

Martingales associated to peacocks using the curtain coupling

Nicolas Juillet*

Abstract

We consider right-continuous peacocks, that is, families of real probability measures $(\mu_t)_{t \in [0,1]}$ that are increasing in convex order. Given a sequence of time partitions we associate the sequence of martingales characterised by the fact that they are Markovian, constant on the partition intervals $[t_k, t_{k+1}[$, and such that the transition kernels at times t_{k+1} are the curtain couplings of marginals μ_{t_k} and $\mu_{t_{k+1}}$. We study the limit *curtain* processes obtained when the mesh of the partition tends to zero and study existence, uniqueness and relevancy with respect to the original data. For any right-continuous peacock we show there exist sequences of partitions such that a limit process exists (for the finite-dimensional convergence).

Under certain additional regularity assumptions, we prove that there is a unique limit curtain process and that it is a Markovian martingale. We first study by elementary methods peacocks whose marginals correspond to uniform distributions in convex order. In this case, the results and techniques complete the results and techniques used in a parallel work by Henry-Labordère, Tan and Touzi [9]. We obtain the same type of results for all limit curtain processes associated to a class of analytic discrete peacocks, i.e., the measures μ_t are finitely supported and vary analytically in t .

Finally, we give examples of peacocks and sequences of partitions such that the limit curtain process is a non-Markovian martingale.

Keywords: convex order; peacocks; martingales; couplings; optimal transport.

AMS MSC 2010: Primary 60E15; 60G44; 60B10, Secondary 49M25; 60G48; 60J25.

Submitted to EJP on March 15, 2016, final version accepted on January 7, 2018.

Since the seminal article [16] by Hans G. Kellerer it has been known exactly which families $(\mu_t)_{t \in \mathbb{R}}$ of real probability measures are of the form $(\text{Law}(X_t))_{t \in \mathbb{R}}$ where $(X_t)_{t \in \mathbb{R}}$ is a martingale. A simple application of the conditional Jensen inequality gives a necessary condition: the functions $t \mapsto \mu_t$ have to be increasing in convex order. Indeed, if φ is a convex function

$$\int \varphi d\mu_s = \mathbb{E}(\varphi(X_s)) \leq \mathbb{E}(\varphi(X_t)) = \int \varphi d\mu_t.$$

On the other hand, that this condition is also sufficient is the most well-known part of Kellerer's Theorem. However, there is a second very important part to this theorem

*Université de Strasbourg et CNRS, France. E-mail: nicolas.juillet@math.unistra.fr

as attested by the title of his paper: *Markov-Komposition und eine Anwendung auf Martingale*. In fact the martingale $(X_t)_{t \in \mathbb{R}}$ associated to $(\mu_t)_{t \in \mathbb{R}}$ can be chosen to be Markovian and this statement completes Kellerer's Theorem.

Theorem (Kellerer, 1972). *Let $(\mu_t)_{t \in \mathbb{R}}$ be a family of integrable probability measures on \mathbb{R} . The following conditions are equivalent*

- *The family of laws $(\mu_t)_{t \in \mathbb{R}}$ is increasing in convex order,*
- *there exists a Markovian martingale $(X_t)_{t \in \mathbb{R}}$ with $\text{Law}(X_t) = \mu_t$ for every $t \in \mathbb{R}$.*

Kellerer's proof is of a topological and functional analytic nature and does not involve many pure probabilistic arguments. It relies on the compactness of certain sets of measures on finitely many copies of \mathbb{R} . In particular, Kellerer does not explicitly construct the martingale $(X_t)_{t \in \mathbb{R}}$ in his theorem. Quite recently, from a more stochastic point of view – motivated by mathematical finance questions – Hirsch, Profeta, Roynette and Yor devoted a book [11] to defining explicit martingales associated to $(\mu_t)_t$, the existence of which is ensured by Kellerer's Theorem. They studied processes having marginal distributions increasing in convex order, in particular classes of processes that generalise their “guiding example” $X_t = t^{-1} \int_0^t \exp(B_s - s/2) ds$. Such processes were called “Processus Croissants pour l'Ordre Convexe”, which has been abbreviated to PCOC and is pronounced “peacock”. In the present paper we investigate a particular martingale construction based on optimal transport.

In keeping with the terminology in later papers [12, 13, 9, 4] where the initial data are families of measures $(\mu_t)_{t \in \mathbb{R}}$ instead of processes, any family of measures increasing in convex order will also be called a *peacock*. In this paper we restrict to peacocks indexed on a compact interval, which most of the time will be $[0, 1]$. We furthermore restrict to right-continuous peacocks (for the weak topology), making it possible to consider càdlàg martingales in this context.

The goal of the present paper is to focus on one particular martingale construction based on (martingale) optimal transport theory. For every peacock $(\mu_t)_{t \in [0,1]}$ we construct a Markovian martingale $(X_t)_{t \in [0,1]}$ attached to a discrete grid $0 \leq t_0 < t_1 < \dots < t_N = 1$ such that

- $\text{Law}(X_t) = \mu_{t_k}$ for every $t \in [t_k, t_{k+1}[$, where $k \in \{0, \dots, N-1\}$,
- $\text{Law}(X_{t_k}, X_{t_{k+1}})$ is the (left-)curtain coupling of μ_{t_k} and $\mu_{t_{k+1}}$ (this coupling comes from martingale transport theory and will be introduced and commented in Subsection 1.1).

Using sequences of these interpolating martingales we investigate the convergence of these processes when the mesh of the grid tends to zero. If a limit process exists we will call it a *limit curtain process*. Such a limit process may not exist, may not be unique if it does and depends on the convergence topology. We say that a limit curtain process of $(\mu_t)_{t \in [0,1]}$ is *relevant* if it is a martingale for its own filtration, has the proper marginals, i.e., $\text{Law}(X_t) = \mu_t$ at every time $t \in [0, 1]$ and is concentrated on càdlàg paths.

Our approach is parallel to the very recent paper by Henry-Labordère, Tan, and Touzi (see [9] and Remark 0.2) where the authors also choose curtain couplings for defining the transitions in the martingales. Our motivation for using the curtain coupling is explained in detail at the end of Subsection 1.1. For now let us just mention that it is canonical in many respects including optimality in relation with optimal transportation theory and monotonicity (see Proposition 1.10). Moreover, the curtain coupling is one of the most simple martingale couplings to work with, if not the unique one.

Our first theorem is on the existence and the relevancy of limit curtain processes.

Theorem A (Propositions 2.4 and 2.6). *Given any right-continuous peacock $(\mu_t)_{t \in [0,1]}$, for the finite dimensional topology (see Subsection 2.4) the set of limit curtain processes is not empty and contains at least one relevant process $(X_t)_{t \in [0,1]}$, that is, a càdlàg martingale with $\text{Law}(X_t) = \mu_t$ at every time.*

For the Skorokhod topology all limit curtain processes are relevant (but it is not known whether or not this set is empty).

Remark 0.1. Our proof of Theorem A is new but the techniques, in particular the technique of considering a martingale on a countable subset of $[0, 1]$, have been used several times in the peacock literature [3, 8, 12] or on related topics since at least 1978 [15].

We study some representative examples of peacocks, both continuous and discrete, where continuous or discrete refer to the distributions μ_t involved in the peacocks. For some of them, we prove the existence and uniqueness of the limit curtain process in the Skorokhod topology, for arbitrary sequences of partitions.

The following result on peacocks for uniform measures is obtained by using direct methods.

Theorem B (Proposition 3.1, Subsection 2.3 and end of Section 3). *Let μ_t be the uniform measure on $[-f(t), f(t)]$, where f is a positive, increasing, and continuous function on $[0, 1]$. There exists a unique limit curtain process of $(\mu_t)_{t \in [0,1]}$ and it is Markovian.*

The process consists of piecewise increasing trajectories with jumps downwards at random arrival times t to the bottom $-f(t)$ of the interval. The intensity of jumps is independent of the position and happens with intensity $2^{-1} d \log(f(t))/dt$.

For $f(t) = \exp(2t)/2$, the set of arrival times for the jumps is a standard Poisson point process.

We now state our results on the existence and the uniqueness of limit curtain processes for discrete peacocks. This type of result is not typical in the peacock literature because the measures μ_t are often absolutely continuous or smooth, for instance the result of a diffusion, and discrete measures are usually seen as general measures that may be obtained as limits.

Theorem C (Theorem 4.4). *Let $(\mu_t)_{t \in [0,1]}$ be a peacock of type $\mu_t = \sum_{i=1}^n a_i(t) \delta_{x_i(t)}$ where $t \mapsto (x_1, \dots, x_n, a_1, \dots, a_n)(t)$ is real analytic and furthermore satisfies*

$$a_1(t), \dots, a_n(t) > 0 \quad \text{and} \quad x_1(t) < \dots < x_n(t).$$

Then, there is a unique limit curtain process and this process is Markovian.

Finally, we give examples of peacocks with non-Markovian limit curtain processes. This may occur both in the discrete and continuous situation.

Theorem D (Sections 5 and 6). *There exist a peacock $(\mu_t)_{t \in [0,1]}$ and a limit curtain process $(X_t)_{t \in [0,1]}$ of $(\mu_t)_{t \in [0,1]}$ such that $\text{Law}(X_t) = \mu_t$ for every t , but $(X_t)_{t \in [0,1]}$ is not Markovian. In fact $(\mu_t)_{t \in [0,1]}$ can be taken either to be discrete or continuous.*

Remark 0.2. In the parallel paper by Henry-Labordère, Tan and Touzi [9], the authors consider *continuous* peacocks satisfying certain assumptions (Assumption 3.1 and Assumption 3.4). In [9, Proposition 3.2] under Assumption 3.1 they prove the existence of a limit curtain process in the Skorokhod topology. This can be compared to our Theorem A: Their convergence is for a stronger topology but the result requires stronger assumptions as well.

The second main result is Theorem 3.7. It holds under Assumptions 3.1 and 3.4. and states the uniqueness and a characterisation by a SDE of the limit curtain coupling.

This limit is a local Lévy process. Our closer result is Theorem B. The peacocks that we consider there do not satisfy Assumptions 3.1 and 3.4. Neither does our main peacock $[\mu]$, with μ_t uniform on $[-1/2 \exp(2t), 1/2 \exp(2t)]$, satisfy Assumption 3.4 (iii) (but it satisfies Assumption 3.1). Nevertheless, this peacock is not far from entering the setting of [9] and its limit curtain process appears prototypical for the local Lévy processes of Henry-Labordère, Tan and Touzi. From this perspective the class of peacocks considered in [9] may appear more general than our uniform peacocks (peacocks with uniform laws). Note finally that the methods used in the proofs are different.

Finally the third main result of [9] is in Section 3.3. It is related to a characterisation of the limit curtain process as the optimiser of a continuous-time martingale transport problem and of its dual problem. It completely departs from our Theorems C and D. In particular, discrete peacocks are not addressed in [9].

Our paper is organised as follows. In Section 1 we introduce the necessary definitions related to martingale transport and convex order, and make explicit the relation between the peacock problem and optimal transport theory. The next sections are devoted to proving the theorems of the introduction: Theorem A in Section 2, Theorem B in Section 3, Theorem C in Section 4 and Theorem D in Section 5 and 6. We conclude the paper with two remarks and by suggesting some open questions (Section 7).

1 Reminders about the convex order and the curtain coupling

On \mathbb{R} , let \mathcal{P} be the space of probability measures and \mathcal{M} the space of positive measures with finite first moments. We denote by \mathcal{P}_1 the space $\mathcal{P} \cap \mathcal{M}$ of probability measures with finite first moments. For every metric space (S, ρ) , we denote by $\mathcal{P}(S)$, $\mathcal{M}(S)$ and $\mathcal{P}_1(S)$ the corresponding spaces. On $\mathcal{P}_1(S)$ we denote by $\mathcal{T}_{cb}(S)$ the usual weak topology of $\mathcal{P}(S)$ induced by the continuous bounded functions.

A measure π on \mathbb{R}^2 with marginals $\mu, \nu \in \mathcal{M}$ is called a *transport plan* from μ to ν or a *coupling* of μ and ν . Let $\Pi(\mu, \nu)$ be the space of transport plans with marginals μ and ν . For probability measures, the subspace $\Pi_M(\mu, \nu) \subseteq \Pi(\mu, \nu) \cap \mathcal{M}(\mathbb{R}^2)$ is defined as follows

$$\Pi_M(\mu, \nu) = \{\pi = \text{Law}(X, Y) \in \Pi(\mu, \nu) : \mathbb{E}(Y|X) = X\},$$

where the constraint $\mathbb{E}(Y|X) = X$ means: $\mathbb{E}(Y|X = x) = x$ for μ -almost every $x \in \mathbb{R}$. If $\mu(\mathbb{R}) = \nu(\mathbb{R}) = m$, one defines this space by normalising

$$\Pi_M(\mu, \nu) = \{\pi \in \Pi(\mu, \nu) : m^{-1}\pi \in \Pi_M(m^{-1}\mu, m^{-1}\nu)\}.$$

Elements of $\Pi_M(\mu, \nu)$ are called *martingale transport plans* or *martingale couplings*. The space $\bigcup_{\mu, \nu \in \mathcal{P}_1} \Pi_M(\mu, \nu)$ is a closed set of $\mathcal{P}(\mathbb{R}^2)$.

For a separable metric space (S, ρ) , the Prokhorov distance Prok_S metrises $\mathcal{T}_{cb}(S)$. For $\mu, \nu \in \mathcal{P}(S)$, the distance $\text{Prok}_S(\mu, \nu)$ is defined as the infimum of those ε such that the inequalities

$$\mu(A) \leq \nu(A^\varepsilon) + \varepsilon \quad \text{and} \quad \nu(A) \leq \mu(A^\varepsilon) + \varepsilon$$

hold for every Borel set A and $A^\varepsilon = \{y \in S, \rho(y, A) \leq \varepsilon\}$. According to Dudley–Strassen theorem [7, Theorem 6.9], we have the equivalent definition.

$$\begin{aligned} \text{Prok}_S(\mu, \nu) &= \inf\{\varepsilon > 0 : \exists \pi \in \Pi(\mu, \nu), \pi(\{(x, y) \in S^2, |y - x| \leq \varepsilon\}) \geq 1 - \varepsilon\} \\ &= \inf\{\varepsilon > 0 : \exists(\Omega, \mathbb{P}), \exists X \sim_{\mathbb{P}} \mu, \exists Y \sim_{\mathbb{P}} \nu, \mathbb{P}(|Y - X| \leq \varepsilon) \geq 1 - \varepsilon\}. \end{aligned}$$

We will use the Prokhorov distance for two types of spaces. First, the space $(\mathbb{R}^j, \|\cdot\|_\infty)$ in relation with the finite dimensional convergence, in Theorem A. Second, the Skorokhod space $D[0, 1]$ of càdlàg real paths defined on $[0, 1]$ with the Skorokhod distance ; see for instance [7, Chapter 3]) where it is defined.

1.1 Definitions of the shadows and the curtain coupling

The aim of this subsection is to introduce the (left-)curtain coupling, which is the essential feature of our approach to the peacock problem. Its definition is based on the shadow projection that is defined in Definition 1.3 and for which we recall important properties and examples.

Definition 1.1 (Usual, convex and extended convex order on \mathcal{M}). *Let μ and ν be measures in \mathcal{M} . The measure ν is greater than μ in convex order if for any convex function $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ we have*

$$\int \varphi d\mu \leq \int \varphi d\nu.$$

We write $\mu \preceq_C \nu$.

The measures are in extended order and we denote it by $\mu \preceq_E \nu$ if the previous relation holds for every nonnegative convex function. If it holds for all nonnegative functions the measures are said in usual order and we denote it by $\mu \leq \nu$. Finally if it holds for every bounded increasing function the measures are said in stochastic order and we denote it by $\mu \preceq_{sto} \nu$.

Proposition 1.2. *For $\mu, \nu \in \mathcal{M}$ the relation $\mu \preceq_C \nu$ holds if and only if $\Pi_M(\mu, \nu)$ is not empty. Moreover, $\mu \preceq_E \nu$ holds if and only if there exist η with $\mu \preceq_C \eta$ and $\eta \leq \nu$.*

Proof. The first assertion is a famous theorem of Strassen [26, Theorem 8] applied to dimension one. The second is Proposition 4.4 in [5]. □

In [5, Lemma 4.6] the following important theorem-definition is proven.

Definition 1.3 (Definition of the shadow). *If $\mu \preceq_E \nu$, there exists a unique measure η such that*

- $\mu \preceq_C \eta$,
- $\eta \leq \nu$,
- if η' satisfies the two first conditions, i.e., $\mu \preceq_C \eta' \leq \nu$, one has $\eta \preceq_C \eta'$.

This measure η is called the shadow of μ in ν and we denote it by $S^\nu(\mu)$.

The shadows are sometimes difficult to determine. An important fact is that they have the smallest variance among the measures η' . Indeed, $\eta \preceq_C \eta'$ implies $\int x d\eta = \int x d\eta'$ and $\int x^2 d\eta \leq \int x^2 d\eta'$ with equality if and only if $\eta = \eta'$ or $\int x^2 d\eta = +\infty$. For every $\mu \in \mathcal{P}$ we denote by G_μ the quantile function of μ . Recall that it is the unique left-continuous increasing function from $]0, 1[$ to \mathbb{R} such that $(G_\mu)_\# \lambda_{[0,1]} = \mu$, where $\#$ is the push-forward.

Example 1.4 (Shadow of an atom, Example 4.7 in [5]). Let δ be an atom of mass α at a point x . Assume that $\delta \preceq_E \nu$. Then $S^\nu(\delta)$ is the restriction of ν between two quantiles, more precisely it is $\nu' = (G_\nu)_\# \lambda_{]s, s']}$ where $s' - s = \alpha$ and the barycentre of ν' is x .

The next lemma describes the tail of the shadows. We denote by spt the support of a positive measure, i.e., the smallest closed set of full measure.

Lemma 1.5. *Let $\mu, \nu \in \mathcal{M}$ satisfy $\mu \preceq_E \nu$. Assume that $x = \sup[\text{spt } \mu]$ is finite and denote $\text{sup}[\text{spt}(S^\nu(\mu))]$ by $y \in \mathbb{R} \cup \{+\infty\}$. Then*

$$S^\nu(\mu)([x, y]) = \nu([x, y]).$$

Proof. The statement $S^\nu(\mu)([x, y]) = \nu([x, y])$ is equivalent to the fact that $S^\nu(\mu)|_{[x, +\infty[}$ is the stochastically leftmost measure θ among the measures of the same mass satisfying $\theta \leq \nu|_{[x, +\infty[}$ (see [14, Subsection 1.2]). This is a reformulation of Lemma 2.3

in [14] because $S^\nu(\mu)|_{[x,+\infty[} = (S^\nu(\mu) - \mu)_+|_{[x,+\infty[} + \min(\mu(x), \nu(x))\delta_x$ and $\nu|_{[x,+\infty[} = (\nu - \mu)_+|_{[x,+\infty[} + \min(\mu(x), \nu(x))\delta_x$. \square

The following result is one of the most important on the structure of shadows. It is Theorem 4.8 of [5].

Proposition 1.6 (Structure of shadows). *Let γ_1, γ_2 and ν be elements of \mathcal{M} and assume that $\mu = \gamma_1 + \gamma_2 \preceq_E \nu$. Then we have $\gamma_2 \preceq_E \nu - S^\nu(\gamma_1)$ and*

$$S^\nu(\gamma_1 + \gamma_2) = S^\nu(\gamma_1) + S^{\nu - S^\nu(\gamma_1)}(\gamma_2).$$

Example 1.7 (Shadow of a finite sum of atoms). Let μ be the measure $\sum_{i=1}^n \alpha_i \delta_{x_i}$ and $\nu = G_{\#} \lambda_{]0, m]}$ such that $\mu \preceq_E \nu$. We can apply Proposition 1.6 to this sum as well as Example 1.4 on the shadow of one atom. We obtain recursively the following description. There exists an increasing sequence of sets $J_1 \subseteq \dots \subseteq J_n \subseteq]0, m]$ satisfying that J_k has measure $\sum_{i=1}^k \alpha_i$ and $J_k \setminus J_{k-1}$ is a pseudo-interval of $]0, m] \setminus J_{k-1}$, that is $J_k \setminus J_{k-1} =]s, t] \setminus J_{k-1}$ for some $0 \leq s, t \leq m$. These sets satisfy $S^\nu(\sum_{i=1}^k \alpha_i \delta_{x_i}) = G_{\#} \lambda_{J_k}$ for every $k \leq n$.

Conversely, any increasing sequence $(J_i)_{i=1, \dots, n}$ such that $J_k \setminus J_{k-1}$ is a pseudo-interval of $]0, m] \setminus J_{k-1}$ is associated with a family of atoms $(\alpha_i \delta_{x_i})_i$ such that $\alpha_i = \lambda(J_i) - \lambda(J_{i-1})$, and point x_i is the barycentre of $G_{\#} \lambda_{J_i \setminus J_{i-1}}$ where $G_{\#} \lambda_{J_k}$ is the shadow of $\sum_{i=1}^k \alpha_i \delta_{x_i}$ in ν .

With the shadow projections, we can introduce the left-curtain coupling. Notice that for an atomic measures μ it is directly related to Example 1.7 where $(x_i)_i$ has to be an increasing sequence.

Definition 1.8 (Left-curtain coupling, Theorem 4.18 in [5]). *Let $\mu, \nu \in \mathcal{M}$ satisfy $\mu \preceq_C \nu$. There exists a unique measure $\pi \in \Pi_M(\mu, \nu)$ such that for any $x \in \mathbb{R}$ the measure $\pi_{]-\infty, x] \times \mathbb{R}}$ has first marginal $\mu_{]-\infty, x]}$ and second marginal $S^\nu(\mu_{]-\infty, x]}\pi_{|c}$ and call it left-curtain coupling or simply curtain coupling.*

One of the main theorems of [5] is the equivalence of three properties of couplings: left-curtain, left-monotone and optimal. Let us define left-monotone couplings.

Definition 1.9 (Left-monotone coupling). *Let π be an element of $\Pi_M(\mu, \nu)$. The coupling π is left-monotone if there exists a Borel set Γ with*

- $\pi(\Gamma) = 1$,
- for every $(x, y^-), (x, y^+)$ and (x', y') elements of Γ satisfying $x < x'$ and $y^- < y^+$, the real y' is not an element of $]y^-, y^+[$.

We can now state the result.

Proposition 1.10 (Theorem 1.9 in [5]). *Let $\pi \in \Pi_M(\mu, \nu)$. Let $c : (x, y) \in \mathbb{R}^2 \rightarrow c(x) \in \mathbb{R}$ be smooth and satisfy $\partial_x \partial_y \partial_y c < 0$. Assume moreover $c \in L^1(\pi)$. Then the properties are equivalent.*

- Left-curtain: the transport plan π is the left-curtain coupling,
- Left-monotone: the transport plan π is left-monotone,
- Optimal: for any $\tilde{\pi} \in \Pi_M(\mu, \nu)$, if $\tilde{\pi} \neq \pi$, then $\int c d\pi < \int c d\tilde{\pi}$.

Remark 1.11. • The cost function $c : (x, y) \rightarrow (1 - \tanh(x))\sqrt{1 + y^2}$ satisfies the assumptions $\partial_x \partial_y \partial_y c < 0$ and $c \in L^1(\pi)$ for every $\pi \in \Pi_M(\mu, \nu)$. The cost functions $-xy^2$ or $(y - x)^3$ satisfy $\partial_x \partial_y \partial_y c < 0$.

- The original condition in [5] is actually $c(x, y) = \phi(x)\psi(y)$ with $\partial_x\phi < 0$ and $\partial_y\psi > 0$, or $c(x, y) = f(y - x)$ with $f^{(3)}(x) > 0$. The condition $\partial_x\partial_y\partial_y c < 0$ appeared in [10] adapting to the martingale setting the twist condition $\partial_x\partial_y c < 0$, also called Spence–Mirrlees condition in the economics and finance literature.

1.2 Pertinence of the curtain coupling and a possible martingale analogue of the constrained Benamou–Brenier transport problem

The next paragraphs aim to introduce more background both on the martingale and non-martingale (or classical) optimal transport theory. We motivate our intension to develop the theory of the curtain coupling, notably in relation with peacocks, from the fact that it is the natural counterpart of this rich and well-developed classical theory. We believe that the curtain coupling could play the same central role in the martingale setting as the quantile coupling does in the classical setting. We stress that the curtain coupling is defined for all pairs $\mu \preceq_C \nu$ and it can be conveniently worked with, which is rare for martingale couplings.

The classical optimal transport problem is the minimisation problem $\pi \mapsto \int c d\pi$ where π varies in $\Pi(\mu, \nu)$. For specific cost functions c the minimiser is uniquely determined. One of the most studied situation is that of a smooth $c : \mathbb{R}^2 \rightarrow \mathbb{R}$ with $\partial_x\partial_y c < 0$, e.g. $c(x, y) = (y - x)^2$ or $-xy$. The two latter examples amount to the maximisation of the covariance for joint laws of μ and ν . In general if $\partial_x\partial_y c < 0$ the minimiser is the celebrated quantile coupling, also known as covariant coupling (see, e.g., [25]). Still in case $\partial_x\partial_y c < 0$, if μ is continuous (its atomic part is zero), the optimal transport plan can be defined as follows by the increasing map $T : x \in \mathbb{R} \rightarrow G_\nu \circ F_\mu(x) \in \mathbb{R}$, where we recall that G_ν is the quantile function of ν and F_μ is the cumulative distribution function of μ . Denoting by $f \otimes g$ the function $x \mapsto (f(x), g(x))$, we have $\pi = (\text{Id} \otimes T)_{\#}\mu$. In particular π is concentrated on the graph of T . In probabilistic terms, if X is a random variable of law μ , the optimal transport π is the deterministic coupling $\text{Law}(X, T(X))$. In the martingale transport problem, in case $\partial_x\partial_y\partial_y c < 0$, when μ is continuous, the shape of the optimal coupling is very similar. In fact the curtain coupling can be written $a(x)(\text{Id} \otimes T_1)_{\#}\mu + b(x)(\text{Id} \otimes T_2)_{\#}\mu$ where $a + b = 1$ and $x = a(x)T_1(x) + b(x)T_2(x)$ with $T_1(x) \leq x \leq T_2(x)$. Moreover, as a consequence of the left-monotonicity, (recall Definition 1.9) T_2 is increasing and for $x' > x$, $T_1(x')$ must not be in the interval $]T_1(x), T_2(x)[$; see [10, 5] for more explicit details. Compared with Proposition 1.10 and Remark 1.11, it appears that the curtain coupling is the natural martingale transport counterpart to the quantile coupling: A first observation is that these are optimal transport plans for cost functions satisfying the same condition up to one derivative or having the same expression up to one power degree ; a second observation is that these couplings are characterised by combinatorial monotonicity conditions on the maps T_1, T_2 or T ; and a third observation is that they can be defined using shadows for the first-order or second-order stochastic dominance (\preceq_{sto} or \preceq_C , respectively). In the last assumption we mean that the quantile coupling can be defined specifying that the measure $\mu_{]-\infty, x]}$ is transported on the lowest measure for \preceq_{sto} among the measures η that have mass $\mu_{]-\infty, x]}$ and satisfy $\eta \leq \nu$.

A very important tool in the optimal transport theory is the *displacement interpolation* of two probability measures $\mu = \mu_0$ and $\nu = \mu_1$, introduced by McCann in [23]. In the case of the real line or of a Euclidean space it is simply the interpolation by the curve $(\mu_t)_{t \in [0,1]}$ defined by $\mu_t = \text{Law}((1-t)X + tY)$ where $\pi = \text{Law}(X, Y)$ is an optimal transport plan from μ to ν . In Euclidean spaces the process $t \mapsto (1-t)X + tY$ is a minimiser of a famous dynamical transport problem introduced by Benamou and Brenier [6]. This dynamical problem makes sense on a wide class of geodesic metric spaces. In probabilistic terms, the problem is to find a process $(X_t)_{t \in [0,1]}$ with $\text{Law}(X_0) = \mu, \text{Law}(X_1) = \nu$

minimising a certain action, which is $\mathcal{A} : X \mapsto \mathbb{E} \int_0^1 \left| \frac{d}{dt} X_t \right|^2 dt$ in Euclidean spaces. For a minimiser $(X_t)_{t \in [0,1]}$, the curve $(\mu_t)_{t \in [0,1]}$ defined by $\mu_t = \text{Law}(X_t)$ is a displacement interpolation between μ and ν . In view of the success of the displacement interpolation, it is natural to search for an appropriate *martingale* displacement interpolation. In [3], the Skorokhod embedding problem has been approached via optimal transport methods and the martingales obtained in this paper can be regarded as martingale displacement interpolations.

From the perspective of optimal transport, the interpretation of the peacock problem is slightly different. In our problem not only μ_0 and μ_1 are fixed as in the Skorokhod problem or the Benamou–Brenier dynamical problem but also all the intermediate marginals. However, for a given curve $(\nu_t)_{t \in [0,1]}$, the Benamou–Brenier minimisation problem still makes sense under the additional constraint $\text{Law}(X_t) = \nu_t$ for every t . This has been studied by Ambrosio–Gigli–Savaré and Lisini [1, 19] in the setting of metric spaces. These authors proved, among other results, that the minimal value $\mathcal{A}(X)$ in this Benamou–Brenier problem under marginal constraints is the energy \mathcal{E}_2 of the curve $(\nu_t)_{t \in [0,1]}$, defined in a suitable sense by $\mathcal{E}_2(\nu) = \int_0^1 |\dot{\nu}_t|^2 dt$. It is shown in [19, proof of Theorem 5] that the minimising process can be approximated using a discretisation of time and the optimal transport plans between ν_{t_k} and $\nu_{t_{k+1}}$ for $t_k = k/2^n$ and $k = 0, \dots, 2^n - 1$. Therefore, both our goal to define a process with prescribed marginals, and our method to consider optimal transport plans for the measures indexed on a partition, are similar to those appearing in optimal transport theory. From this perspective our construction of a martingale associated to a peacock can be regarded as the martingale analogue of the classical optimal transport theory developed in [1, 19]. Note finally that in [9], the authors are able to define a relevant martingale counterpart \mathcal{A}^* to the action \mathcal{A} ; (see [9, Equation (2.5)] where it presented as a maximisation problem) and Remark 0.2 of the present paper. Nevertheless it is neither known which peacocks and processes minimise \mathcal{A}^* for only μ_0, μ_1 fixed, nor what are the martingale transport plans in $\Pi_M(\mu_0, \mu_1)$ derived from it.

For the sake of completeness, we finally mention that a different optimal transport problem with infinitely many marginals on \mathbb{R} has been studied by Pass [24]. The minimiser that is the quantile process alias Kamae–Krengel process [15] is also a minimiser of the Benamou–Brenier problem.

2 From curtain couplings to limit curtain processes

2.1 Setting

Recall that we only consider right-continuous peacocks indexed on a compact interval, that most of the time will be $[0, 1]$. We remind the reader that $(\mu_t)_{t \in [0,1]}$ is a peacock if for every $s, t \in [0, 1]$, μ_t is in \mathcal{P}_1 and $\mu_s \preceq_C \mu_t$. We shall sometimes abbreviate $(\mu_t)_{t \in [0,1]}$ to $[\mu]$. Note that the set of discontinuity $D_{[\mu]} = \{t \in [0, 1], \mu_{t-} \neq \mu_t\}$ is countable. Indeed if $\mu_{t-} \neq \mu_t$, there exists a martingale coupling $\pi \in \Pi_M(\mu_{t-}, \mu_t)$ with $\pi(\{(x, y) \in \mathbb{R}^2 : x = y\}) < 1$. Therefore, if φ is the strictly convex function $x \mapsto \sqrt{1 + x^2}$ we are not in the equality case of the conditional version of Jensen’s inequality and thus $\int \varphi d\mu_{t-} < \int \varphi d\mu_t$. Hence $D_{[\mu]}$ is the set of discontinuity points of $t \mapsto \int \varphi(x) d\mu_t(x)$ and it is countable.

2.2 Limit curtain processes

Given any peacock $[\mu]$, we describe a procedure to construct a process using a discretisation. More precisely, for any interval partition $\sigma = \{t_0, \dots, t_N\}$ of $[0, 1]$ with $0 = t_0 < \dots < t_N = 1$, we denote by $P^{\mu, \sigma}$ the law, on the Skorokhod space $D[0, 1]$, of a process $(X_t^{\mu, \sigma})_{t \in [0,1]}$ that we describe now: A random trajectory is (almost surely)

constant on every $[t_k, t_{k+1}[$ (or $[t_k, t_{k+1}]$ if $k = N - 1$), and for every $k < N$ the joint law $(X_{t_k}^{\mu, \sigma}, X_{t_{k+1}}^{\mu, \sigma})$ is the left-curtain coupling between μ_{t_k} and $\mu_{t_{k+1}}$. In order, for $P^{\mu, \sigma}$, to be uniquely determined we moreover need to specify that $X^{\mu, \sigma}$ is a Markov process. Notice that, as the process is most of time constant, it can be seen as an inhomogeneous Markov chain with transitions at deterministic times $(t_k)_{k=0, \dots, N}$. Note moreover that $(X_t^{\mu, \sigma})_{t \in [0, 1]}$ is a martingale in the canonical filtration $(\mathcal{F}_t)_{t \in [0, 1]}$, where \mathcal{F}_s is generated by the random variables $(X_u^{\mu, \sigma})_{u \leq s}$. Indeed, for $s \leq t$, $\mathbb{E}(X_t^{\mu, \sigma} | \mathcal{F}_s^{\mu, \sigma}) = X_s^{\mu, \sigma}$ because the transitions are martingale transport plans and $\mathbb{E}(X_t^{\mu, \sigma} | \mathcal{F}_s) = \mathbb{E}(X_t^{\mu, \sigma} | \mathcal{F}_s^{\mu, \sigma})$ because the process is Markovian.

We are ready for the definition of limit curtain coupling.

Definition 2.1. Let $[\mu] = (\mu_t)_{t \in [0, 1]}$ be a right-continuous peacock. The probability measure P on $D[0, 1]$ is called a limit curtain process of $[\mu]$ if there exists a sequence of partitions $(\sigma^{(p)})_{p \in \mathbb{N}}$ with mesh $|\sigma^{(p)}|$ going to 0 as p tends to $+\infty$, such that $P^{\mu, \sigma^{(p)}}$ tends to P in the weak convergence sense.

Moreover we denote by $\text{LimCurt}([\mu])$ the set of limit curtain couplings of $[\mu]$.

The definitions readily generalise to right-continuous peacocks defined on other segments than $[0, 1]$.

Actually any element $P \in \text{LimCurt}([\mu])$ is relevant for the peacock problem because any process $(X_t)_{t \in [0, 1]}$ of law P — (i) satisfies $\text{Law}(X_t) = \mu_t$ for every t , (ii) is a martingale (see Proposition 2.6 for the proof of these two points). One of our goals is to provide sufficient conditions on $[\mu]$ to ensure that $\text{LimCurt}([\mu])$ is not empty, or, even better, reduced to a single element. The other goal is to examine whether there are Markovian or non-Markovian processes among the limit processes.

2.3 Transformations

Given the peacock $[\mu]$, a sequence $(\sigma^{(p)})_p$, and ϕ the transformation $\phi : x \mapsto ax + b$ (with $a > 0$), if $(P^{\mu, \sigma^{(p)}})_p$ converges as p goes to infinity, the same holds for the peacock $[\phi_{\#}\mu]$ and $(P^{\phi_{\#}\mu, \sigma^{(p)}})_p$. The corresponding limit processes can be jointly written $(X_t)_{t \in [0, 1]}$ and $(aX_t + b)_{t \in [0, 1]}$. That this relation holds, is due to the fact that affine maps like ϕ transform monotone couplings – recall Definition 1.9 – into monotone couplings.

Let us now consider another natural transformation $t \mapsto \mu_{\tau(t)}$ where τ is a continuous increasing function from $[0, 1]$ to $[0, \tau(1)]$. Using the uniform continuity of τ we see that $\text{LimCurt}([\mu_{\tau}])$ is exactly the set of laws of processes $(X_{\tau(t)})_{t \in [0, 1]}$ where $(X_t)_{t \in [0, \tau(1)]}$ is any process of law $P \in \text{LimCurt}([\mu_t])$.

2.4 Finite dimensional topology

In a similar way as we did for $\text{LimCurt}([\mu])$ we introduce $\text{LimCurtFD}([\mu])$.

Definition 2.2. Let $[\mu] = (\mu_t)_{t \in [0, 1]}$ be a right-continuous peacock. The probability measure P on $\mathbb{R}^{[0, 1]}$, with its cylindrical σ -field, is called a limit curtain process for the finite dimensional topology of $[\mu]$ if there exists a sequence of partitions $(\sigma^{(p)})_{p \in \mathbb{N}}$ with mesh $|\sigma^{(p)}|$ going to 0 as p tends to $+\infty$, such that $P^{\mu, \sigma^{(p)}}$ tends to P in the weak convergence sense.

Moreover we denote by $\text{LimCurtFD}([\mu])$ the set of limit curtain couplings of $[\mu]$.

We continue to denote by $(X_t)_{t \in [0, 1]}$ a process of law P as in Subsection 2.2. Note that the probability measure is now a measure on $\mathbb{R}^{[0, 1]}$, instead of $D[0, 1]$. We have $P \in \text{LimCurtFD}([\mu])$ if and only if there exist a sequence of processes $((X_t^{\mu, \sigma^{(p)}})_{t \in [0, 1]})_p$ of law $(P^{\mu, \sigma^{(p)}})_p$ such that for all $k \in \mathbb{N}$ and $u_1 < \dots < u_k$, $\text{Law}(X_{u_1}^{\mu, \sigma^{(p)}}, \dots, X_{u_k}^{\mu, \sigma^{(p)}})$ weakly converges to $\text{Law}(X_{u_1}, \dots, X_{u_k})$ as p tends to infinity.

The following example shows that a limit curtain process for the finite dimensional topology may not satisfy the condition $\text{Law}(X_t) = \mu_t$. It may also be a measure on $\mathbb{R}^{[0,1]}$ that is not obtained from a measure on $D[0, 1]$.

Example 2.3. We consider $\sigma^{(p)} = \{k/2^p, k = 0, \dots, 2^p\}$ for every $p \in \mathbb{N}$ and

$$\mu_t = \begin{cases} \delta_0 & \text{if } t < 1/3, \\ (\delta_{-1} + \delta_1)/2 & \text{if } t \geq 1/3. \end{cases}$$

The sequence $(P^{\mu, \sigma^{(p)}})_{p \in \mathbb{N}}$ converges to some $P \in \text{LimCurtFD}([\mu])$. If $(X_t)_{t \in [0,1]}$ has law P it satisfies $\text{Law}(X_{1/3}) = \delta_0 \neq \mu_{1/3}$. Moreover, it is impossible to modify $(X_t)_{t \in [0,1]}$ to make it a càdlàg process.

Nevertheless, an element of $\text{LimCurtFD}([\mu])$ is always a martingale as can be seen by adapting the proof of Lemma 3.4 in [12] by Hirsch and Roynette. In the next proposition, we adapt another proof by the same authors [12, Theorem 3.2] that relies on compactness, Cantor’s diagonal argument, and the classical regularisation of martingales. Note that, in order to prevent from non relevant limits as the one in Example 2.3, the sequence of partitions chosen in Proposition 2.4 includes the times of discontinuity of the peacock. As we will see in Proposition 2.6, such a behaviour can not occur for the Skorokhod topology.

Proposition 2.4. *Let $[\mu] = (\mu_t)_{t \in [0,1]}$ be a right-continuous peacock. The set $\text{LimCurtFD}([\mu])$ is not empty and contains at least one relevant process.*

More precisely, for any nested sequence of partitions $(\sigma^{(p)})_{p \in \mathbb{N}}$ such that $\cup_p \sigma^{(p)}$ is dense and contains the points of discontinuity of $t \mapsto \mu_t$, there exists a subsequence of $(P^{\mu, \sigma^{(p)}})_{p \in \mathbb{N}}$ converging to a limit $P \in \text{LimCurtFD}([\mu])$. Moreover, there exists a process $(X_t)_{t \in [0,1]}$ associated with P that is a (maybe non-Markovian) càdlàg martingale with 1-dimensional marginals μ_t at any time $t \in [0, 1]$.

Sketch of proof of Proposition 2.4. We mostly follow the proof of [12, Theorem 3.2] and will cite the numbering of this article. The only important difference concerns the partition sequence $(\sigma^{(p)})_p$ that may not be dyadic and must include the times of discontinuity. This is an essential feature with regard to the finite dimensional convergence that is proved after Lemma 2.5, but it is not for Theorem 3.2 in [12] that is a pure existence theorem. At step (2) we adapt the proof as follows: We take $X_t^{\mu, p} = X_{[t]^p}^{\mu, p}$ instead of 0 where $[t]^p$ is the greatest element of $\sigma^{(p)}$ that is smaller than t . Moreover, we choose left-curtain couplings as Markov transitions instead of an arbitrary martingale provided by Strassen’s Theorem. Observe that after step (2) we are really considering $P^{\mu, p} := P^{\mu, \sigma^{(p)}}$ and an associated process $X^{\mu, p}$, as described in Subsection 2.2. We can follow the steps of Hirsch and Roynette until the end. We obtain a measure P on $D[0, 1]$ as the limit of a subsequence $(P^{\mu, \varphi^{(p)}})_p$ of $(P^{\mu, p})_p$ in a peculiar sense: When finitely many times of $\cup_p \sigma^{(p)}$ are selected the joint projection of $P^{\mu, \varphi^{(p)}}$ on these marginals converge to the corresponding projection for P . Hence, there exists a càdlàg martingale $(X_t)_{t \in [0,1]}$ that has the correct marginal law of P when projected on countably many copies of \mathbb{R} with indices in $\cup_p \sigma^{(p)}$. The fact that such a process exists is granted by the classical theory of continuous martingales (see step (5)). Moreover, using the continuity at times $t \notin \cup_p \sigma^{(p)}$ the 1-marginals are μ_t on this set also.

The only element that must be added to Hirsch and Roynette’s proof is the finite dimensional convergence for finitely many times not all in $\cup_p \sigma^{(p)}$. For this purpose we need a lemma.

Lemma 2.5. *Let $(\nu_t)_{t \in [0,1]}$ be a peacock, continuous at time r . Then for every $\varepsilon > 0$, there exists α such if $\max(|s - r|, |t - r|) < \alpha$, the sets $\Pi_M(\nu_s, \nu_t)$ and $\Pi_M(\nu_t, \nu_s)$ are*

contained in the ball of centre $(\text{Id} \otimes \text{Id})_{\#} \nu_r$ and radius ε (for any distance metrising $\mathcal{D} = \{\pi \in \mathcal{P}(\mathbb{R}^2) : (\text{proj}_i)_{\#} \pi \preceq_C \nu_1, i = 1, 2\}$).

Proof. The result is due to the fact that $\Pi_M(\nu_r, \nu_r) = \{(\text{Id} \otimes \text{Id})_{\#} \nu_r\}$. Let $(\pi_n)_n$ be a sequence of martingale transport plans with both marginals converging to ν_r . The set \mathcal{D} is tight because for every $\pi \in \mathcal{D}$

$$\iint_{|x|+|y| \geq n} 1 \, d\pi(x, y) \leq \iint \frac{|x| + |y|}{n} \, d\pi(x, y) \leq \frac{1}{n} \left(\int |x| \, d\nu_1(x) + \int |y| \, d\nu_1(y) \right).$$

Thus Prokhorov theorem shows that every subsequence of $(\pi_n)_n$ has a converging subsequence in the closed set Π_M . The limits must be in $\Pi_M(\nu_r, \nu_r)$, so that $(\pi_n)_n$ converges to $(\text{Id} \otimes \text{Id})_{\#} \nu_r$. The lemma follows from this remark. \square

Take r_1, \dots, r_j real times and for every $i \in \mathbb{N}$, q_1^i, \dots, q_j^i elements of $\cup_p \sigma^{(p)}$ such that for every $k \leq j$, $(q_k^i)_i$ is a decreasing sequence converging to r_k . We specify moreover that q_k^i has to be r_k if $r_k \in \cup_p \sigma^{(p)}$. Moreover, times $r_k \notin \cup_p \sigma^{(p)}$ are continuity points of the peacock so that in both cases we have $\lim_{i \rightarrow \infty} \mu_{q_k^i} = \mu_{r_k}$. The sequence $u_{i,p} = \text{Law}(X_{q_1^i}^{\mu, \varphi^{(p)}}, \dots, X_{q_j^i}^{\mu, \varphi^{(p)}})$ depends on i and on p . If i tends to infinity for a fixed p , because paths of $D[0, 1]$ are right-continuous, the sequence $u_{i,p}$ converges to $\text{Law}(X_{r_1}^{\mu, \varphi^{(p)}}, \dots, X_{r_j}^{\mu, \varphi^{(p)}})$. If p tends to infinity, for a fixed i , it converges to $\text{Law}(X_{q_1^i}, \dots, X_{q_j^i})$ where $(X_t)_{t \in [0, 1]}$ has law P . Moreover, $\text{Law}(X_{q_1^i}, \dots, X_{q_j^i})$ tends to $\text{Law}(X_{r_1}, \dots, X_{r_j})$ when i goes to infinity. We need to state the commutativity of limits $\lim_p(\lim_i u_{i,p}) = \lim_i(\lim_p u_{i,p})$. Lemma 2.5 will provide the required uniformity. For this lemma we choose the Prokhorov distance on the space of probability measures on $(\mathbb{R}^j, \|\cdot\|_{\infty})$ and denote it by Prok . Take $\varepsilon > 0$. We claim that for p sufficiently large (greater than some p_{ε}) the inequality

$$\text{Prok}(\text{Law}(X_{r_1}^{\mu, \varphi^{(p)}}, \dots, X_{r_j}^{\mu, \varphi^{(p)}}), \text{Law}(X_{q_1^i}^{\mu, \varphi^{(p)}}, \dots, X_{q_j^i}^{\mu, \varphi^{(p)}})) < 4\varepsilon$$

holds as soon as i is large enough (greater than some i that does not depend on p). Indeed, for every $k = 1, \dots, j$, $\text{Law}(X_{r_k}^{\mu, \varphi^{(p)}}, X_{q_k^i}^{\mu, \varphi^{(p)}})$ is an element of $\Pi_M(\mu_s, \mu_t)$ where s and t have distance to r_k smaller than $\max(|\sigma^{\varphi^{(p)}}|, |q_k^i - r_k|)$. With the same notation α as in Lemma 2.5 we choose $p_{\varepsilon, k}$ sufficiently large so that $|\sigma^{\varphi^{(p)}}| \leq \alpha$ if $p > p_{\varepsilon, k}$. If now we also have $|q_k^i - r_k| \leq \alpha$, the Prokhorov distance, associated with the norm $\|\cdot\|_{\infty}$ of \mathbb{R}^2 , between $\text{Law}(X_{r_k}^{\mu, \varphi^{(p)}}, X_{q_k^i}^{\mu, \varphi^{(p)}})$ and $(\text{Id} \otimes \text{Id})_{\#} \mu_{r_k}$ is smaller than ε . Hence considering a coupling as ensured by the Strassen-Dudley theorem one can couple these measures in a close way. Therefore if (Ω, \mathbb{P}) is the probability space on which $X^{\mu, \varphi^{(p)}}$ is defined, $\mathbb{P}(|X_{r_k}^{\mu, \varphi^{(p)}} - X_{q_k^i}^{\mu, \varphi^{(p)}}| < 2\varepsilon) \geq 1 - \varepsilon$, which means that for $p > p_{\varepsilon} := \max_k p_{\varepsilon, k}$ the event $\{\max_{1 \leq k \leq j} |X_{r_k}^{\mu, \varphi^{(p)}} - X_{q_k^i}^{\mu, \varphi^{(p)}}| < 2\varepsilon\}$ holds with probability greater than $1 - j\varepsilon$. For such a value of p , as i is sufficiently large, it holds

$$\begin{aligned} & \text{Prok}(\text{Law}(X_{r_1}^{\mu, \varphi^{(p)}}, \dots, X_{r_j}^{\mu, \varphi^{(p)}}), \text{Law}(X_{r_1}, \dots, X_{r_j})) \\ & \leq \text{Prok}(\text{Law}(X_{r_1}^{\mu, \varphi^{(p)}}, \dots, X_{r_j}^{\mu, \varphi^{(p)}}), \text{Law}(X_{q_1^i}^{\mu, \varphi^{(p)}}, \dots, X_{q_j^i}^{\mu, \varphi^{(p)}})) \\ & \quad + \text{Prok}(\text{Law}(X_{q_1^i}^{\mu, \varphi^{(p)}}, \dots, X_{q_j^i}^{\mu, \varphi^{(p)}}), \text{Law}(X_{q_1^i}, \dots, X_{q_j^i})) \\ & \quad + \text{Prok}(\text{Law}(X_{q_1^i}, \dots, X_{q_j^i}), \text{Law}(X_{r_1}, \dots, X_{r_j})). \\ & \leq j\varepsilon + \varepsilon + \varepsilon \end{aligned}$$

This proves the finite dimensional convergence. \square

With similar arguments as those in Proposition 2.4 we prove now that all elements of $\text{LimCurt}([\mu])$ are relevant.

Proposition 2.6. *Let $[\mu]$ be a peacock and $(\sigma^{(p)})_{p \in \mathbb{N}}$ a sequence of partitions such that $(P^{\mu, \sigma^{(p)}})_{p \in \mathbb{N}}$ converge to some P in the Skorokhod topology. Then any process $(X_t)_{t \in [0, 1]}$ of law P satisfies $\text{Law}(X_t) = \mu_t$ for every $t \in [0, 1]$ and is a martingale.*

Proof. With the convergence in the Skorokhod space, the finite dimensional convergence is also true for finitely many times selected in a set $E = [0, 1] \setminus D$ where D is countable (and $1 \in E$). Without loss of generality, we can assume that the paths of $(X_t)_{t \in [0, 1]}$ are right-continuous and $(\mu_t)_{t \in [0, 1]}$ is right-continuous. It follows that $\text{Law}(X_t) = \mu_t$ is also satisfied for $t \in D$. Moreover $(X_t)_{t \in E}$ is a martingale on E (see for instance [12, Lemma 3.4]). Finally, $(X_t)_{t \in [0, 1]}$ is the regularised martingale obtained at any point as limit on the right in the sense of the classical theory of continuous martingales (same argument as step (5) in [12, proof of Theorem 3.2]). \square

Finally, notice that Example 2.3 illustrates the fact that a sequence $P^{\mu, p}$ may converge to two noncompatible limits for the Skorokhod and the finite dimensional topology.

In the following sections we examine $\text{LimCurt}([\mu])$ for some particular classes of peacocks $[\mu]$.

3 Uniform measures on intervals

In this section we consider a prototypical peacock $[\mu]$: for every $t \in [0, 1]$, the law of μ_t is a simple continuous measure, namely the uniform measure on an interval.

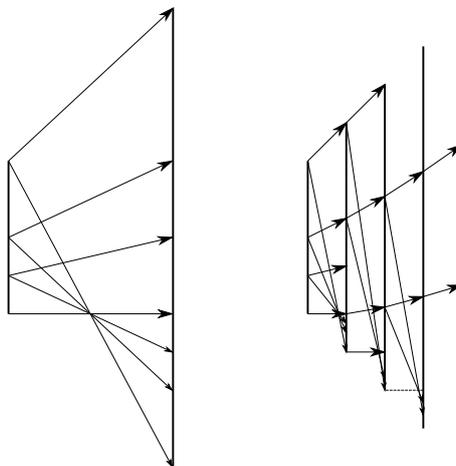


Figure 1: Composition of curtain couplings for uniform measures.

The left-curtain coupling π_{1c} between two uniform measures, μ and ν , can easily be deduced from Definition 1.8. As explained in Subsection 2.3, the resulting coupling is invariant under translation and scaling, so that it is enough to look at $\mu = \lambda_{[0, 1]}$ and $\nu = \frac{1}{1+2a} \lambda_{[-a, 1+a]}$ for some $a > 0$. Note now that, for every $x \in [0, 1]$, the shadow of $\mu|_{]-\infty, x]}$ in ν has the same mass and barycentre, respectively x and $x/2$, as $\mu|_{]-\infty, x]}$. It is readily proved that this shadow is the uniform measure $\frac{1}{1+2a} \lambda_{[-ax, (1+a)x]}$. Therefore, the left-curtain coupling from μ to ν can be described with two linear maps. The submeasure $\frac{1+a}{1+2a} \lambda_{[0, 1]} \leq \mu$ is mapped linearly on $[0, 1+a]$ and the remainder $\frac{a}{1+2a} \lambda_{[0, 1]}$ linearly on $[0, -a]$. The coupling can also be described with random variables. Let X be uniform on $[0, 1]$ and Z be an independent Bernoulli variable $Z \sim \mathcal{B}(a/(1+2a))$. Then define

$Y = (1 + a)X$ if $Z = 0$ and $Y = -aX$ if $Z = 1$. This gives $\pi_{1c} = \text{Law}(X, Y)$. In Figure 1, on the left we illustrate the curtain coupling π_{1c} between two uniform measures. On the picture, an arrow from x to y corresponds to some ordered pair (x, y) in the support of π . Roughly speaking the transport route is open for transporting mass from x to y . On the right of the figure we consider five uniform measures in convex order and draw the same arrows representing the four curtain couplings between them.

The next result is the main part of Theorem B. The remaining part is based on Subsection 2.3 and proved at the end of the present section.

Proposition 3.1. *For every $t \in [0, 1]$, let μ_t be uniform on $[-\exp(2t)/2, \exp(2t)/2]$. Then there is a unique limit curtain process to the peacock $[\mu] = (\mu_t)_{t \in [0,1]}$ and it can be described as follows: Choose an initial point X_0 uniformly on $[-1/2, 1/2]$ and independently a one-dimensional Poisson point process of intensity 1 with jump times $0 = T_0 < T_1 < \dots < T_N < 1$ where $N \sim \mathcal{P}(1)$. The random path is defined as follows.*

$$X(t) = \begin{cases} \exp(2t)/2 - (1/2 - X_0) \exp(t) & \text{if } t \in [0, T_1[, \\ \exp(2t)/2 - \exp(t + T_i) & \text{if } t \in [T_i, T_{i+1}[, \\ \exp(2t)/2 - \exp(t + T_N) & \text{if } t \in [T_N, 1]. \end{cases}$$

In simple words, at successive random times T , the process jumps down and starts new pieces of increasing curves from position $-\exp(2T)/2$.

Let us briefly give an idea on how the Poisson point process can appear. The curtain coupling of μ_t and μ_{t+h} is directed by the quotient of their support lengths. This is $\exp(2h)$, which is independent of t . Therefore, the probability to jump down is $2^{-1}(1 - \exp(-2h))$, which is equivalent to h as h tends to 0. Recall that it is also independent of the position in the support of μ_t . In the limit, the probability to jump down during the period $[t, t + h]$ is still independent of the position and from t . The transition kernel between μ_t and μ_{t+h} is, up to a scaling factor, the same for $h > 0$ fixed, and every $t \in [0, 1 - h]$.

Our proof of Proposition 3.1 relies on the Euler approximation method and the approximation of the classical Poisson point process by Bernoulli processes.

3.1 Euler approximation method

We consider a continuous transition function $T : (s, t, x) \in [0, 1] \times [0, 1] \times \mathbb{R} \mapsto \mathbb{R}$ defined for $s \leq t$ such that the limit

$$V(t, x) = \lim_{h \rightarrow 0^+} T(t, t + h, x)/h$$

exists. We denote by R_T the rest of the Taylor expansion

$$R_T(s, t, x) = T(s, t, x) - (t - s)V(s, x).$$

In the next proposition we will compare, given any partition $\sigma : 0 = t_0 < t_1 < \dots < t_N < 1 = t_{N+1}$ and two initial points x_0 and \bar{x}_0 , the solution $x(t)$ of the ODE

$$\begin{cases} x(0) = x(t_0) = x_0 \\ \dot{x}(t) = V(t, x(t)) \end{cases} \tag{3.1}$$

to the Euler scheme starting in \bar{x}_0 :

$$\begin{cases} \bar{x}_0 = x_0 \\ \bar{x}_{k+1} = \bar{x}_k + T(t_k, t_{k+1}, \bar{x}_k). \end{cases}$$

The comparison between \bar{x}_k and $x(t_k)$ can be done at discrete times t_k but also at continuous times by associating with $(\bar{x}_k)_{k=0,\dots,T}$ the càdlàg function \bar{x} defined by $\bar{x}(t) = \bar{x}_k$ on $[t_k, t_{k+1}[$. The proof follows the classical line for the convergence of the Euler scheme in numerical analysis.

Proposition 3.2. *Let T, V and R_T be the functions introduced above and assume that V is continuous, bounded, and that there exists $L > 0$ such that*

$$|V(y, t) - V(x, t)| \leq L|y - x|.$$

Let R_V be the local truncature error in the approximation of the flow at first order. We assume the uniform estimates $|R_T(t, t + h, x)| := |V(x, t + h) - V(x, t)| \leq Mh^2/2$ and $|R_V(t, t + h, x)| \leq Mh^2/2$ for some $M > 0$.

Then there is a function $F : (\mathbb{R}^+)^2 \rightarrow \mathbb{R}^+$ increasing in both arguments and with $F(0, 0) = 0$ such that

$$\|x - \bar{x}\|_\infty \leq F(|x_0 - \bar{x}_0|, |\sigma|).$$

Note that the hypotheses on V ensure that (3.1) is in the scope of Picard-Lindelöf Theorem.

Proof. We consider the one-step operation starting from \bar{x}_k and x_k on the interval $[t_k, t_{k+1}[$:

$$\bar{x}_{k+1} = \bar{x}_k + (t_{k+1} - t_k)V(t_k, \bar{x}_k) + R_T(t_k, t_{k+1}, \bar{x}_k)$$

and

$$x(t_{k+1}) = x(t_k) + \int_{t_k}^{t_{k+1}} V(s, x(s))ds = x_k + (t_{k+1} - t_k)V(t_k, x_k) + R_V(t_k, t_{k+1}, x_k);$$

We take the difference and obtain

$$\begin{aligned} |\bar{x}_{k+1} - x(t_{k+1})| &\leq |\bar{x}_k - x(t_k)| + (t_{k+1} - t_k)|V(t_k, \bar{x}_k) - V(t_k, x_k)| + |R_V| + |R_T| \\ &\leq |\bar{x}_k - x(t_k)|(1 + L(t_{k+1} - t_k)) + M(t_{k+1} - t_k)^2. \end{aligned}$$

Using the fact that $\prod_{k=1}^n (1 + h_k) \leq \exp(h_1 + \dots + h_n)$ for positive real numbers h_k . It follows for $n \leq N$

$$\begin{aligned} |\bar{x}_n - x_n| &\leq |\bar{x}_0 - x_0| \exp(L.t_n) + \sum_{k=1}^n M(t_k - t_{k-1})^2 \exp(L(t_n - t_k)) \\ &\leq |\bar{x}_0 - x_0| \exp(L.t_n) + M|\sigma| \exp(L.t_n) \leq (|\bar{x}_0 - x_0| + M|\sigma|) \exp(L.t_n). \end{aligned}$$

Thus $\|\bar{x} - x\|_\infty \leq (|\bar{x}_0 - x_0| + M|\sigma|) \exp(L.t_n) + \|V\|_\infty \cdot |\sigma|$ □

3.2 Poisson point process

We state the following result without proof. It states that it is possible to couple a Bernoulli process and a Poisson process. We invite the reader to consult [2] on the Poisson approximation.

Lemma 3.3. *Let $(\sigma^{(p)})_{p \in \mathbb{N}}$ be a sequence of interval partitions with the mesh $|\sigma^{(p)}|$ going to 0 as p tends to $+\infty$. Then there exists a probability space on which one can define an increasing sequence of random variables $(T_i)_{i \in \mathbb{N}^*}$ and for every $p \in \mathbb{N}$, an increasing sequence $(T_i^{(p)})_{i \in \mathbb{N}^*}$, such that*

- $\{T_i : T_i < 1\}$ realises a Poisson point process of intensity 1 on $[0, 1]$, say $T_0 = 0$, the $(T_i - T_{i-1})_{i \in \mathbb{N}^*}$ are independent and have exponential distribution of parameter 1.

- For every $p > 0$, $\{T_i^{(p)} : T_i^{(p)} \leq 1\}$ is a Bernoulli point process on $\sigma^{(p)}$ with the following meaning: Every $T_i^{(p)}$ takes values in $\sigma^{(p)} \cup \{+\infty\}$ where the elements of $\sigma^{(p)}$ are $0 = t_0 < \dots < t_{Q_p} = 1$. The random times $T_1^{(p)} < T_2^{(p)} < \dots$ are ordered and the events $(E_k)_{k \geq 1}$ defined by $E_k = \{t_k \in \{T_i^{(p)} : i \geq 1\}\}$ are independent and have probability $2^{-1}(1 - \exp(2(t_{k-1} - t_k)))$.
- For every $\varepsilon > 0$, if $N = \#\{T_i : T_i < 1\}$ and $N^{(p)} = \#\{T_i^{(p)} : T_i^{(p)} \leq 1\}$ we have

$$\mathbb{P} \left(\{N = N^{(p)}\} \text{ and } \{i \leq N \Rightarrow |T_i^{(p)} - T_i| \leq \varepsilon\} \right) \xrightarrow{n \rightarrow \infty} 1.$$

The next lemma concerns the trajectories of the limit process suggested in Proposition 3.1. These are for $s \in [0, 1]$,

$$g_s : t \in [s, 1] \mapsto \frac{1}{2} \exp(2t) - \exp(s + t) \in \mathbb{R}.$$

Lemma 3.4. Let s, t, s', t' be elements of $[0, 1]^2$, with $s < t$ and $s' < t'$. Then it holds

$$\max_{u \in [0, 1]} |g_s(s + u(t - s)) - g_{s'}(s' + u(t' - s'))| \leq 10(|s' - s| + |t' - t|).$$

Proof. For every $u \in [0, 1]$ we consider the partial derivatives of $(s, t) \mapsto g_s(s + u(t - s))$. The norm of these derivatives is bounded by e^2 so that the mean inequality permits us to conclude. \square

Proof of Proposition 3.1. We prove that for a sequence $(\sigma^{(p)})_p$, the sequence $P^{\mu, p} := P^{\mu, \sigma^{(p)}}$ converges in the Skorokhod topology to the law P of the process described in the statement of the proposition. Our strategy is to use the Prokhorov distance associated with the Skorokhod distance. In different words, for every $\varepsilon > 0$, we want to couple P and $P^{\mu, p}$ in $D[0, 1] \times D[0, 1]$ using a coupling Θ such that with probability greater than $1 - \varepsilon$ the (Skorokhod) distance between the two marginal associated càdlàg processes $(X_t)_{t \in [0, 1]}$ and $(X_t^{\mu, p})_{t \in [0, 1]}$ is smaller than ε .

It is also correct to perform the coupling in another probability space and this is what we will do with the probability space of Lemma 3.3 together with a uniform random value $X_0 \sim \mathcal{U}([-1/2, 1/2])$ independent of this space. We construct the process X as explained in the statement of the proposition. A random path starts from X_0 at time 0 and jumps down at times T_1, \dots, T_N . The piece of trajectory after the i -th jump is defined by g_{T_i} .

Before we describe the piecewise constant process $X^{\mu, p}$, let us introduce an intermediate process Y^p . A random path starts at point X_0 and jumps at each time $t_k^{(p)}$ if and only if the interval $[t_{k-1}^{(p)}, t_k^{(p)}[$ contains some T_i . After a jump the trajectory is g_{t_k} .

The process $X^{\mu, p}$ does not directly follow the trajectories g_s but it is a discretisation of those trajectories in the sense of Proposition 3.2. A random path starts from X_0 . It is constant on each interval $[t_{k-1}^{(p)}, t_k^{(p)}[$. At time $t_k^{(p)}$ it jumps down into $[g_{t_{k-1}^{(p)}}(t_{k-1}^{(p)}), g_{t_k^{(p)}}(t_k^{(p)})]$ with probability $2^{-1}(1 - \exp(2(t_{k-1}^{(p)} - t_k^{(p)})))$, which is small. In the other case it does a small jump upwards from $x = X_{t_{k-1}^{(p)}}^{\mu, p}$ to

$$T(t_{k-1}^{(p)}, t_k^{(p)}, x) = (x + 2^{-1} \exp(2t_{k-1}^{(p)})) \frac{1 + \exp(2\delta t_k^{(p)})}{2} - 2^{-1} \exp(2t_{k-1}^{(p)})$$

where $\delta t_k^{(p)} = t_k^{(p)} - t_{k-1}^{(p)}$. The vector field V corresponding to this transition T with respect to the definitions of Subsection 3.1 is $V(t, x) = x + 2^{-1} \exp(2t)$. Note that V is 1-Lipschitz in x and continuous in t . The solutions of the ODE (3.1) are of the form

$\exp(2t)/2 - C \exp(t)$ where C is a constant. For $C = \exp(s)$ we recover g_s . The trajectories of the flow starting from $[-1/2, 1/2]$ at time 0 or from $-2^{-1} \exp(2s)$ for some $s \in [0, 1]$ at time s are bounded and V is also bounded if (t, x) is in a bounded set.

We can now conclude, explaining that with high probability the trajectories of X are close to those of Y^p and that the trajectories of Y^p are close to those of $X^{\mu,p}$. Of course, this holds if p is large enough. For the first estimate we consider the event $\{N = N^{(p)}\} \cap \{i \leq N \Rightarrow |T_i^{(p)} - T_i| \leq \varepsilon\}$ and define $\lambda : [0, 1] \mapsto [0, 1]$ as the piecewise linear and continuous change of time that fixes $\{0, 1\}$ and maps each T_i for $i \leq N$ on $T_i^{(p)}$. With this λ used in the definition of the Skorokhod distance and Lemma 3.4, we see that the Prokhorov distance between $\text{Law}(X)$ and $\text{Law}(Y^p)$ is smaller than 10ε . As ε can be chosen arbitrary, this first distance tends to 0 as p tends to infinity.

The distance between $\text{Law}(Y^p)$ and $\text{Law}(X^{\mu,p})$ also tends to zero: Proposition 3.2 allows us to compare on each $[t_{k-1}^{(p)}, t_k^{(p)}]$, without time wiggling, the piecewise constant trajectories of $X^{\mu,p}$ to the continuous pieces $g_{T_i^{(p)}}$ of Y^p . The beginning of the first trajectories lies in $[g_{t_{k-1}^{(p)}}(t_{k-1}), g_{t_k^{(p)}}(t_k^{(p)})]$ and beginning of the other ones at $g_{t_k^{(p)}}(t_k^{(p)})$, so that the distance between these points tends to zero together with $|\sigma^{(p)}|$. Moreover, the precise expression of F given in the proof of Proposition 3.2 allows us to bound uniformly from above the supremum norms over all pieces of Y^p . \square

We conclude the section with the proof of Theorem B.

Proof of Theorem B. In Proposition 3.1 we have proved the result in the case $\mu_t = \nu_t$ where ν_t is the uniform measure on $[-g(t), g(t)]$ with $g : t \in [0, 1] \rightarrow \frac{1}{2} \exp(2t)$. Of course, the result also holds if we extend (or restrict) the same peacock to $[0, T]$ for some $T > 0$. We explain now how Subsection 2.3 can be used to deduce the general case. Let f be some positive increasing and continuous function and for every $t \in [0, 1]$ let μ_t be the uniform measure on $[-f(t), f(t)]$. If f is constant the result is trivial. Therefore we assume $f(1) > f(0)$. Let $\phi : x \mapsto ax$ be the linear function with $ag(0) = f(0)$ and $T > 0$ such that $ag(T) = f(1)$. Therefore $\phi_{\#}\nu_{tT} = \mu_t$ for $t \in [0, 1]$. Let $\tau : [0, 1] \rightarrow [0, T]$ be the function such that $\phi_{\#}\nu_{\tau(t)} = \mu_t$. In fact $\tau(t) = g^{-1}(f(t)/a) = \frac{1}{2} \log(2a^{-1}f(t))$. The function τ is increasing and continuous. Hence according to Subsection 2.3 the process $(X_t)_{t \in [0,1]}$ is a limit curtain process whose law is in $\text{LimCurt}([\mu_t])$ if and only if $(aX_{\tau(t)})_{t \in [0,1]}$ has law P for some $P \in \text{LimCurt}([\nu_t])$. Hence the law of the limit curtain process is uniquely determined. The trajectories of the process are piecewise deterministic and obtained by transforming the trajectories in Proposition 3.1. The expectation of the number of jumps on $[t_0, t_1]$ is

$$\begin{aligned} \tau(t_1) - \tau(t_0) &= \frac{1}{2} \log(2a^{-1}f(t_1)) - \frac{1}{2} \log(2a^{-1}f(t_0)) \\ &= \frac{1}{2} \log\left(\frac{f(t_1)}{f(t_0)}\right) = \frac{1}{2} \log(f(t_1)) - \frac{1}{2} \log(f(t_0)). \end{aligned} \quad \square$$

4 Finitely supported measures

Let \mathcal{V} be the set of vectors

$$(X; A) = (x_1, \dots, x_n, a_1, \dots, a_n) \in \mathbb{R}^{2n}$$

such that $\sum a_i x_i = 0$ and $\sum a_i = 1$. Every vector of \mathcal{V} can be associated with a signed measure $\sum_{i=1}^n a_i \delta_{x_i}$. Let $(x_1, \dots, x_n, a_1, \dots, a_n)$ and $(Y; B) = (y_1, \dots, y_n, b_1, \dots, b_n)$ be two elements of \mathcal{V} . As a function of $(X; A)$ and $(Y; B)$ let us now denote by $\Gamma \subseteq M_{n \times n}(\mathbb{R})$ the

subspace of matrices satisfying

$$\begin{cases} M1 = A, \\ 1^T M = B^T, \\ MY = \text{Diag}(a)X, \end{cases} \tag{4.1}$$

where 1 stays for the vector $(1, \dots, 1)^T$, $\text{Diag}(a)$ is the diagonal matrix with entries a_1, \dots, a_n and A, B, Y and X are columns.

Lemma 4.1. *With the notation above, assume that the entries of X are all different and that the same holds for Y . Then the affine space $\Gamma \subseteq M_{n \times n}(\mathbb{R})$ has dimension $(n - 1)(n - 2)$, and, in a neighborhood of $((X; A), (Y; B))$, the map*

$$f : ((X'; A'), (Y'; B')) \in \mathcal{V}^2 \mapsto \Gamma(X', A', Y', B') \subseteq M_{n \times n}(\mathbb{R})$$

is analytic, as a map to the affine Grassmannian of affine spaces of dimension $(n - 1)(n - 2)$ included in $\mathbb{R}^{n \times n} \cong M_{n \times n}(\mathbb{R})$.

Proof. We can prove that the application that maps $M \in \Gamma$ to the submatrix consisting of the $n - 1$ upper rows and the $n - 2$ left-most columns is an affine bijection with $M_{(n-1) \times (n-2)}$. Indeed, there always exists a way to complete such a matrix to an element of Γ and this way is unique. We first consider the $n - 1$ upper rows together with the first and third constraint of (4.1). On each line we obtain a 2×2 linear system to solve and the solution is unique because of $\begin{vmatrix} 1 & x_{n-1} \\ 1 & x_n \end{vmatrix} \neq 0$. We complete the n -th row in the unique possible way according to the second constraint and we have still two relations on the lower row that need to be checked. These relations rely on the definition of \mathcal{V} . First, we already have $\sum \sum m_{ij} = \sum b_j = 1 = \sum a_i$ and $\sum_j m_{ij} = a_i$ for every $i \leq n - 1$. It follows $\sum_j m_{nj} = a_n$. Second, we have $\sum_j m_{ij} y_j = a_i x_i$ for every $i \leq n - 1$ and we want to prove it for $i = n$. This follows by subtracting these $n - 1$ relations to $\sum b_j y_j = \sum a_i x_i$. \square

In this section we are interested in defining a limit curtain coupling for peacocks $(\mu_t)_{t \in [0,1]}$ of type

$$\mu_t = \sum a_i(t) \delta_{x_i(t)}$$

where $(x_1, \dots, x_n, a_1, \dots, a_n)(t) \in \mathcal{V}$ are real analytic functions of time and furthermore satisfy $a_1(t), \dots, a_n(t) > 0$ and $x_1(t) < \dots < x_n(t)$ for every $t \in [0, 1]$. We will denote $(x_1, \dots, x_n, a_1, \dots, a_n)(t)$ by $(X_t; A_t)$ and $\Gamma(X_s, A_s, X_t, A_t)$ by Γ_{st} . The fact that the measures of $[\mu]$ are in the convex order implies that for $s \leq t$ the subspace Γ_{st} associated with $(X_s; A_s)$ and $(X_t; A_t)$ contains at least one matrix with nonnegative entries. The linear equalities defining Γ_{st} are indeed equivalent to those of $\Pi_M(\mu_s, \mu_t)$. More precisely the affine map

$$\sum_{1 \leq i, j \leq n} m_{ij} \delta_{(x_i(s), x_j(t))} \in \Pi_M(\mu_s, \mu_t) \mapsto (m_{ij})_{1 \leq i, j \leq n} \in \Gamma_{st}$$

is onto and has image $\Gamma_{st} \cap (\mathbb{R}^+)^{n \times n}$, the subset of nonnegative matrices of Γ_{st} . Hence we can identify $\Gamma_{st} \cap (\mathbb{R}^+)^{n \times n}$ with $\Pi_M(\mu_s, \mu_t)$. For all s, t we denote by $M(s, t)$ the matrix associated with the left-curtain coupling of μ_s and μ_t . It is extremely useful to write $M(s, t)$ in the form of the stochastic matrix $\tilde{M}(s, t)$ defined as follows

$$\begin{aligned} \sum_{1 \leq i, j \leq n} m_{ij}(s, t) \delta_{(x_i(s), x_j(t))} &\mapsto M(s, t) \\ &\mapsto \tilde{M}(s, t) := \text{Diag}(a)^{-1}(s)M(s, t). \end{aligned}$$

The sum of the entries on the i -th row is no longer $a_i(s)$ but 1. Moreover, $A_t^T = A_s^T \tilde{M}(s, t)$. Each row $i = 1, \dots, n$ is a state of a Markov chain and at every time the vector A_t is a probability measure on these states. However the family $(\tilde{M}(s, t))_{s,t}$ is not compatible.

According to Proposition 1.10 left-curtain couplings are exactly left-monotone couplings, which means that $M(s, t) \in \Gamma_{st}$ is the unique matrix in $(\mathbb{R}^+)^{n \times n}$ that satisfies

$$\forall i < i', \forall j^- < j' < j^+, m_{i,j^-} \times m_{i,j^+} \times m_{i',j'} = 0.$$

This is a closed condition, that is, stochastic matrices are in a bounded set. From this, we recover the main result of [14] in the specific case of finitely supported measure: The function $(s, t) \mapsto M(s, t)$ defined on $\{(s, t) \in [0, 1]^2, s \leq t\}$ is continuous.

According again to Proposition 1.10, left-curtain coupling are the unique optimal solutions of a linear minimisation problem. Therefore, $M(s, t)$ is an extreme point of $\Pi_M(\mu_s, \mu_t) \equiv \Gamma_{st} \cap (\mathbb{R}^+)^{n \times n}$. Hence $M(s, t)$ satisfies at least $(n - 1)(n - 2)$ relations of type $m_{i,j} = 0$ that are independent of the ones defining Γ_{st} .

Given an interval partition $\sigma^{(p)} = \{t_0, \dots, t_{Q_p}\}$ of $[0, 1]$ with $0 = t_0 < \dots < t_{Q_p} = 1$ of the interval $[0, 1]$, we introduce the coherent family $(\bar{R}^{(p)}(s, t))_{0 \leq s \leq t \leq 1}$ in the following way. If $s \in [t_i, t_{i+1}[$ and $t \in [t_j, t_{j+1}[$ the transition matrix between those times is $\bar{R}^{(p)}(s, t) = \tilde{M}(t_i, t_{i+1}) \cdots \tilde{M}(t_{j-1}, t_j)$. It sends the distribution of mass A_{t_i} to A_{t_j} . We will prove in Proposition 4.4 that $\bar{R}^{(p)}(s, t)$ converges to a certain R_{st} when $|\sigma^{(p)}|$ goes to zero. This will in particular prove that R_{st} is a stochastic matrix that sends the distribution A_s to A_t .

Lemma 4.2 and Lemma 4.3 apply to the analytic peacocks introduced above.

Lemma 4.2. *There exists a constant C depending on $[\mu]$ such that for all ξ^-, ξ^+ and any finite increasing sequence $(\theta_k)_{k=0}^K$ in $[\xi^-, \xi^+]$, one has*

$$\left\| \tilde{M}(\theta_0, \theta_1) \cdots \tilde{M}(\theta_{K-1}, \theta_K) - \text{Id}_n \right\| < C(\xi^+ - \xi^-)$$

Proof. Let us first prove the lemma in the case where the finite sequence θ_n is just $\theta_0 = \xi^+$ and $\theta_1 = \xi^-$. The product of transition matrices is $\tilde{M}(\xi^-, \xi^+)$ that we simply note \tilde{M} . Therefore due to the shape of the left-curtain couplings we can claim $\tilde{M}_{ij} = 0$ for $j > i + 1$ if $h := (\xi^+ - \xi^-)$ is sufficiently small. Indeed, the shadow of $\sum_{l=1}^i a_l(\xi^-) \delta_{x_l(\xi^-)}$ must be close to $\sum_{l=1}^i a_l(\xi^+) \delta_{x_l(\xi^+)}$ when h is small. More precisely considering the centres of mass of these measures, only a mass of $O(h)$ is sent to the atoms $a_{i+1}(\xi^+) \delta_{x_{i+1}(\xi^+)}, \dots, a_n(\xi^+) \delta_{x_n(\xi^+)}$ and this bound $O(h)$ can be chosen uniformly in ξ^- . Because of Lemma 1.5 if h is sufficiently small this part of the shadow can only be in x_{i+1} . Hence the claim on \tilde{M}_{ij} holds. With similar arguments and using what has already been proved it follows that the entry \tilde{M}_{ij} is also $O(h)$, uniformly in ξ^- for every $j < i$. Indeed, the measure $a_i \delta_{x_i}$ is transported to a measure of barycentre $a_i(\xi^-)$, only $O(h)$ is transported to a_{i+1} and no mass goes on upper atoms. Therefore we have proved that for any given peacock, there exists some constant $c > 0$ such that

$$\left\| \tilde{M}(\xi^-, \xi^+) - \text{Id}_n \right\| \leq c(\xi^+ - \xi^-).$$

In the general case where $(\theta_k)_k$ is not reduced to two times, using the submultiplicativity of the operator norm, the estimate $1 + x \leq \exp(x)$, and a telescopic sum we obtain

$$\begin{aligned} \left\| (\tilde{M}(\theta_0, \theta_1) \cdots \tilde{M}(\theta_{K-1}, \theta_K) - \text{Id}_n \cdots \text{Id}_n) \right\| &\leq \sum_{k=0}^{n-1} c(\theta_{k+1} - \theta_k) \exp(c(\theta_K - \theta_0)) \\ &\leq c \cdot e^c (\xi^+ - \xi^-). \end{aligned} \quad \square$$

In Theorem 4.4 the limit curtain process is described by a family of transition matrices $(R_{st})_{s \leq t}$ defined by ordinary differential equations in the space of stochastic matrices. However, in Lemma 4.3 we first study the left-curtain matrices M_{st} in place of directly the transition matrices \tilde{M}_{st} .

Lemma 4.3. *There exists a countable and closed set $E \subseteq [0, 1]$ with finitely many accumulation points, and a map $N : [0, 1] \setminus E \rightarrow N(t) = (n_{ij}(t))_{ij}$ to the square matrices of order n such that the following statements are satisfied*

- For every adjacent isolated points $\theta, \theta' \in E$, and S a segment in $] \theta, \theta' [$ there exists $C > 0$ with the uniform estimate

$$\|M(t, t+h) - \text{Diag}(a)(t) - hN(t)\| \leq Ch^2 \tag{4.2}$$

for $t \in S$ and $h > 0$,

- the map $t \mapsto N(t)$ is bounded,
- on $[0, 1] \setminus E$, the map N is analytic, that is, it is analytic on each connected component of $[0, 1] \setminus E$,
- the sum of the entries of each row of N is identically zero,
- for each $j \leq n$, the sum of the entries of the j -th column is da_j/dt ,
- if $j > i + 1$, then $n_{ij} = 0$,
- at every time $t \notin E$, at least $(n - 1)(n - 2)$ entries of $N(t)$ are zero,
- the entries on the diagonal are nonnegative and the other ones are nonpositive.

Proof. We introduce an index $k \geq 1$ such that every $k \leq \binom{n^2}{(n-1)(n-2)}$ is associated with a subset I_k of $(n - 1)(n - 2)$ entries of the matrices of $M_{n \times n}(\mathbb{R})$. Moreover, for all $s, t \in [0, 1]^2$ we only consider the subsets I_k such that the vectorial space Δ_k of matrices with the entries zero on I_k is in direct sum with the vectorial part of Γ_{st} . The spaces Γ_{st} are parallel for different values of s so that we can denote the set of these indices by $\mathcal{I}(t)$. The question whether Γ_{st} is in direct sum with Δ_k at time t is just depending on vector X_t as can be seen in (4.1). The index k will be an element of $\mathcal{I}(t)$ if and only if a certain determinant does not vanish at time t . But it is an analytic function. Hence either k is not an element of $\mathcal{I}(t)$ for every t or it is, except finitely many times on $[0, 1]$.

For $k \in \mathcal{I}(t)$ we can now introduce the analytic map $(s, t) \mapsto M_k(s, t)$ where $\{M_k(s, t)\} = \Gamma_{st} \cap \Delta_k$ is the single point at the intersection.

Let $] \xi^-, \xi^+ [$ be an interval such that $\mathcal{I}(t)$ is the same for every t . The set E of the statement is composed of the ends of these intervals, generically called ξ henceforth, and of points θ defined below, which will be finitely many on each compact subinterval $S \subseteq] \xi^-, \xi^+ [$. Due to the analyticity there are finitely many points ξ in $[0, 1]$. On $] \xi^-, \xi^+ [$, we simply denote $\mathcal{I}(t)$ by \mathcal{I} . Take $t_0 \in] \xi^-, \xi^+ [$. For every (s, t) in a neighbourhood of (t_0, t_0) , the matrix $M(s, t)$ is an extreme point of $\Gamma_{st} \cap (\mathbb{R}^+)^{n \times n}$. It equals at least one $M_k(s, t)$ for $k \in \mathcal{I}$. Recall that M is defined for $s \leq t$ whereas the matrices M_k are defined on $] \xi^-, \xi^+ [^2$. Observe also that $M(t, t) = \text{Diag}(a)(t)$. The maps $(s, t) \mapsto M_k(s, t) - M_l(s, t)$ are analytic and the locus in \mathbb{R}^2 where they vanish is accordingly well-known (see for instance [18, Chapter 6]): In the neighbourhood of a zero there are finitely many curves going out. Each curve may be a half-line or has finitely many intersections with half-lines. Hence we deduce that there exists a neighbourhood of t_0 such that for every s in this neighbourhood, there exist $k(s)$ and $\varepsilon(s) > 0$ with $M(s, s+h) = M_{k(s)}(s, s+h)$ for every $h \in [0, \varepsilon(s)]$. Moreover, the neighbourhood can be restricted, so that k is constant both

for $s < t_0$ and for $s > t_0$. Finally, the function ε can be chosen to be continuous. Using the compactness of the segments $S \subseteq]\xi^-, \xi^+[$ we see that there exists at most finitely many accident times θ_k on S . Those points θ are after the points ξ the remaining elements of E . Between two θ 's there exists $k \in \mathcal{I}$ with $M(t, t+h) = M_k(t, t+h)$ if h is small enough, and ε can be chosen uniformly on every segment included in $]\theta_k, \theta_{k+1}[$. Hence we obtain (4.2) for $N = dM_k(t, t+h)/dh|_{h=0^+}$.

The statements on N now follow from the system (4.1), equation (4.2), the definition of M_k , Lemma 4.2 and the structure of the zeros of $\tilde{M}(t, t+h)$ for small h stated in the proof of this lemma. \square

Theorem 4.4. *For every peacock $[\mu]$ concentrated on n injective and analytic curves $t \in [0, 1] \mapsto x_i(t)$ with analytic weight $a_i(t)$ (recall the setting after Lemma 4.1), there is a unique limit curtain process and this process is Markovian.*

More precisely for $\tilde{N}_t = \text{Diag}(a_1(t), \dots, a_n(t))^{-1}N_t$, the family $(R_{st})_{s \leq t}$ associated with the differential equations

$$\begin{cases} \frac{d}{du}R_{su} \big|_{u=t} = R_{st}\tilde{N}(t), \\ R_{ss} = \text{Id} \end{cases} \tag{4.3}$$

defines a set of coherent transition matrices on a space of n states. Together with the initial measure $(a_1, \dots, a_n)(0)$ on this space, it defines a Markov process with càdlàg trajectories. This process is the limit in both the finite dimensional and the Skorokhod topology of any sequence $(P^{\mu, \sigma^{(p)}})_{p \in \mathbb{N}}$ associated with a sequence $(\sigma^{(p)})_p$ of partitions with mesh going to zero.

Proof. When written for the stochastic matrices, equation (4.2) yields

$$\tilde{M}(t_0, t_0 + h) = \text{Id}_n + h\tilde{N}(t_0) + O(h^2)$$

for $h \geq 0$ going to 0 where $\tilde{N}(t) = \text{Diag}(a_1, \dots, a_n)^{-1}(t)N(t)$. Lemma 4.3 testifies that $\|\tilde{N}\|$ is uniformly bounded on $[0, 1] \setminus E$ by some constant c . It follows that the system of differential equations (4.3) is well defined, with $R_{st} = \text{Id}_n + \int_s^t R_{su}\tilde{N}_u du$ and

$$\|R_{st} - \text{Id}\| \leq c(e-1)(t-s)$$

because of the convexity of \exp and $t-s \leq 1$. As $\tilde{N}(t)$ is uniformly bounded, it is the infinitesimal generator of a non-homogeneous Markov chain with states the curves $(x_i)_{i=1, \dots, n}$. At time t , the rate for jumping from curve x_i to curve x_j is $\tilde{n}_{i,j}(t)$.

Recall that given any interval partition $\sigma^{(p)} = \{t_0, \dots, t_{Q_p}\}$ of $[0, 1]$ with $0 = t_0 < \dots < t_{Q_p} = 1$, we have introduced the coherent family of $(\bar{R}^{(p)}(s, t))_{0 \leq s \leq t \leq 1}$ before Lemma 4.2. Our first task is to prove that $\bar{R}^{(p)}(s, t)$ converges to R_{st} when $|\sigma^{(p)}|$ tends to zero. This proves in particular that R_{st} is a stochastic matrix that sends the row mass A_s to A_t . Note that, due to Lemma 4.2, we also have $\|\bar{R}_{\xi, \xi+h}^{(p)} - \text{Id}\| \leq C(h + |\sigma^{(p)}|)$ for some constant C only depending on the peacock.

In view of the notation introduced in Subsection 3.1 we can fix $s \in [0, 1]$ and denote $\bar{R}^{(p)}(s, u)$ by $x(u)$ for a few paragraphs only. We obtain $T(u, u+h, x) = x(u)\tilde{M}(u, u+h)$ and $V(x, u) = x\tilde{N}(u)$. Proposition 3.2 requires that V is continuous in u and Lipschitz continuous in x . The second condition is satisfied but the first one may not be true on every $[s, t]$. Let $E \subseteq [0, 1]$ be as in Lemma 4.3.

We first consider the case $[s, t] \subseteq]\theta_k, \theta_{k+1}[$ with $(\theta_k)_k$ as in this lemma. Hence up to a time rescaling we can apply Proposition 3.2. Thus $\bar{x}(u) = \bar{R}_p(s, u)$ uniformly converges to R_{su} for every $u \in [s, t]$ as p goes to infinity. One difficulty to overcome is that $\sigma^{(p)}$ may avoid the starting and end times s, t . This problem is fixed by the estimates of

$\|R_{\xi, \xi+h} - \text{Id}\|$ above and $\|\bar{R}_{\xi, \xi+h}^{(p)} - \text{Id}\|$ (see Lemma 4.2) when h is small. More precisely if s' and t' are respectively the greatest and smallest times in the partition $\sigma^{(p)}$ that satisfy $s' \leq s$ and $t \leq t'$, the matrices $R_{s't'}$ and $\bar{R}^{(p)}(s', t')$ tends to R_{st} and $\bar{R}^{(p)}(s, t)$ respectively.

If now $E \cap [s, t]$ is not empty, due to the structure of E , it is possible to find finitely many $[s_k, t_k]$ that do not intersect E such that the cumulated length $\sum (s_{k+1} - t_k)$ is arbitrarily small. Writing now $R_{st} = R_{s, s_1} R_{s_1 t_1} R_{t_1 s_2} \cdots R_{s_K t_K} R_{t_K t}$ and $\bar{R}^{(p)}(s, t)$ in a similar manner we obtain the estimate

$$\|R_{st} - \bar{R}^{(p)}(s, t)\| \leq C \sum_{k=1}^K \|R_{s_k t_k} - \bar{R}^{(p)}(s_k, t_k)\| + C[(s_1 - s) + (t_K - t) + \sum_{k=1}^{K-1} (s_{k+1} - t_k)].$$

It follows from the facts that the first term tends to zero and the second can be chosen arbitrarily small that $\bar{R}^{(p)}(s, t)$ tends to R_{st} as p goes to infinity.

For proving the convergence of $P^{\mu, p}$ in the finite dimensional topology *it is enough to prove the convergence for two times marginals*, which is what we have already done. This is a simple consequence of continuity of the product of finitely many real numbers.

In the last part of this proof we explain how to prove the convergence in the Skorokhod topology. We prove below that for every $\varepsilon > 0$, if p is sufficiently large, there exists Θ a measure on $D[0, 1] \times D[0, 1]$ with marginals $P^{\mu, p}$ and P (the Poisson like process generated by \tilde{N}) such that with probability greater than $1 - \varepsilon$, the Skorokhod distance between the first and the second coordinate of $(D[0, 1])^2$ is smaller than ε for the joint law Θ . In other words we prove that the Prokhorov distance between $P^{\mu, p}$ and P tends to zero. We will wrongly call *constant* on $[s, t]$ any $P^{\mu, p}$ -random trajectory x that starts close to x_k at time s and all the transitions are done from state k to itself. In the case of a P -random trajectory, x is constant on $[s, t]$ if it is continuous. In this case $x = x_k$ for some k . A more concrete way to justify this abuse is to introduce $\alpha > 0$ and consider only partitions with a sufficiently small mesh so that $|x_k(t) - \bar{x}(t)| < \varepsilon$ holds at any time for some unique k . In fact, due to the uniform continuity of the trajectories $(x_k)_{k=1, \dots, K}$, the real ε may be chosen as small as we want. Finally, it is the same to prove that the Prokhorov distance tends to zero with \mathbb{R} or $\{1, \dots, n\}$ as state space. Consider a finite set $S \subseteq [0, 1]$, and, using the convergence in the finite dimensional topology, consider a sufficiently large n and a coupling Θ such that with probability greater than $1 - \varepsilon/10$, we have $\|X_t - X_t^{\mu, p}\| \leq \varepsilon$ for every time $t \in S$. Here X and $X^{\mu, p}$ are the first and second coordinates of Θ and ε is sufficiently small to characterise the state in $\{1, \dots, n\}$. Let us call jump the discontinuities of X_t and the discontinuities in state of $X^{\mu, p}$. Lemma 4.2 and the fact that \tilde{N}_t is bounded allow us to claim that if the mesh of $S \cup \{0, 1\}$ is sufficiently small, the probability that X or $X^{\mu, p}$ has two or more jumps on some interval of the partition is smaller than $\varepsilon/10$. We can also assume that this mesh is smaller than ε , which is important for the horizontal distortion in the definition of the Skorokhod distance. On these conditions, we can easily prove that Θ is a convenient coupling for proving that the Prokhorov distance associated with the Skorokhod distance is smaller than ε . □

5 A discrete counterexample

We show that not every element of $\text{LimCurt}([\mu])$ is Markovian and that this set may have cardinality ≥ 2 . Here we take the setting of the last paragraph with $n = 3$ and

$a_1 = 2a_2 = 2a_3 = 1/2$ but we do not assume $x_1 < x_2 < x_3$. In fact

$$\begin{cases} x_1(t) = 9 - t \\ x_2(t) = 8 + 2t \\ x_3(t) = 10, \end{cases}$$

so that $x_1 < x_2 < x_3$ on $[1/2, 1[$ and $x_1 < x_3 < x_2$ on $]1, 3/2]$. We parametrise the peacock $[\mu]$ on $[1/2, 3/2]$. Note that $x_2(1) = x_3(1) = 10$. We will see that sequences of partitions not including time 1 all generate the same process independently of the sequence and that this is not a Markov process. On the contrary, if the sequence includes time 1 (at least asymptotically) there exists a unique limit curtain process independent of the sequence and this process is Markovian.

A computation allows us to state

$$\begin{aligned} M(1-h, 1+h) &= \begin{pmatrix} 1/2 \frac{20-h}{20+h} & 0 & 1/2 \frac{2h}{20+h} \\ \frac{20h+5h^2}{(20+h)(40+6h)} & 1/4 \frac{2h}{20+3h} & 1/4 \frac{20-3h}{20+h} \\ 1/4 \frac{2h}{20+3h} & 1/4 \frac{20+h}{20+3h} & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 0 & 1/4 \\ 0 & 1/4 & 0 \end{pmatrix} + 2h \begin{pmatrix} -1/40 & 0 & 1/40 \\ 1/80 & 1/80 & -1/40 \\ 1/80 & -1/80 & 0 \end{pmatrix} + o(h), \end{aligned}$$

so that

$$\tilde{M}(1-h, 1+h) = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & 1/4 \end{pmatrix}^{-1} M(1-h, 1+h) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} + O(h).$$

Hence at the order zero a partition including the interval $[1-h, 1+h[$ realises a permutation between the trajectories of x_2 and x_3 . We admits without computation that we obtain the same permutation for any interval $[1-h, 1+h'[$ with $h, h' > 0$ tending to zero.

Let us now see what happens if we have two intervals $[1-h, 1[$ and $]1, 1+h'[$. At order zero the transport plans are

$$\begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/4 & 1/4 \end{pmatrix} \text{ and } \begin{pmatrix} 1/2 & 0 \\ 0 & 1/4 \\ 0 & 1/4 \end{pmatrix},$$

which corresponds to transitions

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \\ 0 & 1/2 \end{pmatrix}.$$

After composition of these transitions we obtain

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Let us sum up. On $[1/2, 1[$ and $]1, 3/2]$ the law of the limit curtain process is exactly the same whatever time 1 is included or not. Namely we observe the same behaviour as in Theorem 4.4. For describing the limit process on $[1/2, 3/2]$ when 1 is included to the partitions, we need furthermore one Bernoulli trial. With probability $1/2$ a locally continuous trajectory will follow x_2 when arriving in 1 and with probability $1/2$ it will follow x_3 . In the other case where the partitions avoid time 1, a random trajectory has probability zero to be one of the functions x_2 or x_3 in a neighbourhood of 1.

6 A continuous counterexample.

We complete the proof of Theorem D started in Section 5. Compared to Section 5, the novelty is that all measures μ_t are absolutely continuous, which means that the non-Markovian behaviour of the limit process is not due to atoms.

Our example $[\mu]$ is illustrated in Figure 2. On the figure the grey area corresponds to the points (t, x) such that x is in the support of μ_t . Let us explain the process in simple words. Several trajectories are depicted, in particular x_1 and x_2 . They start at different points and have first an increasing piece of trajectory similar to the ones in Section 3. After a jump downwards at time t_0 they start a common constant trajectory from the smallest value of the support of μ_{t_0} at point $e^{-2t_0}/2$. The next jump happens at the same time $t_1 = t_0 + 1$ and the two trajectories split again. This behaviour indicates that the Markov property is not satisfied because between the two jumps it is possible to distinguish x_1 from x_2 by having a look at the past.

Our construction is divided into two time intervals – the *stocking* period $[0, 1]$ and *destocking* period $[1, 2]$ – whose particularities explained in the two next subsections are related to the curtain coupling. On $[0, 1]$, we wait for a jump downwards, which happens with positive probability and amount to stocking the trajectories on $[-e^2/2, -1/2]$. For the martingales of the approximating sequence, at each time t_0 of the partition a continuum of trajectories jump downwards into a small interval and start constant distinct pieces of trajectories. For the limit process all the jumps at time t_0 abut on the common value $e^{-2t_0}/2$ that is the minimum of the support of μ_{t_0} . On $[1, 2]$ the mass stocked on $[-e^2/2, -1/2]$ is destocked. For the approximating martingales, the trajectories x that have jumped at time t_0 jump at time $t_1 = t_0 + 1$ with positive probability to distinct positions that depend on their respective values at the beginning of the process. In fact for the limit process $x(t_1) = x(0) - 1/2$. All in all, in the limit we see distinct trajectories that gather at time t_0 , follow a common path and split again at time t_1 . This phenomenon is responsible for the non-Markovian behaviour of the limit process.

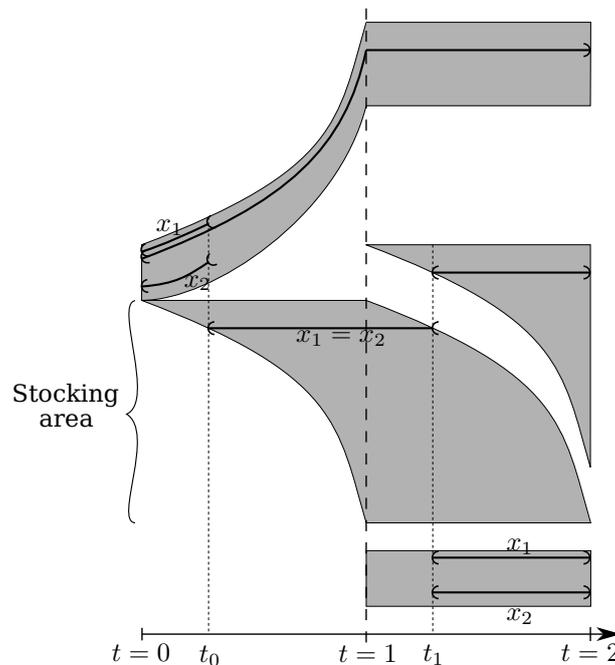


Figure 2: A non-Markovian limit curtain process associated to a peacock with absolutely continuous 1-marginals.

6.1 First step: Stocking

Consider a peacock of the type $\mu_t = \mu_t^1 + \mu_t^2$ where

- for every t , $\text{spt}(\mu_t^1) \subseteq]-\infty, b]$ and it is the restriction of μ_1^1 to $]a(t), b[$ with a a decreasing function,
- for every t , μ_t^2 is concentrated on $]b, +\infty[$.

We call such a peacock a *stocking peacock*. We give later an example that is used for our construction in Subsection 6.3. Let us first describe the shape of the left-curtain coupling π_{lc} between μ_s and μ_t for $s < t$. Due to $\mu_s^1 \leq \mu_t^1$, we have $S^{\mu_t^1}(\mu_s^1) = \mu_s^1$. It follows that π_{lc} is $(\text{Id} \otimes \text{Id})_{\#} \mu_s^1 + \pi$ where the marginals of π are μ_s^2 and $\mu_t^2 + (\mu_t^1 - \mu_s^1)$. We can conclude that a process $X^{\mu, \sigma}$ associated to a stocking peacock $[\mu]$ and a partition σ is constant from the time it meets $] - \infty, b[$.

Example 6.1. Consider the limit curtain process defined in Section 3. If we stop this martingale after the first jump, it is still a martingale. Considering then the time marginals, we obtain a peacock $\mu_t = \mu_t^1 + \mu_t^2$ with

$$\mu_t^1(dx) = \frac{1}{(-2x)^{3/2}} \lambda_{[-e^{2t}/2, -1/2]} dx \quad \text{and} \quad \mu_t^2 = e^{-2t} \lambda_{[e^{2t}/2 - e^t, e^{2t}/2]}.$$

This is clearly a stocking peacock for $a(t) = -\exp(2t)/2$ and $b = -1/2$. See the left part of Figure 2 for an illustration.

We claim without details that this peacock has a unique limit curtain process and that it is the stopped martingale itself. This fact is specific to this peacock. A proof can be derived from the techniques in Section 3.

6.2 Second step: Destocking

Consider a peacock of the following type: $\mu_t = \mu_t^1 + \mu_t^2 + \mu_t^3$ where

- for every t , $\text{spt}(\mu_t^1) \subseteq]-\infty, a[$ and $s \leq t \Rightarrow \mu_s^1 \leq \mu_t^1$,
- μ_t^2 is concentrated on $[a, b]$ and it is the restriction of μ_0^2 to $[a, b(t)]$ where b is a decreasing function,
- for every t , $\text{spt}(\mu_t^3) \subseteq]b, +\infty[$ and it is the restriction of μ_1^3 to $[c(t), +\infty[$ where c is a decreasing function.
- for all s, t if $s \leq t$ we have $\mu_s^2 - \mu_t^2 \preceq_C (\mu_t^1 - \mu_s^1) + (\mu_t^3 - \mu_s^3)$.

We call such a peacock a *destocking peacock*. Note that it may also be a stocking peacock for the writing $\mu_t^1 + (\mu_t^2 + \mu_t^3)$. The name indicates that one is “destocking” the mass in μ_t^2 . Let us describe the transition between μ_s and μ_t given by the left-curtain coupling. For $x \leq b(t)$ we have $\mu_s|_{]-\infty, x]} \leq \mu_t|_{]-\infty, x]}$ so that $S^{\mu_t}(\mu_s|_{]-\infty, x]}) = \mu_s|_{]-\infty, x]}$. Note in particular $S^{\mu_t}(\mu_s|_{]-\infty, b(t)]}) = \mu_s^1 + \mu_t^2$ because $\mu_s|_{]-\infty, b(t)]} = \mu_s^1 + \mu_t^2 \leq \mu_t$. Recall $b(s) \geq b(t)$ and consider now the shadow of $\mu_s|_{]-\infty, b(s)]} = \mu_s^1 + \mu_s^2$. Using Proposition 1.6 and Lemma 1.5 we obtain

$$\begin{aligned} S^{\mu_t}(\mu_s^1 + \mu_s^2) &= S^{\mu_t}(\mu_s^1 + \mu_t^2) + S^{\mu_t - S^{\mu_t}(\mu_s^1 + \mu_t^2)}(\mu_s^2 - \mu_t^2) \\ &= (\mu_s^1 + \mu_t^2) + S^{(\mu_t^1 - \mu_s^1) + \mu_t^3}(\mu_s^2 - \mu_t^2) \\ &= (\mu_s^1 + \mu_t^2) + [(\mu_t^1 - \mu_s^1) + (\mu_t^3 - \mu_s^3)] \\ &= \mu_t - \mu_s^3. \end{aligned}$$

From these computation it follows that the left-curtain coupling is $(\text{Id} \otimes \text{Id})_{\#} (\mu_s^1 + \mu_t^2 + \mu_s^3) + \pi$ where π is a martingale coupling of marginals $\mu_s^2 - \mu_t^2$ and $(\mu_t^1 - \mu_s^1) + (\mu_t^3 - \mu_s^3)$.

See the right part of Figure 2 for an illustration of a locally destocking peacock. The support of μ_t^3 is the union of two intervals. The peacock is not globally destocking on the maximal interval $[1, 2]$ because there is no possible value of b that satisfies $b < c(t)$ for every t . Nevertheless, the behaviour is the same.

Lemma 6.2. *Let $[\mu]$ be a destocking peacock. With the same notation as above for a, b, c and $(\mu_t^i)_{i \in \{1,2,3\}}$, we assume moreover that*

- *the functions $t \mapsto b(t)$ and $t \mapsto c(t)$ are smooth,*
- *For every $i = 1, 2, 3$, $t \mapsto \mu_t^i(\mathbb{R})$ is smooth with nonzero derivative,*
- *$t \mapsto \mu_t^1$ is smooth,*
- *$t \mapsto \mu_t^2(\mathbb{R})$ decreases from 1 to 0.*

Then there is a unique limit curtain process in $\text{LimCurt}([\mu])$. It is a locally constant process with exactly one jump.

Proof. For proving the convergence in the Skorokhod topology, we apply Theorem 12.6 of [7]. First we notice that the processes are concentrated on the path $x_{j,k,l} : t \in [0, 1] \rightarrow \mathbb{R}$ defined by $x_{j,k,l}(t) = k\mathbb{1}_{[0,j](t)} + l\mathbb{1}_{[j,1](t)}$ where $0 < j < 1$, $k \in [a, b]$ and $l \notin [a, b]$. In the subspace consisting of the latter càdlàg paths, pointwise convergence on a countable, dense subset of $[0, 1]$ provides convergence in the Skorokhod topology.

Therefore according to [7, Theorem 12.6] it is enough to check that the sequence of processes has a limit for the finite dimensional convergence. The description of the left-curtain coupling between μ_s and μ_t together with the assumptions of the lemma, provide a candidate limit process that we describe now: Start from a point $k \in [a, b]$ according to $\mu_0^2 = \mu^2$, be constant until time $b^{-1}(k)$. At this time start a second constant trajectory, either at point $c \circ b^{-1}(k)$ or in a point uniformly chosen according to $d\mu_t^1/dt|_{t=b^{-1}(k)}$ with the proper probabilities making this transition a martingale kernel. Given finitely many times t_1, \dots, t_j and a partition σ , it is enough to consider the trajectories that jump outside the intervals containing the times t_i . These trajectories can easily be coupled with the trajectories of the candidate limit process. As is the proof of Theorem 4.4, this proves that the Prokhorov distance associated with the Skorokhod distance on $D[0, 1]$ tends to zero when the mesh $|\sigma|$ tends to zero. \square

6.3 Putting the two steps together

We consider a peacock parametrised on $[0, 2]$ that we illustrate in Figure 2. When restricted to $[0, 1]$, it is simply the peacock of Example 6.1. On $[1, 2]$ it is a (locally) destocking peacock. It is consisting of four terms $\mu_t = \mu_t^1 + \mu_t^2 + (\mu_t^3 + \mu_t^3)$. We start to define the easy parts

$$\mu_t^3 = e^{-2} \lambda_{[e^2/2 - e, e^2/2]} \quad \text{and} \quad \mu_t^2(dx) = \frac{1}{(-2x)^{3/2}} \lambda_{[-e^2/2, -e^{2(t-1)}/2]}(dx)$$

that continue or restrict the two parts of μ_t after $t = 1$. We have also

$$\mu_t^1 = f(t, x) \lambda_{[-6, -5]} \quad \text{and} \quad \mu^3(dx) = g(x) \lambda_{[1 - e^{2(t-1)}/2, 1/2]}(dx).$$

where f and g are chosen in a way that the mass of $-d\mu_t^2 = \mu_{t+dt}^2 - \mu_t^2$ that is at first order $e^{-3(t-1)} \lambda_{[-e^{2(t-1)}/2 - e^{2(t-1)}/2, -e^{2(t-1)}/2]}$ is mapped on the atom $d\mu_t^3 \approx e^{2(t-1)} dt g(1 - \exp(2(t-1))/2) \delta_{2^{-1}(1 - \exp(2(t-1)))}$ for the upper part of the left-curtain coupling or linearly on $d\mu_t^1 \approx (df(t, x)/dt) \lambda_{[-6, -5]} dt$ for the down part. This is obtained for the functions

$$f(t, x) = \int_1^t \frac{e^{-(u-1)}}{1 - x - e^{-2(u-1)}/2} du \quad \text{and} \quad g(x) = [1 + \log(1 - (x - 6)^{-1})](-2(x - 1))^{3/2}.$$

Let $(\sigma^{(p)})_p$ be a sequence of partitions of $[0, 1]$ with mesh $|\sigma^{(p)}|$ going to 0. We associated any $\sigma^{(p)}$ with the partition of $[0, 2]$ consisting of the times t of $\sigma^{(p)}$ together with the times $t + 1 \in [1, 2]$. We denote the latter by $\bar{\sigma}^{(p)}$. Note that 1 is a time of this partition. We have seen in the two former paragraphs that when restricted to $[0, 1]$ or $[1, 2]$ the peacock $[\mu]$ has a unique limit curtain process and one can check that it is Markovian in both cases. For the peacock on $[0, 2]$ and the sequence $(\bar{\sigma}^{(p)})_p$ one also obtains a limit curtain process but rather surprisingly it is not a Markovian process.

The proof is technical but essentially runs as the one of Proposition 3.1. Hence our focus will be on the main ideas. If t_{k-1} and t_k are elements of $\sigma^{(p)}$, different trajectories of $X^{\mu,p}$ are jumping down at time t_k . They are mapped linearly from $[e^{2t_{k-1}}/2 - e^{t_{k-1}}, e^{2t_{k-1}}/2]$ to the small interval $[-e^{2t_{k-1}}/2, -e^{2t_k}/2]$. Between t_k and $t_{k-1} + 1$ nothing can happen to these trajectories because the left-curtain coupling is identity in their regions. At time $t_k + 1$ all the mass contained in the small interval $[-\exp(2t_{k-1})/2, -\exp(2t_k)/2]$ must jump again either to the neighbourhood of $1 - \exp(2t_k)/2$ or somewhere down into the interval $[-6, -5]$. We have parametrised the masses μ_t^1 and μ_t^3 in such a way that the jump down becomes linear when $|t_k - t_{k-1}| \leq |\bar{\sigma}^p|$ is small. Therefore the mapping of the positions between time t_{k-1} and $1 + t_k$ is almost linear. More precisely the left-curtain transitions reverses the orientation at the first jump and make it right again at the second. In the limit curtain process, there are two types of trajectories. The first type is consisting of the continuous trajectories

$$t \mapsto \min(\exp(2t)/2 - (1/2 - X_0)\exp(t), \exp(2)/2 - (1/2 - X_0)\exp(1)).$$

We are interested in the second type of trajectories that start in the same way but jump after a duration $T < 1$ (an exponential random time). After the jump the trajectory has value $-\exp(2T)/2$ on $[T, T + 1[$. There is a second jump at time $T + 1$ either to $-\exp(2T)/2 + 1$ or to a point of $[-6, -5]$ that depends of the past in the simplest manner. Indeed, this point is $X_0 - (6 - 1/2)$. Hence the limit curtain process is not Markovian.

7 Remarks and open questions

7.1 An equivalent topology

We open the section by a remark on the topology used to define LimCurt and LimCurtFD the sets of limit curtain couplings. As in [17, 5, 14] an alternative topology *with first moment* may appear more pertinent with respect to the peacock and martingale transport literature. We explain that our theorems are the same with this topology in place of \mathcal{T}_{cb} .

Let (S, ρ) be a Polish metric space, e.g. $(\mathbb{R}^j, \|\cdot\|)$ or $D[0, 1]$ with the Skorokhod distance. Recall that $\mathcal{T}_{\text{cb}}(S)$ is the topology on $\mathcal{P}(S)$ with $\mu_n \rightarrow \mu$ if and only if for all continuous bounded functions $f : S \rightarrow \mathbb{R}$, it holds $\int f d\mu_n \rightarrow \int f d\mu$. Let $\mathcal{T}_1(S)$ be the topology on $\mathcal{P}_1(S)$ defined by the continuous functions growing at most linearly – we mean that $f/(1 + \rho(x_0, \cdot))$ is bounded for some $x_0 \in S$.

According to [27, Theorem 7.12 and Remark 7.13] the two topologies coincide when restricted to a bounded set or a set $\mathcal{C} \subseteq \mathcal{P}_1(S)$ satisfying a tightness condition.

$$\lim_{R \rightarrow \infty} \limsup_{\mu \in \mathcal{S}} \int_{\rho(x, x_0) > R} \int_S \rho(x, x_0) d\mu = 0$$

for some (and hence every $x_0 \in S$). This holds for

- Any set of measures concentrated on a bounded set $B \subseteq S$. The measure involved in Theorem B and Theorem C and Theorem D are concentrated on a bounded set of $D[0, 1]$. Therefore these theorems also hold for an alternative definition of $\text{LimCurt}([\mu])$ where the topology $\mathcal{T}_1(D[0, 1])$ replaces $\mathcal{T}_{\text{cb}}(D[0, 1])$.

- For sets of uniformly integrable measures.
 - For instance for $S = \mathbb{R}$ and $\mu_1 \in \mathcal{P}_1(\mathbb{R})$ the set $\mathcal{C} = \{\nu \in \mathcal{P}_1, \nu \preceq_C \mu_1\}$ is consisting of uniformly integrable measures. The proof is similar to the proof of uniform integrability in the theory of L^1 martingales (see Exercise 1.1 [11]).
 - for $S = \mathbb{R}^j$ and $\mu_1 \in \mathcal{P}_1(\mathbb{R})$ the set

$$\mathcal{D} = \{\pi \in \mathcal{P}(\mathbb{R}^j), \forall i \leq j, (\text{proj}_i)_\# \pi \preceq_C \mu_1\}$$

is consisting of uniformly integrable measures, as a consequence of the previous example.

Therefore, in Theorem A the set $\text{LimCurtFD}([\mu])$ could have been defined by the topology $\mathcal{T}_1(\mathbb{R}^{[0,1]})$ in place of $\mathcal{T}_1(\mathbb{R}^{[0,1]})$.

7.2 Markov-Lipschitz property

To the best of our knowledge, the proofs of Kellerer's Theorem in the literature all rely on the so-called *Lipschitz-(Markov) property*, that already appeared in Kellerer's paper [16, Definition 2 and 3]. We mean in particular the very elegant theory of Lowther [22, 21, 20] on *almost continuous diffusions* (abbreviated ACD) and the proof by Hirsch, Roynette and Yor that is presented in [13]. The latter is based on Pierre's uniqueness theorem for the Fokker-Planck equation [11, Chapter 6], a convolution technique and the important contributions by Lowther. As examples of the strength of Lowther's results, we mention the uniqueness of ACD martingales $\Phi(\mu)$ for all continuous peacocks $(\mu_t)_t$, the continuity of the map Φ and the fact that any other continuous Ψ associating some not a priori ACD martingale $\Psi(\mu)$ to continuous peacocks $(\mu_t)_t$ satisfies $\Psi = \Phi$ [20, Theorems 1.3, 1.4 and 1.5].

7.3 Open problems

We propose the following open problems:

- Can the set $\text{LimCurt}([\mu])$ be empty for some right-continuous peacock? We conjecture that the answer is no.
- How many Markov processes may $\text{LimCurt}([\mu])$ contain? We conjecture that there is exactly one.

One may ask the same questions for peacocks without continuity assumptions, or the second question for $\text{LimCurtFD}([\mu])$. Also the same questions make sense for all other couplings than the left-curtain coupling, still using the Markov composition. In view of the Kamae–Krengel Theorem [15] and Pass' transport problem [24], the case of the quantile coupling seems to be of particular interest.

References

- [1] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré, *Gradient flows in metric spaces and in the space of probability measures*, second ed., Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008. MR-2401600
- [2] Andrew D. Barbour, Lars Holst, and Svante Janson, *Poisson approximation*, Oxford Studies in Probability, vol. 2, The Clarendon Press, Oxford University Press, New York, 1992, Oxford Science Publications. MR-1163825
- [3] Mathias Beiglböck, Alexander M. G. Cox, and Martin Huesmann, *Optimal transport and Skorokhod embedding*, Invent. Math. **208** (2017), no. 2, 327–400. MR-3639595

- [4] Mathias Beiglböck, Martin Huesmann, and Florian Stebegg, *Root to Kellerer*, Séminaire de Probabilités XLVIII, Lecture Notes in Math., vol. 2168, Springer, Cham, 2016, pp. 1–12. MR-3618124
- [5] Mathias Beiglböck and Nicolas Juillet, *On a problem of optimal transport under marginal martingale constraints*, Ann. Probab. **44** (2016), no. 1, 42–106. MR-3456332
- [6] Jean-David Benamou and Yann Brenier, *A numerical method for the optimal time-continuous mass transport problem and related problems*, Monge Ampère equation: applications to geometry and optimization (Deerfield Beach, FL, 1997), Contemp. Math., vol. 226, Amer. Math. Soc., Providence, RI, 1999, pp. 1–11. MR-1660739
- [7] Patrick Billingsley, *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons Inc., New York, 1999, A Wiley-Interscience Publication. MR-1700749
- [8] Oclide José Dotto, *Dilations and stochastic processes*, J. Math. Anal. Appl. **118** (1986), no. 2, 410–421. MR-852168
- [9] Pierre Henry-Labordère, Xiaolu Tan, and Nizar Touzi, *An explicit martingale version of the one-dimensional Brenier’s theorem with full marginals constraint*, Stochastic Process. Appl. **126** (2016), no. 9, 2800–2834. MR-3522302
- [10] Pierre Henry-Labordère and Nizar Touzi, *An explicit martingale version of the one-dimensional Brenier theorem*, Finance Stoch. **20** (2016), no. 3, 635–668. MR-3519164
- [11] Francis Hirsch, Christophe Profeta, Bernard Roynette, and Marc Yor, *Peacocks and associated martingales, with explicit constructions*, Bocconi & Springer Series, vol. 3, Springer, Milan, 2011. MR-2808243
- [12] Francis Hirsch and Bernard Roynette, *On \mathbb{R}^d -valued peacocks*, ESAIM Probab. Stat. **17** (2013), 444–454. MR-3066388
- [13] Francis Hirsch, Bernard Roynette, and Marc Yor, *Kellerer’s theorem revisited*, Asymptotic Laws and Methods in Stochastics. Volume in Honour of Miklos Csorgo (Springer, ed.), Fields Institute Communications Series, 2014. MR-3409839
- [14] Nicolas Juillet, *Stability of the shadow projection and the left-curtain coupling*, Ann. Inst. Henri Poincaré Probab. Stat. **52** (2016), no. 4, 1823–1843. MR-3573297
- [15] Teturo Kamae and Ulrich Krengel, *Stochastic partial ordering*, Ann. Probab. **6** (1978), no. 6, 1044–1049 (1979). MR-512419
- [16] Hans G. Kellerer, *Markov-Komposition und eine Anwendung auf Martingale*, Math. Ann. **198** (1972), 99–122. MR-0356250
- [17] Hans G. Kellerer, *Integraldarstellung von Dilationen*, Transactions of the Sixth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes (Tech. Univ., Prague, 1971; dedicated to the memory of Antonín Špaček), Academia, Prague, 1973, pp. 341–374. MR-0356249
- [18] Steven G. Krantz and Harold R. Parks, *A primer of real analytic functions*, second ed., Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks], Birkhäuser Boston, Inc., Boston, MA, 2002. MR-1916029
- [19] Stefano Lisini, *Characterization of absolutely continuous curves in Wasserstein spaces*, Calc. Var. Partial Differential Equations **28** (2007), no. 1, 85–120. MR-2267755
- [20] George Lowther, *Fitting Martingales To Given Marginals*, ArXiv e-prints (2008).
- [21] George Lowther, *Properties of Expectations of Functions of Martingale Diffusions*, ArXiv e-prints (2008).
- [22] George Lowther, *Limits of one-dimensional diffusions*, Ann. Probab. **37** (2009), no. 1, 78–106. MR-2489160
- [23] Robert J. McCann, *A convexity principle for interacting gases*, Adv. Math. **128** (1997), no. 1, 153–179. MR-1451422
- [24] Brendan Pass, *On a class of optimal transportation problems with infinitely many marginals*, SIAM J. Math. Anal. **45** (2013), no. 4, 2557–2575. MR-3093868
- [25] Svetlozar T. Rachev and Ludger Rüschendorf, *Mass transportation problems. Vol. I*, Probability and its Applications (New York), Springer-Verlag, New York, 1998, Theory. MR-1619170

- [26] Volker Strassen, *The existence of probability measures with given marginals*, Ann. Math. Statist. **36** (1965), 423–439. MR-0177430
- [27] Cédric Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003. MR-1964483

Acknowledgments. I wish to thank Mathias Beiglböck for helpful discussions that started from the elaboration of [5]. I'm also indebted to a particularly careful referee for numerous valuable suggestions. I thank my colleague Vincent Vigon for reading the paper and valuable suggestions as well as Ana Rechtman, Jean Bérard and Pierre Py for providing references. A lot of thanks go to Harold Gretton for his English and editorial advice, and I also wish to thank Charles Boubel, Emmanuel Opshtein, and Marcus Slupinski. Finally I thank Xiaolu Tan and Michel Émery for making me aware of [9] and [13] respectively.

Electronic Journal of Probability

Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS¹)
- Easy interface (EJMS²)

Economical model of EJP-ECP

- Non profit, sponsored by IMS³, BS⁴, ProjectEuclid⁵
- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund⁶ (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

¹LOCKSS: Lots of Copies Keep Stuff Safe <http://www.lockss.org/>

²EJMS: Electronic Journal Management System <http://www.vtex.lt/en/ejms.html>

³IMS: Institute of Mathematical Statistics <http://www.imstat.org/>

⁴BS: Bernoulli Society <http://www.bernoulli-society.org/>

⁵Project Euclid: <https://projecteuclid.org/>

⁶IMS Open Access Fund: <http://www.imstat.org/publications/open.htm>