X', X'') > \delta] < \delta$, where X' and X'' are the respective copies of X in \mathcal{F}' and in \mathcal{F}'' .*

If $X \in L^1(\mathcal{F}_0; E)$, these conditions are also equivalent to:

- (iii) *X is I-cosy.*

Proof. We first assume that $X \in L^1(\mathcal{F}_0; E)$ and show (i) \iff (iii). If the σ -field $\sigma(X)$ is I-cosy, then X is I-cosy by lemma 3.35. Conversely, assume X to be I-cosy. We know from lemma 2.15 that the set of random variables of the form $f(X)$ with $f: E \rightarrow \mathbb{R}$ Lipschitz, is a dense subset of $L^1(\sigma(X))$. It is easy to see that such a random variable $f(X)$ is I-cosy. Therefore the σ -field $\sigma(X)$ is I-cosy as a consequence of lemma 3.33. The proof that (i) \iff (iii). It is then not hard to derive (i) \iff (ii) by replacing ρ with $\rho \wedge 1$. \square

We will use the following lemma to prove the asymptotic character of I-cosiness (proposition 3.38) and to prove proposition 3.43. This lemma involves I-cosiness of a filtration $(\mathcal{F}_n)_{n \leq N}$ with time-axis $-\mathbb{N} \cap]-\infty, N]$ for some integer $N \leq 0$, whereas I-cosiness is defined for a filtration indexed by $-\mathbb{N}$; but it is clear how to adapt the definition to this time-axis and obviously I-cosiness of $(\mathcal{F}_n)_{n \leq N}$ is equivalent to I-cosiness of $(\mathcal{F}_{N+n})_{n \leq 0}$.

Lemma 3.37. *On $(\Omega, \mathcal{A}, \mathbb{P})$, let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration and $(V_n)_{n \leq 0}$ a superinnovation of \mathcal{F} . Let $N < 0$ be an integer and $\mathcal{E}_N \subset \mathcal{F}_N$ a σ -field. If \mathcal{E}_N is I-cosy with respect to the truncated filtration $(\mathcal{F}_n)_{n \leq N}$, then the σ -field $\mathcal{E}_0 := (\mathcal{E}_N \vee \sigma(V_{N+1}, \dots, V_0)) \cap \mathcal{F}_0$ is I-cosy with respect to \mathcal{F} .*

Proof. We introduce the filtration $\mathcal{G} = (\mathcal{G}_n)_{n \leq 0}$ defined to be equal to \mathcal{F} up to time N and for which (V_{N+1}, \dots, V_0) is an innovation from N to 0 ; precisely, we put

$$\mathcal{G}_n = \begin{cases} \mathcal{F}_n & \text{if } n \leq N; \\ \mathcal{F}_N \vee \sigma(V_{N+1}, \dots, V_n) & \text{if } n \in \{N+1, \dots, 0\}. \end{cases}$$

We can see by lemma 1.6 that \mathcal{F} is immersed in \mathcal{G} . Now we consider a random variable $X \in L^1(\mathcal{E}_0)$, and we shall see that X is I-cosy. Take $\delta > 0$ and take (with lemma 1.1) a random variable $Y_N \in L^1(\mathcal{E}_N)$ such that $\sigma(X) \subset \sigma(Y_N, V_{N+1}, \dots, V_0)$. We put $k = |N| + 1$ and equip \mathbb{R}^k with the ℓ^1 -metric. By lemma 2.15, there exists a Lipschitz function $f: \mathbb{R}^k \rightarrow \mathbb{R}$ such that $\mathbb{E}[|X - R|] < \delta$ where $R = f(Y_N, V_{N+1}, \dots, V_0)$. Let c be a Lipschitz constant for f . As Y_N is I-cosy with respect to the truncated filtration $(\mathcal{F}_n)_{n \leq N}$, there exist, on some probability space $(\overline{\Omega}, \overline{\mathcal{A}}, \overline{\mathbb{P}})$, two jointly immersed isomorphic copies $(\mathcal{F}'_n)_{n \leq N}$ and $(\mathcal{F}''_n)_{n \leq N}$ of $(\mathcal{F}_n)_{n \leq N}$ such that $\overline{\mathbb{E}}[|Y'_N - Y''_N|] < \delta/c$ where Y'_N and Y''_N are the respective copies of Y_N . We introduce the product probability space

$$(\widehat{\Omega}, \widehat{\mathcal{A}}, \widehat{\mathbb{P}}) = (\overline{\Omega}, \overline{\mathcal{A}}, \overline{\mathbb{P}}) \otimes (\Omega, \sigma(V_{N+1}, \dots, V_0), \mathbb{P}),$$

and the canonical embeddings $\iota_1: \overline{\mathcal{A}} \rightarrow \widehat{\mathcal{A}}$ and $\iota_2: \sigma(V_{N+1}, \dots, V_0) \rightarrow \widehat{\mathcal{A}}$ (see example A.2). We use a “hat” to identify random variables and σ -fields through these embeddings: for example we put $\widehat{X}_1 = \iota_1(X_1)$ and $\widehat{X}_2 = \iota_2(X_2)$ for any random variables X_1 and X_2 measurable with respect to $\overline{\mathcal{A}}$ and $\sigma(V_{N+1}, \dots, V_0)$ respectively. Then, for each $n \leq N$ we define $\widehat{\mathcal{G}}'_n = \widehat{\mathcal{F}}'_n$ and $\widehat{\mathcal{G}}''_n = \widehat{\mathcal{F}}''_n$, and for each $n \in \{N+1, \dots, 0\}$ we define the σ -fields

$$\widehat{\mathcal{G}}'_n = \widehat{\mathcal{F}}'_N \vee \sigma(\widehat{V}_{N+1}, \dots, \widehat{V}_n) \quad \text{and} \quad \widehat{\mathcal{G}}''_n = \widehat{\mathcal{F}}''_N \vee \sigma(\widehat{V}_{N+1}, \dots, \widehat{V}_n).$$

Using lemma 2.31, it is a child’s play to verify that $\widehat{\mathcal{G}}' := (\widehat{\mathcal{G}}'_n)_{n \leq 0}$ and $\widehat{\mathcal{G}}'' := (\widehat{\mathcal{G}}''_n)_{n \leq 0}$ are two jointly immersed isomorphic copies of \mathcal{G} . The respective copies of R are $\widehat{R}' = f(\widehat{Y}'_N, \widehat{V}_{N+1}, \dots, \widehat{V}_0)$ and $\widehat{R}'' = f(\widehat{Y}''_N, \widehat{V}_{N+1}, \dots, \widehat{V}_0)$. Thus we have $\widehat{\mathbb{E}}[|\widehat{R}' - \widehat{R}''|] \leq c \widehat{\mathbb{E}}[|\widehat{Y}'_N - \widehat{Y}''_N|]$ because f is c -Lipschitz. Due to isomorphism, we have $\widehat{\mathbb{E}}[|\widehat{Y}'_N - \widehat{Y}''_N|] = \overline{\mathbb{E}}[|Y'_N - Y''_N|]$ and $\widehat{\mathbb{E}}[|\widehat{X}' - \widehat{R}'|] = \widehat{\mathbb{E}}[|\widehat{X}'' - \widehat{R}''|] = \mathbb{E}[|X - R|]$ where \widehat{X}' and \widehat{X}'' are the respective copies of X ; consequently, $\widehat{\mathbb{E}}[|\widehat{X}' - \widehat{X}''|] < 3\delta$, thereby showing that X is I-cosy with respect to \mathcal{G} . As \mathcal{F} is immersed in \mathcal{G} , we see that X is I-cosy with respect to \mathcal{F} (lemma 3.30). \square

Proposition 3.38. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a locally separable filtration. The following conditions are equivalent:*

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

Proof. It is easy to convince oneself that (i) \implies (ii). Obviously, (ii) \implies (iii) is true. We now show that (iii) \implies (i). Let $(V_n)_{n \leq 0}$ be a superinnovation of \mathcal{F} , whose existence is provided by corollary 3.28. We assume that the truncated filtration $(\mathcal{F}_n)_{n \leq N}$ is I-cosy for some N . Then we know from lemma 3.37 that \mathcal{F} is I-cosy because of $\mathcal{F}_0 \subset \mathcal{F}_N \vee \sigma(V_{N+1}, \dots, V_0)$. \square

Remark 3.39. When \mathcal{F} is essentially separable, a result in [12] states that proposition 3.38 is more generally true for a truncation with an \mathcal{F} -stopping time N . We have not attempted to generalize this result to locally separable filtrations.

As another application of lemma 3.37, we shall give a result on I-cosiness for the filtration generated by processes enjoying the properties of the following definition.

Definition 3.40. *Let $(X_n)_{n \leq 0}$ be a process, and let $\phi: -\mathbb{N} \rightarrow -\mathbb{N}$ be a strictly increasing map with $\phi(0) = 0$.*

1. *We say that ϕ 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 ϕ 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 3.41. *Let $(X_n)_{n \leq 0}$ be a process, and let $\phi: -\mathbb{N} \rightarrow -\mathbb{N}$ be a strictly increasing map with $\phi(0) = 0$. If ϕ is a sequence of memory-loss times of type I for $(X_n)_{n \leq 0}$, then ϕ is a sequence of memory-loss times of type II for $(X_n)_{n \leq 0}$.*

In the particular case when $(X_n)_{n \leq 0}$ is a Markov process, this lemma shows that, up to isomorphism, there is a parameterization $(U_n)_{n \leq 0}$ of the filtration \mathcal{F} generated by $(X_n)_{n \leq 0}$ for which $\sigma(X_n) \subset \sigma(X_{n-1}, U_n)$ for every $n \leq 0$. We shall use the following lemma in the proof of lemma 3.41, which is copied from lemma 2.22 in [19].

Lemma 3.42. *Let $\Lambda = \{\Lambda_s\}_{s \in S}$ be a probability kernel from a measurable space S to a Polish space E . Then there exists a measurable function $(s, u) \mapsto \Delta_s(u)$ from $S \times [0, 1]$ to E such that for all $s \in S$, the probability Λ_s is the image of the Lebesgue measure on $[0, 1]$ under the mapping $\Delta_s: [0, 1] \rightarrow E$.*

Proof of lemma 3.41. We can assume that each X_n takes its values in a Polish space, by replacing, if needed, X_n with a real-valued random variable generating the same σ -field. Let $k, n \in -\mathbb{N}$ such that $n \in \{\phi(k-1)+1, \dots, \phi(k)\}$. By lemma 3.42, there exists a measurable function $G_{k,n}(x_{\phi(k-1)}, \dots, x_{n-1}, u)$ such that the function $u \mapsto G_{k,n}(X_{\phi(k-1)}, \dots, X_{n-1}, u)$ almost surely carries the Lebesgue measure to the conditional law $\mathcal{L}(X_n | X_{\phi(k-1)}, \dots, X_{n-1})$. Consider a process $(X'_n, U'_n)_{n \leq 0}$ defined by the following conditions:

- for each time $n \leq 0$, the random variable U_n is independent of the past and has the uniform law $[0, 1]$;
- for each $k \leq 0$, the random variable $X'_{\phi(k)}$ has the same law as $X_{\phi(k)}$;
- we have $X'_n = G_{k,n}(X'_{\phi(k-1)}, \dots, X'_{n-1}, U'_n)$ for each $k, n \in -\mathbb{N}$ satisfying $\phi(k-1) < n \leq \phi(k)$.

Assuming that ϕ is a sequence of memory-loss times of type I for $(X_n)_{n \leq 0}$, it is easy to check that these two conditions uniquely define the law of $(X'_n, U'_n)_{n \leq 0}$, that $(X'_n)_{n \leq 0}$ is a copy of $(X_n)_{n \leq 0}$, and then that ϕ is a sequence of memory-loss times of type II for $(X_n)_{n \leq 0}$. \square

Proposition 3.43. *Let $(X_n)_{n \leq 0}$ be a process, \mathcal{F} the filtration it generates, and $\phi: -\mathbb{N} \rightarrow -\mathbb{N}$ a strictly increasing map with $\phi(0) = 0$, assumed to be a sequence of memory-loss times of type II for $(X_n)_{n \leq 0}$. Then the following conditions are equivalent:*

- (i) \mathcal{F} is I-cosy;
- (ii) for each $n \leq 0$, the σ -field $\sigma(X_{\phi(n)})$ is I-cosy with respect to \mathcal{F} ;
- (iii) for each $n \leq 0$, the σ -field $\sigma(X_{\phi(n)})$ is I-cosy with respect to the truncated filtration $(\mathcal{F}_m)_{m \leq \phi(n)}$.

Proof. One obviously has (i) \implies (ii) \implies (iii). Assuming that (iii) holds for some $n \leq 0$, then it is easy to show with the help of lemma 3.37 that $\sigma(X_{\phi(n)}, X_{\phi(n)+1}, \dots, X_0)$ is I-cosy with respect to \mathcal{F} when ϕ is a sequence of memory-loss times of type II for $(X_n)_{n \leq 0}$. Hence, (iii) \implies (i) follows from the L^1 -closure of the set of I-cosy random variables (lemma 3.33) and lemma 2.12. \square

Remark 3.44. If \mathcal{F} is the filtration generated by a martingale $(M_n)_{n \leq 0}$, it is possible to show that \mathcal{F} is I-cosy if and only if the random variable M_0 is I-cosy. This is deduced from the same result stated in [21] for a general self-joining criterion (see also remark 2.28).

Proposition 3.46 below will be used in the proof of theorem 4.9. Its proof invokes the following lemma, copied verbatim from [14], to which we refer for a proof.

Consequently, letting E be a Polish space and $X \in L^1(\mathcal{F}_0; E)$, the analogous three statements with X instead of \mathcal{E}_0 also are equivalent.

Proof. The consequence follows from proposition 3.36. Inheritability of I-cosiness (lemma 3.30) gives (iii) \implies (i). Let \mathcal{G} be a standard conditionally non-atomic filtration. By lemma 3.22, the product filtration $\mathcal{F} \otimes \mathcal{G}$ is conditionally non-atomic, which shows that (ii) \implies (iii). It remains to show that (i) \implies (ii). Assume \mathcal{E}_0 is I-cosy with respect to \mathcal{F} . As \mathcal{G} is I-cosy (proposition 3.31), the σ -field $\mathcal{E}_0 \otimes \mathcal{G}_0$ is I-cosy with respect to $\mathcal{F} \otimes \mathcal{G}$ by proposition 3.46. In particular, the image of \mathcal{E}_0 under the identification with the first factor is I-cosy with respect to $\mathcal{F} \otimes \mathcal{G}$. \square

Remark 3.48. A stronger result is derived in [21] from the equivalence between I-cosiness and Vershik's standardness criterion: *in the same context of the above corollary, if \mathcal{G} is a locally separable extension of \mathcal{F} , 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} .* We do not know how to prove this result without using Vershik's standardness criterion.

Example: random dynamical systems

We will give in theorem 3.53 a sufficient condition for a stationary Markov process $(X_n)_{n \leq 0}$ to generate an I-cosy filtration. In the proof we give, we will firstly argument that, due to proposition 3.43 and stationarity, it suffices to show I-cosiness of X_0 . Theorem 3.53 firstly requires the stationary process to be *couplable* (definition 3.49). Assuming this condition will allow us to construct some copies $(X'_n)_{n \leq 0}$ and $(X''_n)_{n \leq 0}$ of $(X_n)_{n \leq 0}$ generating jointly immersed filtrations and independent up to some time n_0 sufficiently small for the random variables X'_T and X''_T to be arbitrarily close with high probability for some random time $T \in \{n_0 + 1, \dots, 0\}$. The second condition is the *stochastic self-contractivity* (definition 3.52) of the Markovian kernel. This condition will allow us to maintain the distance between X'_n and X''_n for n going from T to 0.

Definition 3.49. Let (E, ρ) be a Polish metric space. Let $(X_n)_{n \in \mathbb{Z}}$ be a stationary Markov process in E . This process is *couplable* if for every $\delta > 0$, there exist, on some probability space $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$, two jointly immersed copies $(\mathcal{F}'_n)_{n \geq 0}$ and $(\mathcal{F}''_n)_{n \geq 0}$ of the filtration generated by $(X_n)_{n \geq 0}$ such that \mathcal{F}'_0 is independent of \mathcal{F}''_0 and the stopping time $\inf \{n \geq 0 \mid \rho(X'_n, X''_n) < \delta\}$ is almost surely finite.

Example 3.50 (Stationary Markov chain on a denumerable space). Let $(X_n)_{n \in \mathbb{Z}}$ be a stationary Markov process on a denumerable state space, equipped with the 0 – 1 distance. Considering two independent copies $(X'_n)_{n \geq 0}$ and $(X''_n)_{n \geq 0}$ of $(X_n)_{n \geq 0}$, it is known that the so-called product chain $(X'_n, X''_n)_{n \geq 0}$ is recurrent under the assumption that the Markov kernel of $(X_n)_{n \geq 0}$ is *positive recurrent, irreducible, and aperiodic* (see [33]). In particular the product chain almost surely visits the diagonal, and hence $(X_n)_{n \in \mathbb{Z}}$ is couplable under this assumption. In fact, we can see that if $(X_n)_{n \in \mathbb{Z}}$ generates a Kolmogorovian

filtration, then the Markov kernel is positive recurrent, irreducible, and aperiodic. Indeed, it is known (see [19]) that an irreducible Markov kernel is positive recurrent whenever it admits an invariant probability measure, and it is not difficult to check that the Markov kernel is irreducible and aperiodic if $(X_n)_{n \in \mathbb{Z}}$ generates a Kolmogorovian filtration.

Example 3.51 (Random walk on the circle). Let $(X_n)_{n \in \mathbb{Z}}$ be the stationary Markov process on the one-dimensional torus \mathbb{R}/\mathbb{Z} defined as follows. For a given $\alpha \in \mathbb{R}/\mathbb{Z}$:

- X_n has the uniform distribution on the one-dimensional torus \mathbb{R}/\mathbb{Z} ;
- given $X_n = x$, the random variable X_{n+1} takes as possible values $x \pm \alpha$ with equal probability.

When the “step” α is irrational, then the set $\{x + m\alpha \in \mathbb{R}/\mathbb{Z} \mid m \in \mathbb{N}\}$ is a dense subset of the circle \mathbb{R}/\mathbb{Z} . Thus, considering two independent copies $(X'_n)_{n \geq 0}$ and $(X''_n)_{n \geq 0}$ of $(X_n)_{n \geq 0}$, the product chain $(X'_n, X''_n)_{n \geq 0}$ almost surely visits any open set because of the property of recurrence of a random walk on \mathbb{Z}^2 . Hence $(X_n)_{n \in \mathbb{Z}}$ is couplable when the step is irrational.

Before defining stochastic self-contractivity, we need to introduce the following decomposition of probability kernels, which is usual in the theory of random dynamical systems. Let $(X_n)_{n \in \mathbb{Z}}$ be a stationary Markov process in a Polish space E , with transition kernel $\{P_x\}$. According to lemma 3.42, it is always possible to write the kernel as

$$P_x(f) = \int f \circ \nabla_v(x) d\gamma(v) \quad (1)$$

where γ is a probability measure on some Polish space and $(x, v) \mapsto \nabla_v(x)$ is measurable. Given a kernel written in form (1), we can consider a stationary Markov processes $(X'_n, V'_n)_{n \leq 0}$ with probability transition kernel $\{Q_{x,v}\}$ defined by $Q_{x,v}(h) = \int h(\nabla_t(x), t) d\gamma(t)$ and instantaneous law ν defined by $\nu(h) = \int \int h(\nabla_t(x), t) d\gamma(t) d\mu(x)$ where μ is the instantaneous law of $(X_n)_{n \leq 0}$. Thus $(X'_n)_{n \leq 0}$ has the same law as $(X_n)_{n \leq 0}$, the random variable V'_n is distributed according to γ for every n , one has $X'_{n+1} = \nabla_{V'_{n+1}}(X'_n)$ and the process $(V'_n)_{n \leq 0}$ is a superinnovation of the filtration \mathcal{F}' generated by $(X'_n)_{n \leq 0}$ (this is actually our proof of lemma 3.41 in the particular case where ϕ is the identity map).

Definition 3.52. Let (E, ρ) be a Polish metric space. Let $X = (X_n)_{n \in \mathbb{Z}}$ be a stationary Markov process in E with instantaneous distribution μ . If the transition kernel $\{P_x\}$ can be written as in (1) with the property that there exists a probability kernel $\{\Lambda_{x,y}\}$ such that $\Lambda_{x,y}$ is a joining of γ (i.e. the margins are both γ) for all (x, y) and

$$\int \rho(\nabla_v(x), \nabla_{v'}(y)) d\Lambda_{x,y}(v, v') \leq \rho(x, y), \quad (2)$$

then X is said to be stochastically self-contractive.

Example 3.50 continued. Let $(X_n)_{n \in \mathbb{Z}}$ be a stationary Markov process on a denumerable state space. Given any decomposition (1), and setting $A_{x,y}$ to be the joining of γ supported by the diagonal for any x and y , then we see that $(X_n)_{n \in \mathbb{Z}}$ is stochastically self-contractive when we consider the discrete $0 - 1$ distance on the denumerable state space.

Example 3.51 continued. Let $(X_n)_{n \in \mathbb{Z}}$ be the random walk on the circle. The natural decomposition (1) consists in taking $\nabla_v(x) = x + v\alpha$ and γ the law of $\varepsilon_n := \mathbb{1}_{\{X_n = X_{n-1} + \alpha\}}$. Setting $A_{x,y}$ to be the joining of γ supported by the diagonal of $\{0, 1\}^2$ for any x and y , we see that $(X_n)_{n \in \mathbb{Z}}$ is stochastically self-contractive.

Theorem 3.53. *Let $(X_n)_{n \in \mathbb{Z}}$ be a stationary Markov process in a Polish bounded metric space (E, ρ) . If $(X_n)_{n \in \mathbb{Z}}$ is couplable and stochastically self-contractive, then the filtration generated by $(X_n)_{n \leq 0}$ is I-cosy.*

Thus, in view of examples 3.50 and 3.51, we know from this theorem that a Markov chain on a denumerable state space generates an I-cosy filtration whenever this filtration is Kolmogorovian, and we know that the random walk on the circle with an irrational step generates an I-cosy filtration (we still give a remark on this example at the end of this subsection).

Lemma 3.54. *Let $(X_n, V_n)_{n \leq 0}$ be a Markov process such that each X_n and each V_n takes its values in a Polish space and such that the process $(V_n)_{n \leq 0}$ is a superinnovation (definition 3.10) of the filtration \mathcal{F} generated by $(X_n)_{n \leq 0}$ satisfying in addition $\sigma(X_{n+1}) \subset \sigma(X_n) \vee \sigma(V_{n+1})$ for every $n < 0$. We consider a measurable function f_n such that $X_{n+1} = f_n(X_n, V_{n+1})$ for every $n < 0$.*

On $(\overline{\Omega}, \overline{\mathcal{A}}, \overline{\mathbb{P}})$, let \mathcal{F}' be a copy of \mathcal{F} and $(V''_n)_{n \leq 0}$ a sequence of independent random variables having the same law as $(V_n)_{n \leq 0}$. We suppose that the filtration generated by $(V''_n)_{n \leq 0}$ and the filtration \mathcal{F}' are jointly immersed in some filtration $\overline{\mathcal{H}}$.

Let T be a $\overline{\mathcal{H}}$ -stopping time in $-\mathbb{N} \cup \{+\infty\}$. Let $(X''_n)_{n \leq 0}$ be the process defined by

$$\begin{cases} X''_n = X'_n & \text{if } n \leq T; \\ X''_{n+1} = f_n(X''_n, V''_{n+1}) & \text{for } n \text{ from } T \text{ to } -1. \end{cases}$$

Then $(X''_n)_{n \leq 0}$ is a copy of $(X_n)_{n \leq 0}$ and the filtration it generates is immersed in $\overline{\mathcal{H}}$.

Proof. Let $n \in -\mathbb{N}^*$. We denote by $\{P_x^n\}$ a regular version of the conditional law of X_{n+1} given X_n . Let g be a bounded Borelian function. One easily checks that

$$\mathbb{1}_{T > n} \overline{\mathbb{E}}[g(X''_{n+1}) | \overline{\mathcal{H}}_n] = \mathbb{1}_{T > n} P_{X''_n}^n(g).$$

On the other hand, as V''_{n+1} is independent of $\overline{\mathcal{H}}_n$, we have

$$\mathbb{1}_{T \leq n} \bar{\mathbb{E}}[g(X''_{n+1}) | \bar{\mathcal{H}}_n] = \mathbb{1}_{T \leq n} \bar{\mathbb{E}}[g \circ f_n(X''_n, V''_{n+1}) | \bar{\mathcal{H}}_n] = \mathbb{1}_{T \leq n} P_{X''_n}^n(g).$$

Hence, the process $(X''_n)_{n \leq 0}$ is Markovian with respect to $\bar{\mathcal{H}}$ and has the same Markov kernels $\{P_x^n\}$ as the Markov process $(X_n)_{n \leq 0}$.

It remains to check that X''_n has the same law as X_n for all $n \leq 0$. For $m \leq n$, one has

$$\bar{\mathbb{E}}[g(X''_n) \mathbb{1}_{T=m} | \bar{\mathcal{H}}_m] = \mathbb{1}_{T=m} \bar{\mathbb{E}}[h(X'_m, V''_{m+1}, \dots, V''_n) | \bar{\mathcal{H}}_m]$$

where

$$h(\cdot, v_{m+1}, \dots, v_n) = g \circ f_{n-1}(\cdot, v_n) \circ \dots \circ f_{m+1}(\cdot, v_{m+2}) \circ f_m(\cdot, v_{m+1}).$$

But, because $(V''_{m+1}, \dots, V''_n)$ is independent of $\bar{\mathcal{H}}_m$, we see that

$$\bar{\mathbb{E}}[h(X'_m, V''_{m+1}, \dots, V''_n) | \bar{\mathcal{H}}_m] = Q_{X'_m}(g),$$

where $\{Q_x\}$ a regular version of the conditional law of X_n given X_m , and thus we have

$$\bar{\mathbb{E}}[h(X'_m, V''_{m+1}, \dots, V''_n) | \bar{\mathcal{H}}_m] = \bar{\mathbb{E}}[g(X'_n) | \bar{\mathcal{H}}_m].$$

Hence we obtain $\bar{\mathbb{E}}[g(X''_n) \mathbb{1}_{T=m}] = \bar{\mathbb{E}}[g(X'_n) \mathbb{1}_{T=m}]$. As we obviously have $\bar{\mathbb{E}}[g(X''_n) \mathbb{1}_{T \geq n}] = \bar{\mathbb{E}}[g(X'_n) \mathbb{1}_{T \geq n}]$, we finally obtain $\bar{\mathbb{E}}[g(X''_n)] = \bar{\mathbb{E}}[g(X'_n)]$. \square

Proof of theorem 3.53. Let \mathcal{F} be the filtration generated by the stationary Markov process $(X_n)_{n \leq 0}$. To show that \mathcal{F} is I-cosy, it suffices, thanks to proposition 3.43, to prove that for each $n \leq 0$, the σ -field $\sigma(X_n)$, or equivalently (proposition 3.36) the random variable X_n , is I-cosy with respect to the truncated filtration $(\mathcal{F}_m)_{m \leq n}$. We prove this for $n = 0$ only as our construction will obviously adapt to an arbitrary n due to stationarity.

Set $\delta > 0$ and define $\epsilon = \delta / \text{diam}(E)$. As we assume that the stationary Markov process is couplable (definition 3.49), it is possible to find some n_0 small enough and a probability space $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$ with a joining $(\mathcal{F}', \mathcal{F}'')$ of \mathcal{F} independent up to n_0 such that $\mathbb{P}[\bar{T} < +\infty] > 1 - \epsilon$ where \bar{T} is defined by

$$\bar{T} = \begin{cases} \inf \{n \mid n_0 \leq n \leq 0, \rho(X'_n, X''_n) < \delta\} & \text{if this infimum exists;} \\ +\infty & \text{otherwise.} \end{cases}$$

By replacing $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$ with its independent product with a sufficiently rich probability space, we can assume that we have a sequence $(\hat{U}_n)_{n \leq 0}$ of independent random variables \hat{U}_n each uniformly distributed on $[0, 1]$ and that is independent of $\mathcal{F}'_0 \vee \mathcal{F}''_0$. We denote by $\bar{\mathcal{H}}$ the supremum of the filtration $\mathcal{F}' \vee \mathcal{F}''$ with the filtration generated by $(\hat{U}_n)_{n \leq 0}$.

For the sake of convenience, we assume that γ in the decomposition (1) given by the stochastic self-contractivity assumption (2), is the Lebesgue measure on $[0, 1]$. Moreover we write $f(x, u)$ instead of $\nabla_u(x)$. Let $\Lambda_{x,y}$ be given by

the stochastic self-contractivity assumption, and $\Delta_{x,y}$ by lemma 3.42 applied with $A_{x,y}$. We define the processes $(\hat{X}'_n)_{n \leq 0}$, $(\hat{X}''_n)_{n \leq 0}$, $(\tilde{U}'_n)_{n \leq 0}$ and $(\tilde{U}''_n)_{n \leq 0}$ by letting, for $n \leq T$,

$$\begin{cases} (\tilde{U}'_n, \tilde{U}''_n) = \Delta_{X'_{n-1}, X''_{n-1}}(\hat{U}_n), \\ \hat{X}'_n = X'_n, \\ \hat{X}''_n = X''_n, \end{cases}$$

and, for n from T to -1 ,

$$\begin{cases} (\tilde{U}'_{n+1}, \tilde{U}''_{n+1}) = \Delta_{\hat{X}'_n, \hat{X}''_n}(\hat{U}_{n+1}), \\ \hat{X}'_{n+1} = f(\hat{X}'_n, \tilde{U}'_{n+1}), \\ \hat{X}''_{n+1} = f(\hat{X}''_n, \tilde{U}''_{n+1}). \end{cases}$$

Clearly, each of \tilde{U}'_{n+1} and \tilde{U}''_{n+1} is independent of $\overline{\mathcal{H}}_n$, hence each of the filtrations generated by $(\tilde{U}'_n)_{n \leq 0}$ and $(\tilde{U}''_n)_{n \leq 0}$ is immersed in $\overline{\mathcal{H}}$ (lemma 1.6). Therefore, according to lemma 3.54, the processes $(\hat{X}'_n)_{n \leq 0}$ and $(\hat{X}''_n)_{n \leq 0}$ are two copies of $(X_n)_{n \leq 0}$ and generate jointly immersed isomorphic filtrations, and we know that they are independent up to n_0 .

By construction, due to the stochastic self-contractivity (2), we have

$$\mathbb{E}[\rho(\hat{X}'_0, \hat{X}''_0) \mid \mathcal{F}'_n \vee \mathcal{F}''_n] \mathbb{1}_{\overline{T}=n} \leq \rho(X'_n, X''_n) \mathbb{1}_{\overline{T}=n}.$$

Hence we obtain

$$\mathbb{E}[\rho(\hat{X}'_0, \hat{X}''_0) \mid (\mathcal{F}' \vee \mathcal{F}'')_T] \mathbb{1}_{\overline{T} \neq +\infty} \leq \rho(X'_T, X''_T) \mathbb{1}_{\overline{T} \neq +\infty} \leq \delta,$$

and consequently we have $\mathbb{E}[\rho(\hat{X}'_0, \hat{X}''_0) \mathbb{1}_{\overline{T} \neq +\infty}] \leq \delta$. As we have in addition $\mathbb{E}[\rho(\hat{X}'_0, \hat{X}''_0) \mathbb{1}_{\overline{T}=+\infty}] \leq \delta$, we finally obtain $\mathbb{E}[\rho(\hat{X}'_0, \hat{X}''_0)] \leq 2\delta$; so X_0 is I-cosy. \square

Remark on example 3.51. Let $(X_n)_{n \leq 0}$ be the stationary random walk on the circle with an irrational step. We have seen that theorem 3.53 applies and thus we know that the filtration \mathcal{F} generated by $(X_n)_{n \leq 0}$ is I-cosy. Note that this filtration is of local product type: for each $n \leq 0$, the random variable $\varepsilon_n := \mathbb{1}_{\{X_n = X_{n-1} + \alpha\}}$ is an independent complement of \mathcal{F}_{n-1} in \mathcal{F}_n . As ε_n takes two possible values with equal probability, \mathcal{F} is a *dyadic* filtration, a particular case of homogeneous filtrations (defined in the introduction and in definition 4.1). Therefore, according to theorem A stated in the introduction (or to corollary 4.5), \mathcal{F} is actually a filtration of product type. A generating innovation for this filtration is constructed in [22].

3.5 Example continued: split-word processes

We discuss about result 2.43 and we show how to deduce part (a) of theorem 2.39 assuming this result. In fact, in references [20] and [7], result 2.43 is deduced from the following result:

Result 2.43'. *If A is finite, then \mathcal{F} is not I-cosy under (Δ) , unless μ is degenerate.*

Then result 2.43 follows from the fact that any filtration of product type is I-cosy (proposition 3.31). Now we shall prove part (a) of theorem 2.39 assuming result 2.43'. Consider an alphabet (A, \mathfrak{A}, μ) containing at least two letters a and b . Let $f: A \rightarrow \{a, b\}$ be any measurable function such that $f(\mu)$ assigns positive measure to each of a and b . For a given splitting sequence, consider the split-word process $(X_n, \varepsilon_n)_{n \leq 0}$ on A and define the process $(f(X_n), \varepsilon_n)_{n \leq 0}$. This latter is the split-word process on the alphabet $\{a, b\}$ with the same splitting sequence, and we know that its generated filtration is not I-cosy under condition (Δ) on this splitting sequence. By lemma 1.6, we can see that the filtration generated by $(f(X_n), \varepsilon_n)_{n \leq 0}$ is immersed in the one generated by $(X_n, \varepsilon_n)_{n \leq 0}$, and thus $(X_n, \varepsilon_n)_{n \leq 0}$ itself does not generate an I-cosy filtration under (Δ) , due to inheritance of I-cosiness by immersion (lemma 3.30).

4 Theorems

In this section, we restate and prove the theorems stated in the introduction, and we return to the example of split-word processes in order to finish the proof of theorem 2.39. Let us first recall the notion of homogeneous filtrations given in the introduction.

Definition 4.1. *A filtration $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ is homogeneous if there exists an innovation $(V_n)_{n \leq 0}$ of \mathcal{F} such that for each $n \leq 0$, V_n either has a diffuse law or is uniformly distributed on some finite set.*

Thus, any homogeneous filtration is of local product type (definition 2.3), and a conditionally non-atomic filtration (definition 3.1) is a particular homogeneous filtration. But note that no homogeneity in time is required: some V_n may be diffuse, others may take two values, others three values, etc.

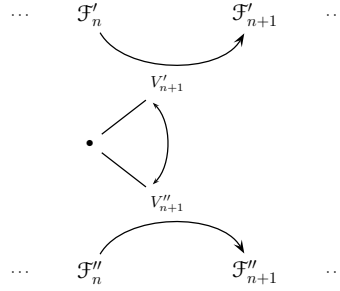
Theorem 4.4 states the equivalence, separately for each random variable, between I-cosiness and Vershik's first level criterion for homogeneous filtrations. As an intermediate step we use Vershik's self-joining criterion, which has been shown to be equivalent to Vershik's first level criterion in theorem 2.38. Under the context of essentially separable filtrations, we have seen that Vershik's first level criterion is equivalent to productness (theorem 2.25), thus theorem A stated in the introduction follows as a consequence of our theorem 4.4. Our generalization to locally separable filtrations has no practical interest;

however, even in the essentially separable case, theorem 4.4 is more precise than theorem A: it asserts that, for a homogeneous filtration, I-cosiness and Vershik's first level criterion are equivalent for a random variable, not only for the whole filtration. For a locally separable filtration, corollary 4.8 states that I-cosiness for a random variable is equivalent to Vershik's first level criterion in a conditionally non-atomic extension of the filtration; this result is still interesting when restricted to the context of essentially separable filtrations. Theorem 4.9 states the equivalence between I-cosiness and standardness or weak standardness, according as we consider essentially separable filtrations or locally separable filtrations.

4.1 Theorems

The key theorem is theorem 4.4. All other theorems stated in the introduction will easily derive therefrom. The key step in the proof is to consider Vershik's self-joining criterion (definition 2.37 and theorem 2.38) as an intermediate step between I-cosiness and Vershik's first level criterion.

The next two lemmas involve joinings $(\mathcal{F}', \mathcal{F}'')$ of \mathcal{F} *permutational after an integer* (definition 2.35), as those appearing in Vershik's self-joining criterion. We recall the picture to be kept in mind for such joinings:



Lemma 4.2. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a filtration of local product type and $(V_n)_{n \leq 0}$ an innovation of \mathcal{F} . We assume that for some integer $n_0 \leq 0$, each innovation V_n is uniformly distributed on a finite set for all $n \in \{n_0 + 1, \dots, 0\}$. Let (E, ρ) be a Polish metric space and $X \in L^1(\mathcal{F}_0; E)$, and let $(\mathcal{F}', \mathcal{F}'')$ be a joining of \mathcal{F} on a probability space $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$.*

Then there exists a filtration \mathcal{F}''' on $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$ such that $(\mathcal{F}', \mathcal{F}''')$ is a joining of \mathcal{F} permutational after n_0 , and such that we have $\bar{\mathbb{E}}[\rho(X', X''')] \leq \bar{\mathbb{E}}[\rho(X', X'')]$ where X' , X'' and X''' are the respective copies of X in \mathcal{F}' , \mathcal{F}'' and \mathcal{F}''' .

Proof. By induction on n , it suffices to show the lemma when assuming that $(\mathcal{F}', \mathcal{F}'')$ is a joining of \mathcal{F} permutational after $n_0 + 1$ for an integer $n_0 < 0$. Let $(V'_n)_{n \leq 0}$ and $(V''_n)_{n \leq 0}$ be the respective copies of $(V_n)_{n \leq 0}$ in \mathcal{F}' and \mathcal{F}'' . Thus we are assuming that $(V'_{n_0+2}, \dots, V'_0)$ and $(V''_{n_0+2}, \dots, V''_0)$ are such that

$V_n'' = T_n(V_n')$ for $n \in \{n_0 + 2, \dots, 0\}$ where the T_n are the $(\mathcal{F}' \vee \mathcal{F}'')_{n-1}$ -measurable random transformations given in definition 2.35.

We shall firstly write the random transformations T_n and the random variables X' and X'' in a convenient form in order to express the conditional expectation $\mathbb{E}[\rho(X', X'') | \mathcal{F}'_{n_0} \vee \mathcal{F}''_{n_0}]$ as a linear function of the conditional law $\mathcal{L}(V'_{n_0+1}, V''_{n_0+1} | \mathcal{F}'_{n_0} \vee \mathcal{F}''_{n_0})$. By lemma 2.36, there are two \mathcal{F}_{n_0+1} -measurable random variables C_{n_0+1}' and D_{n_0+1}'' such that

$$T_n(\cdot) = \psi_0^n(C_{n_0+1}', D_{n_0+1}'', V'_{n_0+2}, \dots, V'_{n-1}, \cdot)$$

for every $n \in \{n_0 + 2, \dots, 0\}$, where the ψ_0^n are measurable functions. As we have

$$(\mathcal{F}' \vee \mathcal{F}'')_{n_0+1} = (\mathcal{F}' \vee \mathcal{F}'')_{n_0} \vee \sigma(V'_{n_0+1}, V''_{n_0+1}),$$

we can take (lemma 1.1) an $(\mathcal{F}'_{n_0} \vee \mathcal{F}''_{n_0})$ -measurable random variable \bar{H}_{n_0} such that $\sigma(C'_{n_0+1}, D''_{n_0+1}) \subset \sigma(\bar{H}_{n_0}, V'_{n_0+1}, V''_{n_0+1})$. On the other hand, we have $\sigma(X) \subset \sigma(X_{n_0}, V_{n_0+1}, \dots, V_0)$ for some \mathcal{F}_{n_0} -measurable random variable X_{n_0} (lemma 1.1). Finally we take a random variable \bar{Z}_{n_0} such that $\sigma(\bar{Z}_{n_0}) = \sigma(\bar{H}_{n_0}, X'_{n_0}, X''_{n_0})$. Thus, we can write

$$T_n(\cdot) = \psi^n(\bar{Z}_{n_0}, V'_{n_0+1}, V''_{n_0+1}, V'_{n_0+2}, \dots, V'_{n-1}, \cdot) \quad (1)$$

for each $n \in \{n_0 + 2, \dots, 0\}$, where ψ^n is measurable, and we can write

$$X' = f(\bar{Z}_{n_0}, V'_{n_0+1}, \dots, V'_0)$$

and

$$\begin{aligned} X'' &= f(\bar{Z}_{n_0}, V''_{n_0+1}, \dots, V''_0) \\ &= f(\bar{Z}_{n_0}, V''_{n_0+1}, T_{n_0+2}(V'_{n_0+2}), \dots, T_0(V'_0)), \end{aligned}$$

for some Borelian function f . We denote by h the function such that

$$X'' = h(\bar{Z}_{n_0}, V'_{n_0+1}, V''_{n_0+1}, V'_{n_0+2}, \dots, V'_0) \quad (2)$$

which is obtained by combining the later equality with (1).

Hence, we have

$$\mathbb{E}[\rho(X', X'') | \mathcal{F}'_{n_0} \vee \mathcal{F}''_{n_0}] = L_{\bar{Z}_{n_0}}(\mathcal{L}(V'_{n_0+1}, V''_{n_0+1} | \mathcal{F}'_{n_0} \vee \mathcal{F}''_{n_0}))$$

where, letting F be the state space of V_{n_0+1} and ν a probability on $F \times F$,

$$L_z(\nu) = \int \mathbb{E}[g_z(v', v'', V'_{n_0+2}, \dots, V'_0)] d\nu(v', v''),$$

where

$$\begin{aligned} g_z(v', v'', V'_{n_0+2}, \dots, V'_0) &= \rho(f(z, v', V'_{n_0+2}, \dots, V'_0), h(z, v', v'', V'_{n_0+2}, \dots, V'_0)). \end{aligned}$$

The set of all probability measures on $F \times F$ having both margins equal to the uniform probability measure on the finite set F is convex; by Birkhoff-Von Neumann's theorem (see [1]), its extreme points are the probability measures supported on graphs of permutations of F . The map $\nu \mapsto L_z(\nu)$ is linear on this convex set, thus it reaches its minimum at such an extremal probability measure. For each z , we measurably select a permutation ϕ_z such that L_z attains its minimum at the measure supported by the graph of ϕ_z .

Then we define $V_{n_0+1}''' = \phi_{\bar{Z}_{n_0}}(V'_{n_0+1})$, and $V_n''' = T'_n(V'_n)$ for $n \in \{n_0 + 2, \dots, 0\}$ where T'_n is obtained by replacing V_{n_0+1}'' with V_{n_0+1}''' in the expression (1) of T_n ; that is, we put

$$V_n''' = \psi_{\bar{Z}_{n_0}, V'_{n_0+1}, V_{n_0+1}''', V'_{n_0+2}, \dots, V'_{n-1}}(V'_n).$$

By lemma 2.4, $(V_{n_0+1}''', \dots, V_0''')$ is an innovation of $\mathcal{F}' \vee \mathcal{F}''$ from n_0 to 0. Finally we define the filtration \mathcal{F}''' as the filtration equaling \mathcal{F}'' up to time n_0 and for which $(V_{n_0+1}''', \dots, V_0''')$ is an innovation from n_0 to 0; that is, we put $\mathcal{F}_n''' = \mathcal{F}_n''$ for $n \leq n_0$ and $\mathcal{F}_n''' = \mathcal{F}_{n_0}''' \vee \sigma(V_{n_0+1}''', \dots, V_n''')$ for $n \in \{n_0 + 1, \dots, 0\}$.

Thus $(\mathcal{F}', \mathcal{F}''')$ is a joining of \mathcal{F} permutational after n_0 . The copy X''' of X in \mathcal{F}''' is obtained by replacing V_{n_0+1}'' by V_{n_0+1}''' in (2), and consequently $\mathbb{E}[\rho(X', X''') \mid \mathcal{F}'_{n_0} \vee \mathcal{F}''_{n_0}]$ is the minimum value of $L_{\bar{Z}_{n_0}}$. So we have

$$\mathbb{E}[\rho(X', X''') \mid \mathcal{F}'_{n_0} \vee \mathcal{F}''_{n_0}] \leq \mathbb{E}[\rho(X', X'') \mid \mathcal{F}'_{n_0} \vee \mathcal{F}''_{n_0}],$$

and then $\mathbb{E}[\rho(X', X''')] \leq \mathbb{E}[\rho(X', X'')]$. \square

Lemma 4.3. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a homogeneous filtration. Let (E, ρ) be a Polish metric space and $X \in L^1(\mathcal{F}_0; E)$. Let $(\mathcal{F}', \mathcal{F}'')$ be a joining of \mathcal{F} on $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$ and $n_0 \leq 0$ an integer.*

Then for any $\epsilon > 0$, there exists a filtration \mathcal{F}''' on $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$ such that $(\mathcal{F}', \mathcal{F}''')$ is a joining of \mathcal{F} permutational after n_0 such that $\mathbb{E}[\rho(X', X''')] \leq \mathbb{E}[\rho(X', X'')] + \epsilon$ where X' , X'' and X''' are the respective copies of X in \mathcal{F}' , \mathcal{F}'' and \mathcal{F}''' .

Proof. Let F be the subset of $\{n_0 + 1, \dots, 0\}$ consisting of those integers n for which V_n is non-atomic. Without loss of generality, we assume that V_n has the uniform law $[0, 1]$ for $n \in F$. Let $\epsilon > 0$. Let k sufficiently large so that, by putting

$$\eta'_n = \sum_{i=0}^{2^k-1} \frac{i}{2^k} \mathbb{1}_{\{\frac{i}{2^k} < V'_n \leq \frac{i+1}{2^k}\}} \quad \text{for } n \in F$$

and $\eta'_n = V'_n$ for $n \in \{n_0 + 1, \dots, 0\} \setminus F$, then there exists a random variable R' measurable with respect to $\mathcal{F}'_{n_0} \vee \sigma(\eta'_{n_0+1}, \dots, \eta'_0)$ and such that $\mathbb{E}[\rho(X', R')] \leq \epsilon/4$. Define a filtration $\mathcal{E}' = (\mathcal{E}'_n)_{n \leq 0}$ by letting $\mathcal{E}'_n = \mathcal{F}'_n$ for $n \leq n_0$ and $\mathcal{E}'_n = \mathcal{F}'_{n_0} \vee \sigma(\eta'_{n_0+1}, \dots, \eta'_n)$ for $n \in \{n_0 + 1, \dots, 0\}$. By applying the preceding lemma to \mathcal{E}' , we obtain some random permutations T_n for

$n \in \{n_0 + 1, \dots, 0\}$ such that by putting $\eta_n''' = T_n(\eta_n')$ for those n , we have $\mathbb{E}[\rho(R', R''')] \leq \mathbb{E}[\rho(R', R'')] where R''' is the copy of R' in the filtration \mathcal{E}''' defined by $\mathcal{E}_n''' = \mathcal{F}_n''$ for $n \leq n_0$ and $\mathcal{E}_n''' = \mathcal{F}_{n_0}'' \vee \sigma(\eta_{n_0+1}', \dots, \eta_n')$ for $n \in \{n_0 + 1, \dots, 0\}$. The T_n are naturally extended to the interval $[0, 1]$ so that we can define $V_n''' = T_n(V_n')$ for every $n \in \{n_0 + 1, \dots, 0\}$. Finally one has $\mathbb{E}[\rho(X', X''')] \leq \mathbb{E}[\rho(X', X'')] + \epsilon$ where X''' is the copy of X' in the filtration \mathcal{F}''' defined by $\mathcal{F}_n''' = \mathcal{F}_n''$ for $n \leq n_0$ and $\mathcal{F}_n''' = \mathcal{F}_{n_0}'' \vee \sigma(V_{n_0+1}', \dots, V_n')$ for $n \in \{n_0 + 1, \dots, 0\}$. $\square$$

Theorem 4.4. *Let $\mathcal{F} = (\mathcal{F}_n)_{n \leq 0}$ be a homogeneous filtration, E a Polish space and $X \in L^1(\mathcal{F}_0; E)$. Then X is I-cosy if and only if X satisfies Vershik's first level criterion. Consequently, a σ -field $\mathcal{E}_0 \subset \mathcal{F}_0$ is I-cosy if and only if it satisfies Vershik's first level criterion, and the filtration \mathcal{F} is I-cosy if and only if it satisfies Vershik's first level criterion.*

Proof. The last sentence is obvious from definitions. We have seen in proposition 3.31 that the 'if' part holds more generally true for any filtration of local product type. The preceding lemma shows that I-cosiness of X implies that X satisfies Vershik's self-joining criterion (definition 2.37), hence the 'only if' part follows from theorem 2.38. \square

Corollary 4.5. *An essentially separable homogeneous filtration is I-cosy if and only if it is of product type.*

Proof. This results from theorem 4.4 and Vershik's first level criterion (theorem 2.25). \square

The following two corollaries justify some terminology introduced in section 3.

Corollary 4.6. *A filtration is weakly standard conditionally non-atomic according to definition 3.19 if and only if it is both weakly standard and conditionally non-atomic.*

Proof. Obviously, a weakly standard conditionally non-atomic filtration is weakly standard and conditionally non-atomic. Conversely, let \mathcal{F} be a weakly standard filtration which is conditionally non-atomic. Then \mathcal{F} is I-cosy by corollary 3.32, and therefore \mathcal{F} satisfies Vershik's first level criterion by theorem 4.4, thus \mathcal{F} is weakly standard conditionally non-atomic according to definition 3.19. \square

Corollary 4.7. *A filtration is standard conditionally non-atomic according to definition 3.1 if and only if it is both standard and conditionally non-atomic.*

Proof. The proof is similar to the one of corollary 4.6 by using corollary 4.5 instead of theorem 4.4. \square

The following corollary of theorem 4.4 gives a characterization of I-cosiness of a random variable with respect to a locally separable filtration. In assertions (ii) and (iii), we commit the same slight abuse of language as in corollary 3.47.

Corollary 4.8. *Let \mathcal{F} be a locally separable filtration and $\mathcal{E}_0 \subset \mathcal{F}_0$ a σ -field. Then the following facts are equivalent.*

- (i) \mathcal{E}_0 satisfies the I-cosiness criterion with respect to \mathcal{F} ;
- (ii) \mathcal{E}_0 satisfies Vershik's first level criterion with respect to the independent product of \mathcal{F} with a standard non-atomic filtration;
- (iii) \mathcal{E}_0 satisfies Vershik's first level criterion with respect to a conditionally non-atomic extension of \mathcal{F} .

Letting E be a Polish space and $X \in L^1(\mathcal{F}_0; E)$, these three statements with \mathcal{E}_0 replaced by X are still equivalent.

Proof. This follows from theorem 4.4 and corollary 3.47. \square

Theorem 4.9. *Let \mathcal{F} be a locally separable filtration. The following assertions are equivalent:*

- (a) \mathcal{F} is I-cosy (definition 3.29);
- (b) The independent product of \mathcal{F} with a standard conditionally non-atomic filtration is weakly standard conditionally non-atomic (definition 3.19);
- (c) \mathcal{F} is weakly standard (definition 3.21).

If in addition \mathcal{F} is essentially separable, then these assertions also are equivalent to:

- (b)' The independent product of \mathcal{F} with a standard conditionally non-atomic filtration is itself standard conditionally non-atomic (definition 3.1);
- (c)' \mathcal{F} is standard (definition 3.6)

Proof. The independent product of a filtration \mathcal{F} with a standard conditionally non-atomic filtration is an extension of \mathcal{F} (lemma 3.5); that shows that (b) \implies (c) and (b)' \implies (c)'. Corollary 3.32 shows that (c) \implies (a) and (c)' \implies (a). It remains to show that (a) \implies (b) and (a) \implies (b)'. A standard conditionally non-atomic filtration is of product type, hence is I-cosy by proposition 3.31. Therefore, if \mathcal{F} is I-cosy, then so is its independent product with a standard conditionally non-atomic filtration in view of proposition 3.46. By lemma 3.22, this product filtration is conditionally non-atomic. Hence, theorem 4.4 shows that (a) \implies (b) and corollary 4.5 and lemma 3.2 show that (a) \implies (b)'. \square

Corollary 4.10. *A filtration is standard if and only if it is weakly standard and essentially separable.*

Proof. We have already noticed that a standard filtration obviously is weakly standard and essentially separable. The converse follows from theorem 4.9. \square

4.2 Example continued: split-word processes

We show how to deduce part (b) of theorem 2.39 assuming result 2.45. Consider a split word process $(X_n, \varepsilon_n)_{n \leq 0}$ on the alphabet $[0, 1]$ equipped with the Lebesgue measure. Let μ be a probability measure on \mathbb{R} and f be the right-continuous inverse of the cumulative distribution function of μ . Then the process $(f(X_n), \varepsilon_n)_{n \leq 0}$ is the split-word process on the alphabet $(\mathbb{R}, \mathfrak{B}_{\mathbb{R}}, \mu)$ with the same splitting sequence as $(X_n, \varepsilon_n)_{n \leq 0}$. By lemma 1.6, the filtration \mathcal{F} generated by $(f(X_n), \varepsilon_n)_{n \leq 0}$ is immersed in the filtration \mathcal{G} generated by $(X_n, \varepsilon_n)_{n \leq 0}$. Consequently, if \mathcal{G} is I-cosy then so is \mathcal{F} (lemma 3.30), and \mathcal{F} is of product type by corollary 4.5. If (Δ) does not hold, we know from result 2.45 that \mathcal{G} is of product type, and consequently is I-cosy (proposition 3.31). Finally \mathcal{F} is of product type if (Δ) does not hold. As a consequence, this is also true if (A, \mathfrak{A}, μ) is Polish because every Polish probability space is Lebesgue isomorphic to a probability space on \mathbb{R} .

A Isomorphisms

This annex gives definitions and elementary lemmas about isomorphisms between probability spaces. The classical definition of an embedding between from a probability space $(\Omega, \mathcal{B}, \mathbb{P})$ into a probability space $(\Omega', \mathcal{A}', \mathbb{P}')$ is given in terms of a Boolean morphism from \mathcal{B}/\mathbb{P} into \mathcal{A}'/\mathbb{P}' which preserves probabilities; such an embedding extends uniquely to random variables (see [2]) and then the definition is equivalently rephrased as follows.

Definition A.1. *Let $(\Omega, \mathcal{B}, \mathbb{P})$ and $(\Omega', \mathcal{A}', \mathbb{P}')$ be two probability spaces. We say that an application $\Psi: L^0(\Omega, \mathcal{B}, \mathbb{P}) \rightarrow L^0(\Omega', \mathcal{A}', \mathbb{P}')$, is an embedding from $(\Omega, \mathcal{B}, \mathbb{P})$ into $(\Omega', \mathcal{A}', \mathbb{P}')$ if the following two conditions hold:*

- (i) *for all integer $n \geq 1$, for all random variables X_1, \dots, X_n on $(\Omega, \mathcal{B}, \mathbb{P})$, and all Borelian applications $f: \mathbb{R}^n \rightarrow \mathbb{R}$, one has $\Psi(f(X_1, \dots, X_n)) = f(\Psi(X_1), \dots, \Psi(X_n))$;*
- (ii) *each random variable X on $(\Omega, \mathcal{B}, \mathbb{P})$ has the same law as $\Psi(X)$.*

The random variable $\Psi(X)$ is also called the *copy* of the random variable X by the embedding Ψ . We shortly say that Ψ is an embedding from $(\Omega, \mathcal{B}, \mathbb{P})$ into $(\Omega', \mathcal{A}', \mathbb{P}')$, and we shortly write $\Psi: \mathcal{B} \rightarrow \mathcal{A}'$. We say that an embedding $\Psi: \mathcal{B} \rightarrow \mathcal{A}'$ is an *isomorphism* from \mathcal{B} onto \mathcal{A}' if it is surjective. It is trivial that an embedding is linear, injective, and continuous for the topology of the convergence in probability. If $X = \mathbb{1}_B$ is the indicator function of an event $B \in \mathcal{B}$, one can verify that $\Psi(X)$ is the indicator function of an event $B' \in \mathcal{A}'$, which we denote by $\Psi(B)$. One easily verifies that the set $\Psi(\mathcal{B}) := \{\Psi(B) \mid B \in \mathcal{B}\}$ is a σ -field and that Ψ defines an isomorphism from $(\Omega, \mathcal{B}, \mathbb{P})$ into $(\Omega', \Psi(\mathcal{B}), \mathbb{P}')$; thus any $\Psi(\mathcal{B})$ -measurable random variable X' has form $X' = \Psi(X)$ for some \mathcal{B} -measurable random variable X . If $\mathcal{B} = \sigma(Y)$ for some random variable Y , it is also easy to see that $\Psi(\mathcal{B}) = \sigma(\Psi(Y))$.

Example A.2. Let $(\Omega, \mathcal{B}, \mathbb{P})$ and $(\Omega^*, \mathcal{C}^*, \mathbb{P}^*)$ be two probability spaces, and let $(\widehat{\Omega}, \widehat{\mathcal{A}}, \widehat{\mathbb{P}}) = (\Omega, \mathcal{B}, \mathbb{P}) \otimes (\Omega^*, \mathcal{C}^*, \mathbb{P}^*)$. The identification with the first factor is the canonical embedding $\iota: \mathcal{B} \rightarrow \widehat{\mathcal{A}}$ defined by $\iota(X): (\omega, \omega^*) \mapsto X(\omega)$.

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

Lemma A.3. *Let $(\Omega, \mathcal{B}, \mathbb{P})$ and $(\Omega', \mathcal{A}', \mathbb{P}')$ be two probability spaces and $\Psi: \mathcal{B} \rightarrow \mathcal{A}'$ be an embedding. Let $X \in L^1(\mathcal{B})$ and $\mathcal{C} \subset \mathcal{B}$ be a σ -field. Then $\Psi(\mathbb{E}[X | \mathcal{C}]) = \mathbb{E}'[\Psi(X) | \Psi(\mathcal{C})]$.*

As shown below, an embedding also defines uniquely a copy of a random variable taking its values in a Polish space (a topological space is said to be *Polish* if it is separable and admits a complete metrization).

Definition A.4. *Let E be a separable metric space, $(\Omega, \mathcal{B}, \mathbb{P})$ and $(\Omega', \mathcal{A}', \mathbb{P}')$ two probability spaces, $X \in L^0(\mathcal{B}; E)$, and $\Psi: \sigma(X) \rightarrow \mathcal{A}'$ an embedding. A random variable $X' \in L^0(\mathcal{A}'; E)$ is denoted by $\Psi(X)$ if one has $\Psi(f(X)) = f(X')$ for every Borelian function $f: E \rightarrow \mathbb{R}$.*

It is straightforward to verify that, given another separable metric space and a Borelian function $g: E \rightarrow F$, one has $\Psi(g(X)) = g(\Psi(X))$ provided that $\Psi(g(X))$ and $\Psi(X)$ exist.

Lemma A.5. *With the same notations as the preceding definition, when E is Polish, there exists a unique random variable $X' = \Psi(X)$.*

Proof. Any Polish probability space is Lebesgue isomorphic to a probability space on \mathbb{R} (see [6], [25], [29]). Hence, there exist a bimeasurable bijection T from a set $E_0 \subset E$ of full \mathbb{P}_X -measure, where \mathbb{P}_X is the law of X , into a set $F_0 \subset \mathbb{R}$ of full μ -measure, where μ is a probability distribution on \mathbb{R} , and T satisfies $T(\mathbb{P}_X) = \mu$. If X' is a random variable such that $\Psi(f(X)) = f(X')$ for all Borelian functions f , then its law is the same as the one of X . Thus the random variables $T(X)$ and $T(X')$ are well-defined and we have $\Psi(T(X)) = T(X')$. It makes sense to take the image under T^{-1} of this random variable and this yields $X' = T^{-1}(\Psi(T(X)))$. Thus there is at most one random variable X' satisfying the desiderata. Finally, putting $X' = T^{-1}(\Psi(T(X)))$ it is easy to verify that the equality $\Psi(f(X)) = f(X')$ is indeed satisfied for each Borelian function $f: E \rightarrow \mathbb{R}$. \square

Example A.6. 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 X_1, \dots, X_n be random variables on $(\Omega, \mathcal{B}, \mathbb{P})$ taking their values in possibly different Polish spaces, and consider (X_1, \dots, X_n) as a random variable in the product Polish space. Then $\Psi(X_1, \dots, X_n) = (\Psi(X_1), \dots, \Psi(X_n))$. Indeed each σ -field $\sigma(X_i)$ is essentially separable so it is possible to write $X_i = g_i(Z_i)$ where Z_i takes its values in \mathbb{R} . Thus, $\Psi(f(X_1, \dots, X_n)) = f(g_1(\Psi(Z_1)), \dots, g_n(\Psi(Z_n))) = f(\Psi(X_1), \dots, \Psi(X_n))$.

Lemma A.7. *Let E be a separable metric space, $(\Omega, \mathcal{B}, \mathbb{P})$ and $(\Omega', \mathcal{A}', \mathbb{P}')$ be two probability spaces, $X \in L^0(\mathcal{B}; E)$ and $X' \in L^0(\mathcal{A}'; E)$. If X and X' have the same law, then $\sigma(X)$ and $\sigma(X')$ are isomorphic and there exists a unique isomorphism $\Psi: \sigma(X) \rightarrow \sigma(X')$ such that $\Psi(X) = X'$ (in the sense of definition A.4).*

Proof. By Doob's functional representation theorem, any real-valued random variable measurable with respect to $\sigma(X)$ is of the form $f(X)$ for a Borelian function $f: E \rightarrow \mathbb{R}$. One easily verifies that we define an isomorphism $\Psi: \sigma(X) \rightarrow \sigma(X')$ by putting $\Psi(f(X)) = f(X')$. \square

Copies of conditional laws

Let E be a Polish space. On $(\Omega, \mathcal{A}, \mathbb{P})$, if X is a random variable taking its values in E , and if $\mathcal{C} \subset \mathcal{A}$ is a σ -field, then the conditional law $\mathcal{L}[X | \mathcal{C}]$ is a \mathcal{C} -measurable random variable taking its values in the set $\mathcal{P}(E)$ of probability measures on E , which is Polish in the *weak topology* (see [4]), generated by the maps $\pi_f: \mu \mapsto \mu(f)$ for bounded continuous functions $f: E \rightarrow \mathbb{R}$. The associated Borel σ -field on $\mathcal{P}(E)$ is itself generated by the maps π_f for bounded continuous functions $f: E \rightarrow \mathbb{R}$. Therefore, the σ -field $\sigma(\mathcal{L}[X | \mathcal{C}])$ is generated by the conditional expectations $\mathbb{E}[f(X) | \mathcal{C}] = \pi_f(\mathcal{L}[X | \mathcal{C}])$ for all bounded continuous functions $f: E \rightarrow \mathbb{R}$, and $\mathbb{E}[f(X) | \mathcal{C}]$ is $\sigma(\mathcal{L}[X | \mathcal{C}])$ -measurable for all suitable functions $f: E \rightarrow \mathbb{R}$.

Lemma A.8. *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 and $\mu: \Omega \times \mathfrak{B}_E \rightarrow [0, 1]$ a probability kernel from (Ω, \mathcal{B}) to \mathfrak{B}_E . Thus μ defines a random variable taking its values in the Polish space of probability measures on E . Then one has $\mu' = \Psi(\mu)$ according to definition A.4 if and only if $\mu'(f) = \Psi(\mu(f))$ for all bounded continuous functions $f: E \rightarrow \mathbb{R}$.*

Proof. If $\mu' = \Psi(\mu)$ then from definition A.4 we know that $\mu'(f) = \Psi(\mu(f))$ for all suitable functions f . Conversely, it is well-known that any measure m on a Polish space E is uniquely determined by the values of $m(f)$ for bounded continuous functions $f: E \rightarrow \mathbb{R}$. \square

The proof of the following lemma is easily derived from lemma A.8; we leave it to the reader.

Lemma A.9. *On a probability space $(\Omega, \mathcal{C}, \mathbb{P})$, let μ 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 one has $\iota(\mu) = \mathcal{L}[\hat{V} | \hat{\mathcal{C}}]$ where $\hat{\mathcal{C}} = \iota(\mathcal{C})$ and \hat{V} is the random variable defined by $\hat{V}(\omega, t) = t$.

Lemma A.10. *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 σ -field. Then $\Psi(\mathcal{L}[X | \mathcal{C}]) = \mathcal{L}[\Psi(X) | \Psi(\mathcal{C})]$.*

Proof. By lemma A.8, it suffices to check that the equalities $\Psi(\mathbb{E}[f(X) | \mathcal{C}]) = \mathbb{E}'[f(\Psi(X)) | \Psi(\mathcal{C})]$ hold for all bounded continuous functions $f: E \rightarrow \mathbb{R}$. This stems from lemma A.3. \square

Proposition A.11. *Let $(\Omega, \mathcal{B}, \mathbb{P})$ be a probability space and let $\mathcal{C}_1, \mathcal{C}_2$ be two sub- σ -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 Ψ_1 and Ψ_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 .*

Proof. The ‘only if’ part follows from lemma A.10. We show the ‘if’ part. Let X be a $\mathcal{C}_1 \vee \mathcal{C}_2$ -measurable random variable. Then, by lemma 1.1, there exist a \mathcal{C}_1 -measurable random variable C_1 , a \mathcal{C}_2 -measurable random variable C_2 , and a Borelian function f such that $X = f(C_1, C_2)$. If Ψ exists, one must have $\Psi(X) = f(\Psi_1(C_1), \Psi_2(C_2))$. The condition $\Psi_1(\mathcal{L}[C_2 | \mathcal{C}_1]) = \mathcal{L}[\Psi_2(C_2) | \Psi_1(\mathcal{C}_1)]$ shows that the pair (C_1, C_2) has the same distribution as $(\Psi_1(C_1), \Psi_2(C_2))$, so X has the same law as $f(\Psi_1(C_1), \Psi_2(C_2))$. To show that Ψ is defined without ambiguity, consider that $X = g(D_1, D_2)$ where D_1 is a \mathcal{C}_1 -measurable random variable, D_2 is a \mathcal{C}_2 -measurable random variable, and g a Borelian function. The assumption implies $\Psi_1(\mathcal{L}[C_2, D_2 | \mathcal{C}_1]) = \mathcal{L}[\Psi_2(C_2, D_2) | \Psi_1(\mathcal{C}_1)]$, which implies that the four-tuple (C_1, D_1, C_2, D_2) has the same distribution as $(\Psi_1(C_1), \Psi_1(D_1), \Psi_2(C_2), \Psi_2(D_2))$; so if $f(C_1, C_2) = g(D_1, D_2)$ almost surely then $f(\Psi_1(C_1), \Psi_2(C_2)) = g(\Psi_1(D_1), \Psi_2(D_2))$ almost surely. Checking condition (i) in definition A.1 is left to the reader. \square

Corollary A.12. *Let $(\Omega, \mathcal{B}, \mathbb{P})$ be a probability space, $\mathcal{C} \subset \mathcal{B}$ a σ -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' an \mathcal{A}' -measurable 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 Ψ_0 and sending V to V' if and only if one has $\Psi_0(\mathcal{L}[V | \mathcal{C}]) = \mathcal{L}[V' | \Psi_0(\mathcal{C})]$.*

Proof. The ‘only if’ part follows from lemma A.10. The ‘if’ part follows from lemma A.7, lemma A.8, and proposition A.11. \square

Corollary A.13. *On $(\Omega, \mathcal{A}, \mathbb{P})$, let $\mathcal{C} \subset \mathcal{A}$ be a σ -field and let V be a random variable. We put $\mu = \mathcal{L}[V | \mathcal{C}]$. Then, with the notations of lemma A.9, there exists a unique isomorphism $\Psi: \mathcal{C} \vee \sigma(V) \rightarrow \hat{\mathcal{C}} \vee \sigma(\hat{V})$ such that the restriction of Ψ to \mathcal{C} equals the canonical embedding ι and $\Psi(V) = \hat{V}$.*

References

1. R.B. Bapat, T.E.S. Raghavan: *Nonnegative Matrices and Applications* (Cambridge University Press, Cambridge 1997)
2. M. Barlow, M. Émery, F. Knight, S. Song, M. Yor: Autour d'un théorème de Tsirelson sur des filtrations browniennes et non-browniennes. Séminaire de Probabilités XXXII (Springer Lectures Notes in Math. 1686, Berlin 1998) pp 264–305
3. S. Beghdadi-Sakrani, M. Émery: On certain probabilities equivalent to coin-tossing, d'après Schachermayer. Séminaire de Probabilités XXXIII (Springer Lectures Notes in Math. 1709, Berlin 1999) pp 240–256
4. P. Billingsley: *Convergence of Probability Measures* (John Wiley & Sons, Inc., New York 1968)
5. J. Blum, D. Hanson: Further Results on the Representation Problem for Stationary Stochastic Processes with Trivial Tail Field. *Journal of Mathematics and Mechanics*, **12**:6, 935–943 (1963)
6. V.I. Bogachev: *Measure Theory*, Volume II (Springer-Verlag, Berlin 2007)
7. G. Ceillier: The filtration of the split-word processes. Preprint (2009)
8. C. Dellacherie, P-A. Meyer: *Probabilités et potentiel*, Chapitres I à IV (Hermann, Paris 1975)
9. J.L. Doob: *Classical Potential Theory and Its Probabilistic Counterpart* (Springer-Verlag, New-York 1984)
10. L.E. Dubins, J. Feldman, M. Smorodinsky, B. Tsirelson: Decreasing sequences of σ -fields and a measure change for Brownian motion. *Annals of Probability*, **24**, 882–904 (1996)
11. R.M. Dudley: *Real Analysis and Probability* (Wadsworth and Brooks/Cole Math Series, Pacific Grove 1989)
12. M. Émery: Old and new tools in the theory of filtrations. In: *Dynamics and Randomness*, ed by A. Maass, S. Martinez, J. San Martin (Kluwer Academic Publishers, Massachusetts 2002) pp 125–146
13. M. Émery: On certain almost Brownian filtrations. *Annales de l'I.H.P. Probabilités et statistiques*, **41**:3, 285–305 (2005)
14. 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, Berlin 2001) pp 265–305
15. J. Feldman, M. Smorodinsky: Decreasing sequences of measurable partitions: product type, standard and prestandard. *Ergodic Theory and Dynamical Systems*, **20**:4, 1079–1090 (2000)
16. J. Feldman, M. Smorodinsky: Addendum to our paper 'Decreasing sequences of sigma fields: product type, standard, and substandard'. *Ergodic Theory and Dynamical Systems*, **22**:4, 1329–1330 (2002)
17. D.L. Hanson: On the Representation Problem for Stationary Stochastic Processes with Trivial Tail Field. *Journal of Mathematics and Mechanics*, **12**:2, 294–301 (1963).
18. D. Hecklen: Bernoullis are standard when entropy is not an obstruction. *Israel Journal of Mathematics*, **107**:1, 141–155 (1998).
19. O. Kallenberg: *Foundations of Modern Probability* (Springer, Berlin Heidelberg New York 1997)
20. S. Laurent: Filtrations à temps discret négatif. PhD Thesis, Université de Strasbourg, Strasbourg (2004)

21. S. Laurent: On Vershikian and I-cosy random variables and filtrations. Preprint (2009)
22. C. Leuridan: Filtration d'une marche aléatoire stationnaire sur le cercle. Séminaire de Probabilités XXXVI (Springer Lectures Notes in Math. 1801, Berlin 2002) pp 335–347
23. P. Major: On the invariance principle for sums of independent identically distributed random variables. *Journal of Multivariate Analysis* **8**, 487–517 (1978)
24. W. Parry: Decoding with Two Independent Processes. In: *Measure and Measurable Dynamics*, Contemporary Mathematics 94, ed by R.D. Mauldin, R.M. Shortt, C.E. Silva (American Mathematical Society, Providence 1989) pp 207–209
25. V. A. Rokhlin: On the fundamental ideas of measure theory. *Amer. Math. Soc. Translation* **71**, 1–53 (1952).
26. M. Rosenblatt: Stationary processes as shifts of functions of independent random variables. *Journal of Mathematics and Mechanics*, **8**:5 (1959)
27. M. Rosenblatt: Stationary Markov Chains and Independent Random Variables. *Journal of Mathematics and Mechanics*, **9**:6, 945–949 (1960)
28. M. Rosenblatt: The Representation of a Class of Two State Stationary Processes in Terms of Independent Random Variables. *Journal of Mathematics and Mechanics*, **12**:5, 721–730 (1963).
29. T. de la Rue: Espaces de Lebesgue. Séminaire de Probabilités XXVII (Springer Lectures Notes in Math. 1557, Berlin 1993) pp 15–21
30. W. Schachermayer: On Certain Probabilities Equivalent to Wiener Measure d'après Dubins, Feldman, Smorodinsky and Tsirelson. Séminaire de Probabilités XXXIII (Springer Lectures Notes in Math. 1709, Berlin 1999) pp 221–239
31. W. Schachermayer: Addendum to the paper 'On Certain Probabilities Equivalent to Wiener Measure d'après Dubins, Feldman, Smorodinsky and Tsirelson'. Séminaire de Probabilités XXXVI (Springer Lectures Notes in Math. 1801, Berlin 2002) pp 493–497
32. M. Smorodinsky: Processes with no standard extension. *Israel Journal of Mathematics*, **107**, 327–331 (1998)
33. H. Thorisson: *Coupling, stationarity, and regeneration*, (Springer-Verlag, New York 2000)
34. B. Tsirelson: About Yor's problem. Unpublished preprint, Tel Aviv University
35. B. Tsirelson: Triple points: from non-Brownian filtrations to harmonic measures. *Geometric And Functional Analysis (GAFA)*, **7**, 1096–1142 (1997)
36. 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)
37. A.M. Vershik: Decreasing sequences of measurable partitions, and their applications. *Soviet mathematics - Doklady*, **11**, 1007–1011 (1970)
38. 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)
39. A.M. Vershik: Approximation in measure theory (in Russian). PhD Thesis, Leningrad University, Leningrad (1973)
40. 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).

41. A.M. Vershik: The theory of decreasing sequences of measurable partitions (in Russian). *Algebra i Analiz*, **6**:4, 1–68 (1994). English translation: *St. Petersburg Mathematical Journal*, **6**:4, 705–761 (1995)