INSTITUT POLYTECHNIQUE DE PARIS

Journal de l'École polytechnique
Mathématiques

Charles Boubel & Nicolas Juillet

The Markov-quantile process attached to a family of marginals

Tome 9 (2022), p. 1-62.

<http://jep.centre-mersenne.org/item/JEP_2022__9__1_0>

© Les auteurs, 2022.
Certains droits réservés.

Cet article est mis à disposition selon les termes de la licence
LICENCE INTERNATIONALE D'ATTRIBUTION CREATIVE COMMONS BY 4.0.
https://creativecommons.org/licenses/by/4.0/

L'accès aux articles de la revue « Journal de l'École polytechnique — Mathématiques » (http://jep.centre-mersenne.org/), implique l'accord avec les conditions générales d'utilisation (http://jep.centre-mersenne.org/legal/).

Publié avec le soutien
du Centre National de la Recherche Scientifique

CENTRE MERSENNE

Publication membre du
Centre Mersenne pour l'édition scientifique ouverte
www.centre-mersenne.org

Journal de l'École polytechnique
Mathématiques

Tome 9, 2022, p. 1–62 DOI: 10.5802/jep.177

THE MARKOV-QUANTILE PROCESS ATTACHED TO A FAMILY OF MARGINALS

by Charles Boubel & Nicolas Juillet

Abstract. — Let $\mu = (\mu_t)_{t\in\mathbb{R}}$ be any 1-parameter family of probability measures on $\mathbb{R}$. Its quantile process $(G_t)_{t\in\mathbb{R}} :]0;1[\to \mathbb{R}^{\mathbb{R}}$, given by $G_t(\alpha) = \inf\{x \in \mathbb{R} : \mu_t(]-\infty;x]) > \alpha\}$, is not Markov in general. We modify it to build the Markov process we call "Markov-quantile": there is a *unique* Markov process X with marginals μ_t, being a limit for the finite dimensional topology of quantile processes where the past is made independent of the future at finitely many times (many non-Markovian limits exist in general). Strikingly, no regularity is required for the family μ. Moreover, if μ is increasing for the stochastic order, X has increasing trajectories. This is an analogue of a result of Kellerer dealing with the convex order. In a companion paper [8] it is also proved that if μ is absolutely continuous in the Wasserstein space $\mathrm{P}_2(\mathbb{R})$, X is solution of a Benamou–Brenier transport problem with marginals μ_t, providing a Markov representation of the continuity equation, unique in the sense above.

Résumé (Le processus Markov-quantile attaché à une famille de marges)

Soit $\mu = (\mu_t)_{t\in\mathbb{R}}$ une famille à un paramètre de mesures de probabilité sur $\mathbb{R}$. Son processus quantile $(G_t)_{t\in\mathbb{R}} :]0;1[\to \mathbb{R}^{\mathbb{R}}$, donné par $G_t(\alpha) = \inf\{x \in \mathbb{R} : \mu_t(]-\infty;x]) > \alpha\}$, n'est en général pas markovien. Nous le modifions pour construire le processus markovien que nous nommons « Markov-quantile » : il existe un *unique* processus markovien X de marges μ_t qui est limite, pour la convergence fini-dimensionnelle, de processus quantiles dont le passé est rendu indépendant du futur en un nombre fini d'instants (beaucoup de limites non markoviennes existent en général). Il est frappant qu'aucune hypothèse de régularité sur la famille μ n'est nécessaire. En outre, si μ est croissante pour l'ordre stochastique, les trajectoires de X sont croissantes. Ceci est un analogue d'un résultat de Kellerer traitant de l'ordre convexe. Dans un article associé [8] on montre aussi que si μ est absolument continue dans l'espace de Wasserstein $\mathrm{P}_2(\mathbb{R})$, X est solution d'un problème de transport de Benamou–Brenier avec marges μ_t, et fournit donc une représentation markovienne de l'équation de continuité, unique dans le sens donné plus haut.

Contents

Mathematical subject classification (2020). — 60A10, 28A33, 60J25, 35Q35, 60G44, 49J55.
Keywords. — Markov process, quantile process, optimal transport, continuity equation, increasing process, Kellerer's theorem, martingale optimal transport, peacocks, copula.

1. Introduction

Rather surprisingly, the basic question of the existence of a *Markov* martingale with prescribed marginals $(\mu_t)_{t\in\mathbb{R}}$ remains open since the partial result [28] of 1972. More generally, only the method of [28], rediscovered or revisited by several authors, seems to be available to provide Markov processes with prescribed marginals—and satisfying an additional constraint, as being a martingale.

In this paper we give two new methods—one of them rests on order arguments—to solve this type of question. We apply them to measures $(\mu_t)_{t\in\mathbb{R}}$ in stochastic order.

This detailed introduction explains precisely this context and all our results.

1.1. Formal structure of the article. — We prove three main theorems, stated and labeled as A, B, and C in this introduction and proved in the order C, A, B in the paper; see also their interdependence in Figure 1, where appears also the Main Theorem of a companion paper [8]. Theorem A answers the Main Problem p. 3 below, and is a general theoretical result in Probability Theory; it builds a certain stochastic process with given marginals: the *Markov-quantile* process. Theorem B gives a convergence result to it. Theorems C presents an application of the Markov-quantile process to another context (Martingales and a theorem of Kellerer [28, 29]), giving by the way the Main Problem additional motivations, see Section 1.3 and Figure 1. We also prove Theorem 2.26, linked with Theorem C. Being a bit more technical, it is not stated in this introduction. Another application (to Optimal Transport and the Continuity Equation) in presented in [8].

In this introduction we first give the very necessary notions to state Theorems A–B as quickly as possible in Section 1.2, then state Theorem C in Section 1.3. We quickly comment on [8] in Section 1.4. We slow the flow in Section 1.5 to give a qualitative insight into the Main Problem and its difficulties, that also shows why it is a natural problem in itself. We give in Section 1.6 the structure of the article and in Section 1.7 an index of our notation.

This paper treats of Measure, Probability, and Transport Theories. To be understood by a large panel of readers, we give the definitions of the more specific notions of each of these fields, or *Reminders* about them if needed.

1.2. Our results and their motivation. — Take $(E_\tau)_{\tau\in\mathsf{T}}$ a (finite or not) family of measurable spaces; $\prod_{\tau\in\mathsf{T}} E_\tau$ is endowed with its cylindrical σ-algebra, generated by the preimages of those of the factors by the projections.

Reminder 1.1. — A *process* is a family $X = (X_\tau)_{\tau\in\mathsf{T}}$ of measurable maps

$$X_\tau : \Omega \to E_\tau,$$

called random variables, defined on the same probability space (Ω, P). In this article, contrary to what may be considered usual, no measurability condition is required on $\Omega \times \mathsf{T}$. For every $\mathsf{T}' \subset \mathsf{T}$, $(X_\tau)_{\tau\in\mathsf{T}'}$ defines a map $F_{\mathsf{T}'}$ from Ω to $\prod_{\tau\in\mathsf{T}'} E_\tau$. The *law* of $(X_\tau)_{\tau\in\mathsf{T}'}$ is the pushed-forward probability measure $(F_{\mathsf{T}'})_\#\mathsf{P}$ on $\prod_{\tau\in\mathsf{T}'} E_\tau$, which is also called the *marginal law* of the measure $(F_{\mathsf{T}})_\#\mathsf{P}$ on $\prod_{\tau\in\mathsf{T}'} E_\tau$.

Now let μ_τ be a probability measure on E_τ, for each τ.

Notation 1.2. — For all measurable space E, $\mathrm{M}(E)$ and $\mathrm{P}(E)$ are the spaces of measures and probability measures on E. If $\mathrm{T}' \subset \mathrm{T}$, $\mathrm{proj}^{\mathrm{T}'}$ is the projection $\prod_{\tau\in\mathrm{T}} E_\tau \to \prod_{\tau\in\mathrm{T}'} E_\tau$; in case $\mathrm{T} = \{\tau_1,\dots,\tau_m\}$ is finite, $\mathrm{proj}^{\tau_1,\dots,\tau_m}$ means $\mathrm{proj}^{\{\tau_1,\dots,\tau_m\}}$. When $P \in \mathrm{P}\big(\prod_{\tau\in\mathrm{T}} E_\tau\big)$ and $s < t$, P^s stands for $(\mathrm{proj}^s)_\# P$ and $P^{s,t}$ for $(\mathrm{proj}^{s,t})_\# P$, and $\mathrm{Marg}((\mu_\tau)_{\tau\in\mathrm{T}})$ denotes $\{P \in \mathrm{P}\big(\prod_{\tau\in\mathrm{T}} E_\tau\big) : \forall\tau \in \mathrm{T}, (\mathrm{proj}^\tau)_\# P = \mu_\tau\}$. When not otherwise specified, what we call the marginals of P are its marginals P^s on a single factor.

Reminder 1.3

(a) If $P \in \mathrm{Marg}((\mu_\tau)_{\tau\in\mathrm{T}})$, setting $\Omega := (\prod_{\tau\in\mathrm{T}} E_\tau, P)$ and $X = (X_\tau)_{\tau\in\mathrm{T}} := (\mathrm{proj}^\tau)_{\tau\in\mathrm{T}}$ we get a process called the *canonical process*, of law P. For this reason, by an abuse of language—e.g., in the title of this article—, we sometimes call *process* a probability measure on a product space. For the same reason we may also see $\mathrm{Marg}((\mu_\tau)_{\tau\in\mathrm{T}})$ as the set of the processes $(X_\tau)_{\tau\in\mathrm{T}}$ such that $\mathrm{Law}(X_\tau) = \mu_\tau$ for all τ.

(b) If $\sharp\mathrm{T} = 2$, i.e., if $\mu \in \mathrm{P}(E)$ and $\nu \in \mathrm{P}(E')$, a measure $P \in \mathrm{Marg}(\mu,\nu)$ is called a *transport (plan)* from μ to ν, or a *coupling* between μ and ν.

Here is our problem. It is stated for any family of measures, without any assumption of regularity in the parameter t.

Main Problem. — If $\mu = (\mu_t)_{t\in\mathrm{R}}$ is a one-parameter family of probability measures on R, we want to build a measure $\mathrm{MQ} \in \mathrm{Marg}(\mu)$ that at once:

(a) is Markov,

(b) resembles as much as possible the quantile measure $\mathrm{Q} \in \mathrm{Marg}(\mu)$.

Let us explain these two points. Take $P \in \mathrm{Marg}((\mu_t)_{t\in\mathrm{R}})$ and $(X_t)_{t\in\mathrm{R}}$ a process of law P; point (a) means:

$$\forall s \in \mathrm{R}, \forall t > s, \quad \mathrm{Law}(X_t | (X_u)_{u\leqslant s}) = \mathrm{Law}(X_t | X_s), \tag{1}$$

where $\mathrm{Law}(X_t | (X_u)_{u\leqslant s})$ is the law of X_t conditionally to the σ-algebra generated by the X_u for $u \leqslant s$. See also Definition 2.11, where the Markov property is introduced only through notions defined in this article.

For point (b), $\mathrm{Q} \in \mathrm{Marg}((\mu_t)_{t\in\mathrm{R}})$ is defined as follows.

Reminder 1.4 (The quantile process). — The quantiles of a measure $\mu \in \mathrm{P}(\mathrm{R})$ generalize the notion of the median, which is the quantile of level $1/2$. The *quantile of μ of level α* is the smallest real number $x_\mu(\alpha)$ such that $\mu(]-\infty, x_\mu(\alpha)]) \geqslant \alpha$ and $\mu([x_\mu(\alpha), +\infty[) \geqslant 1-\alpha$. The quantile process $X = (X_\tau)_{\tau\in\mathrm{T}}$, defined on $\Omega = [0,1]$ with the Lebesgue measure, is given by $X_t(\alpha) = x_{\mu_t}(\alpha)$, and we denote $\mathrm{Law}(X)$ by $\mathrm{Q} \in \mathrm{Marg}((\mu_t)_{t\in\mathrm{T}})$. In particular, $\mathrm{Law}(X_t) = \mu_t$ for every $t \in \mathrm{T}$. See Definition 3.23 for full details.

Alternatively, Q is also implicitly defined by the fact that it is the unique measure such that for all $x \in \mathbb{R}$ and all $t > s$, if X is a process of law Q:

$$\text{(2)} \qquad \mathrm{Law}(X_t | X_s \leqslant x) = \min\{\mathrm{Law}(Y_t | Y_s \leqslant x) : \mathrm{Law}(Y) \in \mathrm{Marg}((\mu_t)_{t\in\mathbb{R}})\},$$

where this minimum is with respect to the stochastic order:

Reminder 1.5. — The *stochastic order* on $\mathcal{P}(\mathbb{R})$ is defined by: $\mu \preccurlyeq_{\mathrm{sto}} \nu$ if, for all $x \in \mathbb{R}$, $\mu(]-\infty, x]) \geqslant \nu(]-\infty, x])$.

The idea of the minimality property (2) is the following. Among the processes with marginals $(\mu_t)_t$, the quantile process is the one that perturbs the least possible the order in which the elements of mass are distributed at each time: for any two times $s < t$, it transports the mass below any constant x, at time s, to the smallest possible measure, for the stochastic order, at time t.

The problem is that Q is not Markov in general, see Remark 1.8(a) a bit below. In view of its definition through (2), we seek some Markov process MQ, if it exists, such that:

– processes of law MQ satisfy a version of (2) *among Markov processes*,

– like for $\mathrm{Q} \in \mathrm{Marg}((\mu_t)_{t\in\mathbb{R}})$, if $(\mu_t)_{t\in\mathbb{R}}$ is increasing for $\preccurlyeq_{\mathrm{sto}}$, some processes $(X_t)_{t\in\mathbb{R}}$ of law $\mathrm{MQ} \in \mathrm{Marg}((\mu_t)_{t\in\mathbb{R}})$ are increasing, i.e., the functions $t \mapsto X_t(\omega)$ are,

– like for Q, the couplings $(\mathrm{proj}^{s,t})_\# \mathrm{MQ}$ of MQ have an *increasing kernel* (or briefly, MQ has increasing kernels), as follows (see Definition 3.11 for alternative definitions not resorting to conditional laws):

Definition 1.6. — Take $P \in \mathrm{Marg}((\mu_t)_t)$ and for $s < t$, set $P^{s,t} = (\mathrm{proj}^{s,t})_\# P$. We call *kernel* of $P^{s,t}$ the data of the conditional measures $\mathrm{Law}(X_t | X_s = x)$, where X is a process of law P; just below we denote it by $(P_x^{s,t})_{x\in\mathbb{R}}$.

We say that $P^{s,t}$ has *increasing kernel* if $x \leqslant y$ implies $P_x^{s,t} \preccurlyeq_{\mathrm{sto}} P_y^{s,t}$. We say that P has increasing kernels if $P^{s,t}$ has for every $s < t$.

Be careful that the following convention is now used throughout. Our answer to the Main Problem is Theorem A below.

Convention 1.7. — When we introduce finite sets $\{r_1, \ldots, r_m\}$ or m-tuples $(r_k)_{k=1}^m$ of real numbers, we mean implicitly that $r_1 < \cdots < r_m$.

Theorem A. — *Let $(\mu_t)_{t\in\mathbb{R}}$ be a family of probability measures on $\mathbb{R}$.*

(a) *There exists a unique measure $\mathrm{MQ} \in \mathrm{Marg}((\mu_t)_{t\in\mathbb{R}})$ such that:*

(i) MQ *is Markov,*

(ii) MQ *has increasing kernels,*

(iii) MQ *has minimal couplings (alias transports) among the measures satisfying* (i) *and* (ii)*, in the sense that it satisfies* (2)*, where the minimum is taken among processes $(Y_t)_t$ satisfying* (i) *and* (ii)*.*

It is also the unique process satisfying (i) *above and:*

(iv) *For each $s < t$, there is a sequence $(R_n^{s,t})_{n\in\mathbb{N}}$ of finite subsets of $[s,t]$ such that $\mathsf{MQ}^{s,t}$ is the weak limit of the sequence $(\mathsf{Q}^{s,t}_{[R_n^{s,t}]})_{n\in\mathbb{N}}$, where for $R = \{r_1, \dots, r_m\} \subset [s,t]$, $\mathsf{Q}^{s,t}_{[R]}$ is the result of the following composition, sometimes also called "product":*

$$\mathsf{Q}^{s,r_1} . \mathsf{Q}^{r_1,r_2} . \cdots . \mathsf{Q}^{r_{m-1},r_m} . \mathsf{Q}^{r_m,t}$$

of the quantile couplings $\mathsf{Q}^{r_i,r_j} \in \mathrm{Marg}(\mu_{r_i}, \mu_{r_j})$.

Moreover:

(b) *If $(\mu_t)_{t\in\mathbb{R}}$ is increasing for the stochastic order, i.e., $s \leqslant t \Rightarrow \mu_s \preceq_{\mathrm{sto}} \mu_t$, there exists a process $X = (X_t)_{t\in\mathbb{R}} : \Omega \to \mathbb{R}^{\mathbb{R}}$ of law MQ with increasing trajectories, i.e., such that $t \mapsto X_t(\omega)$ is an increasing function, for all $\omega \in \Omega$.*

We recall the notion of weak limit in Reminder 1.11, and that of composition, which corresponds to the composition of kernels, in Section 2.1 and in particular 2.1.2. Besides, one may also see $\mathsf{Q}^{s,r_1} . \mathsf{Q}^{r_1,r_2} . \cdots . \mathsf{Q}^{r_m,t}$ as $(\mathrm{proj}^{s,t})_\# \mathsf{Q}_{[R]}$, where $\mathsf{Q}_{[R]}$ is introduced in Proposition 1.10 below.

Informally, we may interpret Theorem A(a) as the following answer to the Main Problem: MQ is the Markov process whose "infinitesimal transitions" are those of the quantile process. Besides, an immediate consequence of Theorem A(a) is the important point (b) of Remark 1.8. Finally, we neither proved nor disproved that MQ is strongly Markovian. The question remains hence open and we think it is a significant one—see Open problem 5.5.4 and Example 5.4 (where it is strongly Markovian).

Remark 1.8

(a) In general, the quantile measure $\mathsf{Q} \in \mathrm{Marg}((\mu_t)_{t\in\mathbb{R}})$ is not Markov. Take, e.g., $\mu_t = \frac{1}{2}(\delta_0 + \delta_1)$ for $t \neq 0$ and $\mu_0 = \delta_0$, then $\mathsf{Q} = \mathrm{Law}((X_t)_{t\in\mathbb{R}})$ with $(X_t)_{t\in\mathbb{R}} = 0$ or $(X_t)_{t\in\mathbb{R}} = \mathbb{1}_{\mathbb{R}^*}$, both with probability 1/2. Hence for all $t > 0$, $\mathrm{Law}(X_t | X_0) = \mu_t$. Now $\mathrm{Law}(X_t | X_0 = 0, X_{-1} = i) = \delta_i$ for $i \in \{0, 1\}$, so that (1) is false for $u = 0$. As proved in [23, Prop. 3], Q is Markov except if a phenomenon of this type happens, see details in Example 5.11. In particular, Q is Markov when the measures μ_t have no atoms (see a direct proof in Remark 4.35).

(b) When $\mathsf{Q} \in \mathrm{Marg}((\mu_t)_{t\in\mathbb{R}})$ is Markov, $\mathsf{MQ} = \mathsf{Q}$. Indeed, then, $\mathsf{Q}^{s,t} = \mathsf{Q}^{s,r_1} . \cdots . \mathsf{Q}^{r_n,t}$ for every $s \leqslant r_1 < \cdots < r_m \leqslant t$. Hence according to (iv), $\forall s, \forall t > s$, $\mathsf{MQ}^{s,t} = \mathsf{Q}^{s,t}$. Since both processes are Markov they coincide in law (see Corollary 2.13).

Remark 1.9 (Justification of the name "Markov-quantile"). — While Properties (i) and (ii) of Theorem A(a) are satisfied by the product measure (law of the independent process) $\bigotimes_{t\in\mathbb{R}} \mu_t$, the quantile process Q satisfies (ii) and (iii) in the sense that it satisfies (ii) and its couplings $\mathsf{Q}^{s,t}$ are minimal among those of the measures satisfying (ii). In fact, Theorem A is constructive and builds MQ as a modification of Q; therefore we call this measure MQ the *"Markov-quantile" measure* attached to $(\mu_t)_{t\in\mathbb{R}}$.

In fact, a deeper convergence statement holds than that resulting from point (iv) above. Indeed we introduce the following notion of a measure in $\mathrm{Marg}((\mu_t)_{t\in\mathbb{R}})$ "turned into a Markov law at a finite set of instants", denoted in a way that is consistent with the notation of Theorem A(iv).

Proposition/Notation 1.10. — *If $P \in \mathrm{Marg}((\mu_t)_{t\in\mathbb{R}})$ and $R \subset \mathbb{R}$ is finite, there is a unique measure in $\mathrm{Marg}((\mu_t)_{t\in\mathbb{R}})$, denoted by $P_{[R]}$, such that:*

– *$P_{[R]}$ is the law of a family of variables $(X_t)_{t\in\mathbb{R}}$ that is "Markov at the instants of R" i.e., (1) holds with "$\forall s \in R$" instead of "$\forall s \in \mathbb{R}$",*

– *for the closure I of each connected component of $\mathbb{R} \smallsetminus R$, $(\mathrm{proj}^I)_\# P_{[R]} = (\mathrm{proj}^I)_\# P$.*

This proposition follows from the way we define $P_{[R]}$ in Definition 4.18, using the concatenation of transport plans given by Definition 2.8. We show:

Theorem B. — *There is an increasing sequence $(R_n)_{n\in\mathbb{N}}$ of finite subsets of $\mathbb{R}$ such that $\mathsf{Q}_{[R_n]} \in \mathrm{Marg}((\mu_t)_{t\in\mathbb{R}})$ converges weakly to $\mathrm{M}\mathsf{Q}$.*

Reminder 1.11. — A sequence $(P_n)_n$ of (probability) measures on some measurable topological space E *converges weakly* to P if, for all bounded continuous function f, $\int f \mathrm{d}P_n \to \int f \mathrm{d}P$. For $E = \mathbb{R}^{\mathbb{R}}$ with the weak topology, this convergence amounts to the weak convergence of all finite marginals.

Remark 1.12. — Of course, not every sequence $(R_n)_{n\in\mathbb{N}}$ is admissible for Theorem B. In fact, we prove a more precise version of Theorem B, see Theorem 4.21. It introduces the notion of *essential atomic times* of $(\mu_t)_t$ that turn out to be the times contained in $\bigcup_n R_n$ for all sequence $(R_n)_{n\in\mathbb{N}}$ admissible for Theorem B; see also Remark 1.26 in this introduction.

Our problem: a classical type of question. — The problem of defining a measure or a process P with given marginals and additional properties is a general problem that includes the Main Problem and has already been explored several times in pure and applied Probability Theory as well as in Analysis or Dynamics. Without claiming exhaustiveness on this rich topic we review some research streams and provide references.

A result related to Theorem A(b) is proved by Kamae and Krengel in [26]. The measures $(\mu_t)_{t\in\mathbb{R}}$ are in $\mathcal{P}(E)$, where E is a partially ordered Polish space. Assuming the measures in stochastic order, in a suitable sense, the authors prove that there exists an increasing process in $\mathrm{Marg}(\mu)$. Other orders can be considered together with expected properties on the processes. For $E = \mathbb{R}^d$, Chapter 8 of [43] proposes plenty of orders. In Stochastic Analysis and Mathematical Finance, the topic of peacocks and their associated martingales is closely related to our problem. "Peacock" stands for PCOC: Processus Croissant pour l'Ordre Convexe (French), that is, increasing process for the convex order. One aims at defining a martingale in $\mathrm{Marg}(\mu)$, where $\mu = (\mu_t)_{t\in\mathbb{R}}$ are the marginals of some peacock, using various techniques. Most of the time the peacock is part of a specific class, so the purpose is more specific than the work of Kellerer presented in Section 1.3. In most of this literature (see e.g., [14, 11, 9, 36, 24, 16, 19, 22, 39] and the references therein) the martingales may or not be Markov. The papers by Lowther [35, 34] on limits of diffusion processes for the finite dimensional convergence permitted some authors to refocus on the Markov setting (see, e.g., [3, 20, 23]), rediscovering Kellerer's work by the way. Lowther's proof consists in adapting the local volatility coefficient of a SDE without drift—as indicated by Dupire in his very influential note [11] on financial engineering—in order to match the marginals of $(\mu_t)_{t\in\mathbb{R}}$ mollified in time and space. The solution being Markov-Lipschitz the "demollification" happens in such a way that the Markov property is preserved at the limit.

At last, an important example in the topic are the fake Brownian motions, that are processes sharing some of the properties characterizing the standard Brownian motion: they are continuous Markov martingales with marginals $\mu_t = \mathrm{N}(0,t)$. See, e.g., [36, 15, 1, 38, 21, 16] for examples of constructions.

1.3. Relations to Kellerer's Theorem. — If you forget about MQ itself, Theorem A(b) gives the following existence property.

Corollary 1.13. — *If $(\mu_t)_{t\in\mathbb{R}} \in \mathcal{P}(\mathbb{R})^{\mathbb{R}}$ is increasing for $\preceq_{\mathrm{sto}}$, there exists a Markov process $X = (X_t)_{t\in\mathbb{R}} : \Omega \to \mathbb{R}^{\mathbb{R}}$ such that $\mathrm{Law}(X) \in \mathrm{Marg}((\mu_t)_t)$ and that the trajectories $t \mapsto X_t(\omega)$ are increasing.*

This extends to the case of the stochastic order the famous theorem of Kellerer on martingales and submartingales with given marginals, Theorem 1.17 below. Our Theorem C recovers, with a different proof, Kellerer's result, as well as (simultaneously) Corollary 1.13 on increasing processes. The proof of Theorem C is also completely independent of that of Theorem A. Moreover, the method used to show it leads to an existence statement for certain Markov processes, Theorem 2.26, omitted in this introduction. To state Theorem C we need to recall two definitions.

Definition 1.14. — Two measures μ and ν on $\mathbb{R}$, with finite first moments, are said to be in convex order $\preceq_C$, respectively in convex stochastic order $\preceq_{C,\mathrm{sto}}$, if for every convex, respectively convex increasing function φ:

$$\int \varphi \,\mathrm{d}\mu \leqslant \int \varphi \,\mathrm{d}\nu. \tag{3}$$

Notice that $\mu \preceq_{\mathrm{sto}} \nu$ if and only if (3) holds for (bounded) increasing functions φ. Now we define a martingale. We do it in the Markov framework (all we need), where this is a bit simpler. We add in (c) a terminology of our own.

Definition 1.15

(a) A measure P on $(\mathbb{R}^d)^2$ is a *martingale coupling* if for every non-negative continuous bounded function $f : \mathbb{R}^d \to \mathbb{R}$:

$$\iint f(x)(y - x)\,\mathrm{d}P(x,y) = 0. \tag{4}$$

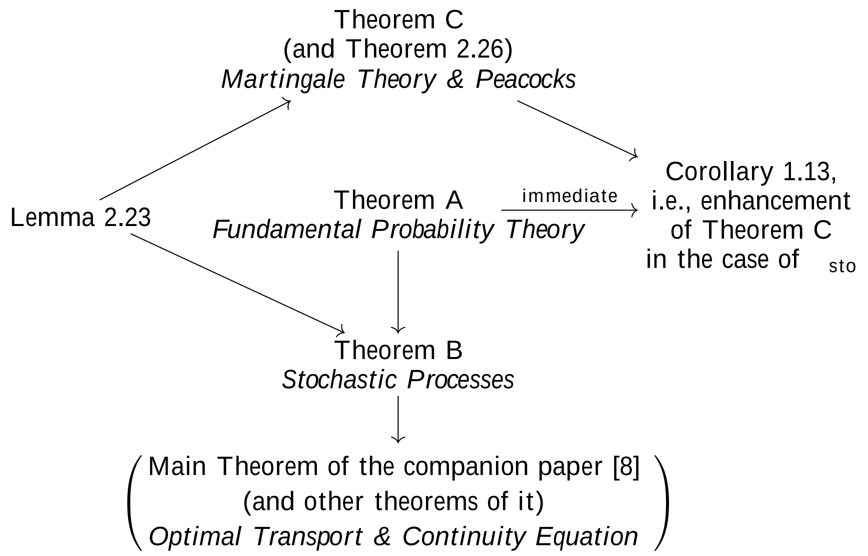

Figure 1. Main theorems: interdependence and field of application.

For $\mathsf{T} \subset \mathbb{R}$, a Markov measure P on $(\mathbb{R}^d)^{\mathsf{T}}$ is a Markov martingale if for every $\{s,t\} \subset \mathsf{T}$ with $s<t$, the coupling $(\mathrm{proj}^{s,t})_\# P$ is.

(b) When $d=1$, *submartingale couplings* and *Markov submartingales* are defined alike, the integral in (4) being non-negative instead of null.

(c) A measure P on $\mathbb{R}^2$ is called an *increasing coupling* if

$$P(\{(x,y)\in\mathbb{R}^2 : x \leq y\}) = 1,$$

i.e., $P = \mathrm{Law}((X_i)_{i\in\{1,2\}})$ where $X_1 \leq X_2$. For $\mathsf{T} \subset \mathbb{R}$, we say that a measure P on $\mathbb{R}^{\mathsf{T}}$ has increasing couplings if all $(\mathrm{proj}^{s,t})_\# P$ have so.

Remark 1.16

(a) If a measure P on $\mathbb{R}$ is the law of a process with increasing trajectories, as in Theorem A(b), P is in case (c) of Definition 1.15. Actually a classical type of reasoning shows the converse, see Lemma 1.19 below. So the result in Corollary 1.13 amounts to giving the existence of a Markov measure $P \in \mathrm{Marg}((\mu_t)_t)$ with increasing couplings.

(b) Take $\mathsf{T} \subset \mathbb{R}$. If there exists $P \in \mathrm{Marg}((\mu_t)_{t\in\mathsf{T}})$ as defined in case (a), (b), or (c) of Definition 1.15, then it is immediate that $(\mu_t)_t$ is respectively increasing for $_{C}$, $_{C,\mathrm{sto}}$ or $_{\mathrm{sto}}$. When $\sharp\mathsf{T} = 2$, by Strassen's theory [44], the converse is true. More generally, it is also true when $\mathsf{T} = \mathbb{N}$; one can deduce it by a quite simple induction based on the Markov concatenation (Definition 2.8).

The converse part of Remark 1.16(b) is a much more delicate question when $\mathsf{T} = \mathbb{R}$ than when T is $\{0,1\}$ or $\mathbb{N}$. Kellerer answered it for $_C$ and $_{C,\mathrm{sto}}$, as follows.

Theorem 1.17 (Kellerer, [28, Th. 3], [29]). — *If $(\mu_t)_{t\in\mathbb{R}}$ is a family of probability measures on* $\mathbb{R}$*, increasing for* $\preceq_C$ *(or* $\preceq_{C,\mathrm{sto}}$*), there is a Markov measure $P \in \mathrm{Marg}((\mu_t)_t)$ which is a (sub)martingale.*

Kellerer's Theorem remains unproved for vectorial measures, see Open question 5.5.2. This is a major motivation to search new methods to construct or to establish the existence of certain Markov processes, as it is done to prove Theorems A and C. Following Kellerer's line of proof, but replacing one of its key lemmas by another one (see details below), we prove its following generalization.

Theorem C. — *If $(\mu_t)_{t\in\mathbb{R}}$ is a family of probability measures on* $\mathbb{R}$*, increasing for* $\preceq_C$*,* $\preceq_{C,\mathrm{sto}}$ *or* $\preceq_{\mathrm{sto}}$*, there is a Markov measure $P \in \mathrm{Marg}((\mu_t)_t)$ which respectively is a martingale, is a submartingale or has increasing couplings.*

Using Remark 1.16(a), one sees that this theorem proves Corollary 1.13, that follows from Theorem A, by another way.

Kellerer's proof uses a continuity result for certain kernels, recalled in Lemma 2.19. We replace it by Lemma 2.23, a continuity result for the increasing kernels of Definition 1.6. A form of Lemma 2.19 appears in every proof of Kellerer's theorem we know [28, 29, 35, 20, 3], so that in this respect our proof, resting on another type of kernels, is new. About Lemma 2.23, the following comments, that also appear in Figure 1, are in order.

– It is a significant result of this article; Section 3 is devoted to its proof.

– It is not used in the proof of Theorem A, so that the proofs of our Theorems A and C are really independent. They bring separately Corollary 1.13, an enhancement Theorem C in the (new with respect to Kellerer's work) case of $\preceq_{\mathrm{sto}}$.

– It plays a prominent role in Theorem B, see p. 51, and as a consequence of it, in the results concerning Optimal Transport and the Continuity Equation established in the companion paper [8].

Remark 1.18. — The Doob–Meyer decomposition theorem of some submartingales in a sum of an increasing process and a martingale is another reason our generalization of Theorem 1.17 to $\preceq_{\mathrm{sto}}$ is a natural work.

We also mention that Kellerer seems to have never considered the question of the extension to $\preceq_{\mathrm{sto}}$ in his papers. However in [30] with application in [31], he explored the related question of the existence of increasing couplings $P \in \mathrm{Marg}(\mu,\nu)$ that are as independent as possible, in a suitable sense.

Finally, here is the lemma announced in Remark 1.16(a), yielding Corollary 1.13 from Theorem C. It is proved p. 30 in Section 3.4.

Lemma 1.19. — *If P is a probability measure on $\mathbb{R}^{\mathbb{R}}$ with increasing couplings (see Definition 1.15(c)), there exists an increasing process $X = (X_t)_{t\in\mathbb{R}} : \Omega \to \mathbb{R}$, i.e., the functions $t \mapsto X_t(\omega)$ are increasing, such that $P = \mathrm{Law}(X)$.*

1.4. Application to the continuity equation and its treatment in Optimal Transport Theory. — A striking commutation between the notions of curves (continuous functions defined on intervals) on the space of probability measures and the probability measures on the space of curves is showing up in Optimal Transport Theory. In a parallel paper [8], as an application of the present paper, we show how the Markov-quantile can enhance this connexion in the direction of Stochastic Processes, see Figure 1. In this paragraph we explain this briefly. To show how MQ enters this field, we need first a few words of context. All the rigorous definitions are given in [8]. It can be deduced from Ambrosio, Gigli and Savaré's major contribution [2] that the Lagrangian formulation of Benamou–Brenier formula [6] is as follows. For any two given measures μ_0, μ_1 (with finite second moment) over a metric space (X, d) there exists a process $X = (X_t)_{t\in[0,1]}$ for which: (i) $X_0 \sim \mu_0$ and $X_1 \sim \mu_1$; (ii) the infimum value of $\mathbb{E}(\mathsf{E}(X))$ is attained, and equals $W_2(\mu_0, \mu_1)^2$ the square of the famous Wasserstein distance, also known as transport distance. The (random) quantity $\mathsf{E}(X)$ is called the kinetic energy of the (random) curve $X = (X_t)_{t\in[0,1]}$. Informally, it is "$\int_0^1 |\dot{X}_t|^2\, dt$" and makes sense for absolutely continuous curves of order two on a metric space. In particular, for a curve $\mu = (\mu_t)_{t\in[0,1]}$ interpolating $\{\mu_0, \mu_1\}$, the energy $\mathsf{E}(\mu)$ is computed in the Wasserstein-space $(\mathcal{P}_2(\mathsf{X}), W_2)$. One has $\mathbb{E}(\mathsf{E}(X)) \geqslant \mathsf{E}(\mu)$ as soon as $X_t \sim \mu_t$ for every $t \in [0,1]$. Lisini proved [33] that this lower bound $\mathsf{E}(\mu)$ is attained by some X.

We prove in the Main Theorem of [8] that, for $\mathsf{X} = \mathbb{R}$, the Markov-quantile process is such a minimizer, and hence that the minimizer can be made Markov. After adaptation, Theorem B provides moreover a uniqueness statement that we would find extremely interesting to extend to the Wasserstein space over any (Polish) metric space (the open problem is formally stated in [8]). For a taste of it, in dimension 1, see the examples in Section 5.2 and the related open problem in Section 5.5.4).

1.5. A first insight in the Markov-quantile process. — We build the Markov-quantile process $\mathrm{MQ}((\mu_t)_{t\in\mathbb{R}})$ answering the Main Problem in elementary examples of increasing difficulty, where what it shall be is clear. This makes the generalization of this construction, i.e., solving the Main Problem, a natural goal. Trying to achieve it by a naive strategy will then reveal its difficulties.

Notation 1.20. — In Section 1.5 we denote the Lebesgue measure on $[0,1]$ by λ, on $\mathbb{R}$ by $\mathrm{d}x$, $\lambda_{\mathbb{R}}$ or also λ when there is no ambiguity, on $\mathbb{R}^d$ by $\lambda_{\mathbb{R}^d}$.

Example 1.21. — Define $\mu = (\mu_t)_{t\in\mathbb{R}}$ by:

$$\mu_t = \delta_0 \text{ if } t = 0 \quad \text{and:} \quad \mu_t = \frac{1}{|t|}\mathbb{1}_{[0,|t|]}\mathrm{d}x \text{ if } t \in \mathbb{R} \smallsetminus \{0\}.$$

The quantile trajectory $(X_t(\alpha))_{t\in\mathbb{R}}$ associated with the level $\alpha \in [0,1]$ on the probability space $\Omega = ([0,1], \lambda)$ is $t \in \mathbb{R} \mapsto \alpha|t|$. The process X is not Markov because at time $t = 0$, with information on $(X_t)_{t<0}$, we better know $(X_t)_{t>0}$—actually here, we determine it completely. A modification makes X Markov. Namely, consider the concatenation $X^{[0]}$ at time 0 of $t \in \mathbb{R}^- \mapsto \alpha(-t)$ and $t \in \mathbb{R}^+ \mapsto \beta t$, where (α, β) is uniformly picked in $[0,1]^2$, in place of $\alpha=\beta$ picked in $[0,1]$. Then $X^{[0]}$ is the only possible answer to the Main Problem: it is Markov (hence $(X_t)_{t<0}$ and $(X_t)_{t>0}$ are independent) and equal to the quantile process where the latter is Markov. Moreover, it satisfies the properties of Theorem A, so it is the Markov-quantile process.

Definition 1.22. — A measure without atom is said to be diffuse. We say that t is an *atomic time* of a family $(\mu_t)_{t\in\mathbb{R}}$ of measures if μ_t is not diffuse.

Example 1.23. — Now take μ_0 any non-diffuse measure and, for $t\neq 0$, μ_t any diffuse measure, for instance $\mu_0 * \theta_t$, where θ_t stands for μ_t of Example 1.21, or $\mu_t=\lambda$. The quantile process is $(X_t)_t=(\min\{x\in\mathbb{R}:\mu_t(]-\infty,x])>\alpha\})_t$; restricted to $\mathbb{R}^-$ or $\mathbb{R}^+$ it is Markov. Pick β uniformly on $[0,1]$. On $([0,1]^2,\lambda\otimes\lambda)$, we consider the modified process $X^{[0]}$ defined by:

$$X_t^{[0]}(\alpha,\beta)=\begin{cases}X_t(\alpha) & \text{if } t\leqslant 0\\ X_t(\alpha^-+\beta(\alpha^+-\alpha^-)) & \text{it } t>0,\end{cases}$$

where $\alpha^+=\alpha^-=\alpha$ if $\mu_0(X_t(\alpha))=0$, and otherwise $]\alpha^-,\alpha^+[$ is the maximal open interval such that $X_0(\alpha')=X_0(\alpha)$ for all α' in it. Let us introduce $I_\alpha=\{\alpha\}$ in the first case and $I_\alpha=]\alpha^-,\alpha^+[$ in the second one. Conditionally on the fact that $X_0^{[0]}$ equals some atom $x_0=X_0(\widetilde{\alpha})$ of μ_0, the values of $X_t^{[0]}$ for $t>0$ are made independent of those of $X_t^{[0]}$ for $t<0$. Indeed, the vector $(\alpha,\alpha^-+\beta(\alpha^+-\alpha^-))$ is uniformly distributed on $I_{\widetilde{\alpha}}\times I_{\widetilde{\alpha}}$: this generalizes Example 1.21; $X^{[0]}$ is Markov and has law $\mathrm{MQ}\in\mathrm{Marg}(\mu)$.

Example 1.24. — When the set R of atomic times of $(\mu_t)_{t\in\mathbb{R}}$ is finite, one may repeat at each $r\in R$ the independence operation described above at time 0, to produce a Markov process. Moreover, one may check that it does not matter in which order, because these operations commute. The resulting process is indeed our Markov-quantile process. With the notation of the paper, its law is $\mathrm{Q}_{[R]}$. Similarly it is easy to imagine the Markov-quantile process when the atomic times form a locally finite set, like, e.g., $\mathbb{Z}$.

Remark 1.25. — See Section 5 for more examples of Markov-quantile processes.

The situation becomes complicated when the set A of atomic times is not locally finite or even worse, uncountable. Consider the following a priori reasonable approach. Let $(R_n)_{n\in\mathbb{N}}$ be a nested family of finite sets such that $R_\infty=\bigcup_n R_n$ is dense in A. We consider the sequence $(\mathrm{Q}_{[R_n]})_n$ and hope for a limit Q. Then we encounter three problems:

– By compactness of $\mathrm{Marg}((\mu_t)_t)$, $(\mathrm{Q}_{[R_n]})_n$ has an accumulation point (see similar reasonings in [19, 24, 26, 33, 46]), but it has no reason to be unique.

– If A is uncountable, this limit needs not to satisfy the Markov property (1) at times $s\in\mathrm{A}\smallsetminus R_\infty$, at which we did not perform the modification of Examples 1.21–1.23. A continuity assumption on $t\mapsto\mu_t$ could let hope to yield it (this was used, with other goals, for measures in stochastic or convex order, see, e.g., [3, 20]) but we do not make such an assumption. Also in the "space of quantile levels", the irregularity may be maximal: the set $\{(t,\alpha)\in\mathbb{R}\times[0,1] : X_t(\alpha)$ is an atom of $\mu_t\}$ needs not to be measurable.

– Anyway, limits of Markov processes are in general not Markov so here, property (1) is not ensured even at $s\in R_\infty$. To our knowledge, before the present paper this type of problem had principally one solution, based on Lipschitz kernels (see Lemma 2.19), first discovered by Kellerer [28, 35, 20, 5, 3]. However, see [37, Lem. 5.3] for a different statement.

Another consequence of this non stability of the Markov property is that it is also not possible to consider the sequence of quantile processes for mollified curves $\mu^{(n)} = (\mu_t * \theta_n)_t$, relying on the fact that all the measures $\mu_t^{(n)}$ are diffuse, so that each $\mathrm{Q}\in\mathrm{Marg}((\mu_t^{(n)})_{t\in\mathbb{R}})$ is Markov.

Remark 1.26

(a) In fact, the convergence in our Theorem A(iv), and hence in its enhancement Theorem B, rests on the *order* $\preceq_{\mathrm{lo}}$ introduced in Definition 3.4: for all s and $t>s$, the choice of the times r_i follows from that of a sequence in some set of measures, tending to the supremum of this set for $\preceq_{\mathrm{lo}}$, see Lemma 4.10, in particular its point (c).

(b) As the examples above suggest, for all s and $t>s$, the composition appearing in point (iv) of Theorem A needs only using couplings $\mathrm{Q}_{r_i,r_{i+1}}$ where the times r_i are atomic. Adding non-atomic times has no effect. Similarly, in Theorem B, that provides nested finite sets R_n such that $\mathrm{Q}_{[R_n]}\to\mathrm{MQ}$, $R=\bigcup_n R_n$ may avoid all non-atomic times.

Now it appears moreover, but only as a consequence of Theorem A once it is proved, that all the atomic times of $(\mu_t)_t$ do not play the same role:

– Some are "essential" (see Definition 4.25). All of them that lie in $]s,t[$ must eventually appear among the r_i in Theorem A(iv), and all of them must belong to R in Theorem B. This is possible as they turn out to be at most countable (Proposition 4.26).

– One may choose the r_i in Theorem A, or R in Theorem B, so that they avoid any fixed finite set of the other atomic times.

Therefore, the intersection of the sets R satisfying the convergence property of Theorem B is the set of the essential atomic times.

The existence, for any sequence $(\mu_t)_t$, of the set of its essential atomic times, at most countable even if the set $\{(t,\alpha)\in\mathbb{R}\times[0,1] : X_t(\alpha)$ is an atom of $\mu_t\}$ is not measurable, is in itself a significant result of this article. Perhaps does this notion admit generalizations when the set of parameters or the measurable space, both equal to $\mathbb{R}$ in this work, are more general spaces.

Remark 1.27. — It is also very important to notice that as soon as the set

$$\{(t,\alpha)\in\mathbb{R}\times[0,1] : X_t(\alpha) \text{ is an atom of } \mu_t\}$$

is regular enough (see the examples of Section 5 for clearly stated instances of this), the Markov-Quantile process may be explicitly computed, as it is done in the important

Example 5.1. More generally, certain properties of this set and of MQ are linked, see the whole of Section 5. Through its various examples, that section gives also an intuition of how MQ behaves.

1.6. Organization of the paper. — In Section 2 we introduce in Section 2.1 kernels and transport plans, their composition and concatenation, and the Markov property expressed in this language. We give in Section 2.2 the structure of Kellerer's work [28, 29], explain why it motivates our reasoning towards Theorem C, and prove the latter. However, we postpone the introduction of one auxiliary notion, and the proof of Lemma 2.24 and of the essential Lemma 2.23 to Section 3. In Section 2.3 we state and prove the "Markovinification" Theorem 2.26.

In Section 3 we introduce the auxiliary notions and results leading to the proofs of Lemmas 1.19, 2.23 and 2.24, then also used in Section 4, namely:

(a) the "lower orthant" and stochastic orders, related suprema and the notion of increasing kernel in Section 3.1,

(b) the quantile transport and the notion of minimal coupling in Section 3.2,

(c) two distances, ρ and $\widetilde{\rho}$, inducing the weak topology on spaces of transport plans $\mathrm{Marg}(\mu,\nu)$ in Section 3.3.

Lemma 1.19 is proved in Section 3.4 and Lemmas 2.23 and 2.24 are in Section 3.5.

Along the way, Section 2.1, Section 3.1, and Section 3.2 also give all the background to understand in detail the three properties (i)–(iii) characterizing MQ in Theorem A.

In Section 4 we prove Theorems A and B: in Section 4.1 we explain how the situation may be pushed forward to the space $[0,1]$ of "levels of quantiles", in Section 4.2 we prove Theorem A, i.e., build the Markov-quantile process, and in Section 4.3 we state and prove Theorem 4.21 which is a more precise and complete version of Theorem B. To do this we introduce the essential atomic times of $(\mu_t)_t$.

In Section 5 we exhibit the Markov-quantile process in a series of examples, state three last remarks about Theorem A, and give open questions.

Note. — When we introduce various tools, sometimes classical, we do it in a way and with remarks adapted to our context. The reader already knowing them may read quickly, taking notice of our few specific remarks, which are useful in the rest of the article.

1.7. Notation

(a) We gather in Table 1 the places where the notation we use widely is introduced, so that the reader can find them quickly if needed.

(b) In this article, $\mathbb{R}^*$ or $\mathbb{N}^*$ means $\mathbb{R}\setminus\{0\}$ or $\mathbb{N}\setminus\{0\}$. If $E'\subset E$, $\mathbb{1}_{E'}:E\to\{0,1\}$ is the indicator function of E' and, if $\mu\in\mathcal{M}(E)$, $\mu|_{E'}$ stands for $\mathbb{1}_{E'}\mu$. The Dirac measure at $x\in E$ is denoted by δ_x and λ stands for the Lebesgue measure on $[0,1]$, $\mathbb{R}$ or $\mathbb{R}^d$. Most of the time we deal with $\lambda|_{[0,1]}$ so, when there is no ambiguity, we simply write λ in this case.

We introduce:	in:
$\mathrm{M}(E)$, $\mathrm{P}(E)$, $\mathrm{proj}^{\mathcal{T}'}$, P^t (similarly ^{t}), $P^{s,t}$, the set $\mathrm{Marg}((\mu_\tau)_{\tau\in\mathcal{T}})$	Notation 1.2
the quantile process Q	Reminder 1.4
Q again, together with the quantile coupling $\mathrm{Q}(\mu,\nu)$	Definitions 3.18 and 3.20, Notation 3.21
the stochastic order $_{\mathrm{sto}}$	Reminder 1.5
M Q	Definition 4.16(b)
if P is some process, R R and $\sharp R < \infty$, $P_{[R]}$	Notation 1.10, Definition 4.18
$_C$ and $_{C,\mathrm{sto}}$	Definition 1.14
λ	Notation 1.20
the composition $k \,.\, k'$ of kernels	Section 2.1.1
the kernels k_P	Notation 2.2
id_E	Notation 2.4
$\mathrm{Joint}(\mu,k)$	Notation 2.5
the transport plans $\mathrm{Id}_{2,\mu}$ and $\mathrm{Id}_{n,\mu}$	Notation 2.6
if P is some process, $^{\mathrm{t}}P$	Definition 2.7
P P'	Definition 2.8
$\mathrm{N}^{\mathrm{LK}}_{s,t}$	Definition 2.16
$\mathrm{N}^{\mathrm{IK}}_{s,t}$	Notation 2.25,
$x \leq y$ when $(x,y) \in (\mathrm{R}^d)^2$	Notation 3.1
F_μ or $F[\mu]$, if μ is some measure	Definition 3.2
μ $_{\mathrm{lo}}$ ν	Definition 3.4
$\mathrm{losup}_\tau\, P_\tau$	Definition 3.5
$\mathrm{M}(\mu)$, $\mathrm{P}(\mu)$, $\mathrm{M}^{\&}(\mu)$ and $\mathrm{P}^{\&}(\mu)$	Notation 3.13
G_μ	Definition 3.23
d	Notation 3.1
the distance ρ	Notation 3.27
$\widetilde{\rho}$	Section 3.3.2
the kernels q_r, k_r and $^{\mathrm{t}}k_r$	Notation 4.2
$A_{r,x}$, A_r and ℓ_r	Notation 4.4
L_R, for R R	Notation 4.10
ℓ_R	Notation 4.11
Lev	Definition 4.19

Table 1. Places where our notation is introduced

If $f : E \to F$ is measurable and $\mu \in \mathrm{M}(E)$, $f_\#\mu \in \mathrm{M}(F)$ is defined by $f_\#\mu(B) = \mu(f^{-1}(B))$. Product measures are denoted by $\mu \otimes \nu$. If f and g are functions, $f \otimes g$ stands for $(x,y) \mapsto (f(x), g(y))$.

(c) Recall also Convention 1.7: introducing $\{r_1,\dots,r_m\}$ or m-tuples $(r_k)_{k=1}^m$ of real numbers, we mean implicitly that $r_1 < \cdots < r_m$.

Vocabulary 1.28

(a) Similarly, Table 2 gathers our common vocabulary.

(b) When transport plans P and Q are composed, we call "composition" the operation, and "product" its result, see 2.1.2.

(c) We need a name for curves $(\mu_t)_{t\in\mathrm{R}}$ with the property $\mu_s \leq \mu_t$ for any $s < t$. If we call them "non decreasing", the standard terminology when the order is total,

For:	See:
process and *marginal (law)*	Reminder 1.1
canonical process, coupling & *transport (plan)*	Reminder 1.3
martingale, submartingale & *increasing* couplings	Definition 1.15
atomic & *essential atomic times*	Definitions 1.22 and 4.25
increasing kernel	Definitions 1.6 and 3.11
Lipschitz kernel	Definition 2.16
quantile coupling & *quantile measure*	Definitions 3.18 and 3.20
atomic levels	Notation 4.4
a process "M *made Markov at the points of* R"	Definition 4.18

Table 2. Places where our common vocabulary is introduced

the readers may imagine that the relation we consider is the contrary of $\mu_s > \mu_t$ (for any $s < t$). But in a poset the contrary of $a > b$ is ($a \leq b$ or "a and b are not comparable"). To avoid ambiguity, we prefer to call it with a positive term, namely "increasing". Similarly, we use "non increasing" instead of "decreasing". This convention is chosen in many articles dealing with our topic, e.g., [43, 26, 34, 20]

Acknowledgements. — The authors wish to thank Jiří Černý, Martin Huesmann, Christian Léonard, Emmanuel Opshtein and Xiaolu Tan for bibliographic or editorial suggestions as well as Michel Émery and Erwan Hillion for discussions on examples related to this work.

2. An extension of a theorem of Kellerer

2.1. The Markov property, composition and concatenation of kernels and transport plans.

— Everywhere E, E', E'' etc. are topological spaces (or sometimes Polish spaces) and $\mathcal{B}(E)$, $\mathcal{B}(E')$ and $\mathcal{B}(E'')$ their Borel σ-algebras.

Definition 2.1. — A probability kernel, or kernel k from E to E' is a map

$$k : E \times \mathcal{B}(E') \longrightarrow [0,1]$$

such that $k(x,\cdot)$ is a probability measure on E' for every x in E and $k(\cdot, B)$ is a measurable map for every $B \in \mathcal{B}(E')$.

Probability kernels are usually interpreted as transition matrices, see Remark 2.3: after one step a particle at x in E arrives at a random position in E', distributed with respect to $k(x,\cdot)$. We often have that interpretation in mind.

Remark/Notation 2.2. — Every transport plan $P \in \mathcal{P}(E \times E')$ can be disintegrated with respect to its first marginal $P^1 := (\mathrm{proj}^1)_\# P$ and a kernel that we denote by k_P, defined from E to E', so that:

$$\iint f(x,y)\,\mathrm{d}P(x,y) = \int \left(\int f(x,y)\, k_P(x,\mathrm{d}y) \right) \mathrm{d}P^1(x)$$

for every bounded continuous function f. Observe that $x \mapsto k(x,\cdot)$ is P^1-almost surely uniquely determined.

2.1.1. *Composition and action of kernels*. — Kernels k from E to E' and k' from E' to E'' can be composed as follows:

$$(k\,.\,k')(x,A) = \int_{E'} k'(y,A)k(x,\mathrm{d}y).$$

Similarly, acting on the right, kernels from E to E' *transport*, or *send* positive measures θ on E on positive measures on E'. Acting on the left, they send (adequately integrable) functions $f : E' \to \mathbb{R}$, on functions $E \to \mathbb{R}$:

$$(\theta\,.\,k)(A) = \int_{E'} k(y,A)\theta(\mathrm{d}y) \quad \text{and:} \quad (k\,.\,f)(x) = \int_{E'} f(y)k(x,\mathrm{d}y).$$

Associativity holds, e.g., $(k\,.\,k')\,.\,k'' = k\,.\,(k'\,.\,k'')$, and $\theta\,.\,(k\,.\,f) = (\theta\,.\,k)\,.\,f$, where the action of measures on functions is the obvious one. This is consistent with the following remark.

Remark 2.3. — We recall the usual interpretation of the composition as matrix multiplication. If $E = \{x_i\}_{i=1}^n$, $E' = \{y_j\}_{j=1}^{n'}$, $E'' = \{z_k\}_{k=1}^{n''}$ are finite, a measure $\theta \in \mathrm{M}(E)$ is a row vector $(\theta(\{x_i\}))_{i=1}^n = (\theta_i)_{i=1}^n$, a kernel k from E to E' is a matrix $k = ((k_{i,j})_{i=1}^n)_{j=1}^{n'}$, where $(k_{i,j})_{j=1}^{n'}$ is the measure $k(x_i,\cdot) \in \mathrm{M}(E')$, viewed as a vector, and a function f from E'' to $\mathbb{R}$ is a (column) vector $f = (f(z_j))_{j=1}^{n''}$. Then, taking $k' = ((k'_{j,k})_{j=1}^{n'})_{k=1}^{n''}$ a kernel from E' to E'', $\theta\,.\,k$, $k\,.\,k'$ and $k'\,.\,f$ introduced above have the same sense as products of matrices.

Notation 2.4. — We denote by id_E the identity kernel (that acts trivially) $(x,B) \mapsto \delta_x(B) = \mathbb{1}_B(x)$.

Notation 2.5. — With $\mu \in \mathrm{M}(E)$ and k a kernel from E to E' is naturally associated the law $\mathrm{Joint}(\mu,k) \in \mathrm{M}(E \times E')$, having μ as first marginal and the family $(k(x_0,\cdot))_{x_0 \in E}$ as laws (on E') conditioned by $x_0 \in E$:

$$\forall B, B' \in \mathrm{B}(E) \times \mathrm{B}(E'), \quad \mathrm{Joint}(\mu,k)(B \times B') = \int_B k(x,B')\mathrm{d}\mu(x).$$

In particular, $P = \mathrm{Joint}(P^1, k_P)$.

2.1.2. *Composition of transport plans*. — If $P \in \mathrm{Marg}(\mu,\mu')$ and $Q \in \mathrm{Marg}(\mu',\mu'')$, we can compose them in a similar way as we compose kernels, getting the product:

$$P\,.\,Q := \mathrm{Joint}(\mu, k_P\,.\,k_Q) \in \mathrm{Marg}(\mu,\mu''), \quad \text{so that: } k_{P.Q} = k_P\,.\,k_Q.$$

Notation 2.6. — We denote $((\mathrm{Id}_E)_{i=1}^n)_\#\mu \in \mathrm{Marg}((\mu)_{i=1}^n)$ by $\mathrm{Id}_{n,\mu}$ or simply Id_n when there is no ambiguity. With $n = 2$, $\mathrm{Id}_{2,\mu} = \mathrm{Joint}(\mu, \mathrm{id}_E)$. It is moreover the identity transport: $\mathrm{Id}_{2,\mu}\,.\,P = P = P\,.\,\mathrm{Id}_{2,\mu'}$.

2.1.3. *Action of transport plans on measures and functions*. — If $\mu \in \mathrm{M}(E)$ and $\mu' \in \mathrm{M}(E')$, transport plans $P \in \mathrm{Marg}(\mu,\mu')$ have an action similar to that of kernels from E to E', on (μ'-almost surely defined) classes of functions f and on measures

$\theta \in \mathrm{M}(E;\mu)$ absolutely continuous with respect to μ. For instance, the latter are transported in $\mathrm{M}(E', \mu')$, as follows:

$$\text{if } \theta \ll \mu \text{ has density } g \text{ and } B' \in \mathrm{B}(E'),\quad (\theta . P)(B') = \int_{B'}\int_E g(x)\mathrm{d}P(x,y).$$

If k is a kernel from E to E', $\theta . k = \theta . \mathrm{Joint}(\mu, k)$. Conversely $\theta . P = \theta . k_P$.

Definition 2.7. — Take $P \in \mathrm{Marg}(\mu_1, \ldots, \mu_k)$, or $P \in \mathrm{Marg}((\mu_t)_{t\in\mathbb{R}})$. We define its transpose ${}^{\mathrm{t}}P \in \mathrm{Marg}(\mu_k, \ldots, \mu_1)$, resp. ${}^{\mathrm{t}}P \in \mathrm{Marg}((\mu_{-t})_{t\in\mathbb{R}})$ by: ${}^{\mathrm{t}}P(B_1 \times \cdots \times B_k) = P(B_k \times \cdots \times B_1)$, resp. ${}^{\mathrm{t}}P(\prod_i B_{t_i}) = P(\prod_i B_{-t_i})$.

The notation ${}^{\mathrm{t}}P$ is a reference to a transposed matrix in Remark 2.3. For $P = \mathrm{Marg}(\mu, \nu)$ we will often consider the bilinear map

$$B : (\theta, f) \longmapsto (\theta P)f = \theta(Pf) = \iint g(x)f(y)\, P(\mathrm{d}x, \mathrm{d}y),$$

where g is the density of θ with respect to μ. The case $\theta = \mu\llcorner_{]-\infty,x]}$, $f = \mathbb{1}_{]-\infty,y]}$ with $B(\theta,\nu) = P(]-\infty,x] \times]-\infty,y])$ is of special interest, see Section 3.1.

2.1.4. *Concatenation of transport plans*. — (See, e.g., [28, p. 111], [46, p. 23].)

Definition 2.8. — If $\mu_i \in \mathrm{P}(E_i)$ for $i \in \{1,2,3\}$, if $P_{1,2} \in \mathrm{Marg}(\mu_1,\mu_2)$ and $P_{2,3} \in \mathrm{Marg}(\mu_2,\mu_3)$, their concatenation $P_{1,2} \circ P_{2,3}$ is the unique $R \in \mathrm{P}(\mathbb{R}^3)$ such that for every $(B_1,B_2,B_3) \in \mathrm{B}(E_1) \times \mathrm{B}(E_2) \times \mathrm{B}(E_3)$:

$$R(B_1 \times B_2 \times B_3) = \int_{x\in B_1}\int_{y\in B_2}\int_{z\in B_3} \mathrm{d}\mu_1(x)k_{1,2}(x,\mathrm{d}y)k_{2,3}(y,\mathrm{d}z). \tag{5}$$

In particular, $R \in \mathrm{Marg}((\mu_1,\mu_2,\mu_3)$, $(\mathrm{proj}^{1,2})_\# R = P_{1,2}$, and $(\mathrm{proj}^{2,3})_\# R = P_{2,3}$.

Remark 2.9. — Let $k_{i,j}$ be a disintegration kernel for $P_{i,j}$ and $P_{2,1} := {}^{\mathrm{t}}P_{1,2}$. The right side of (5) also reads: $\iint_{B_1\times B_2}\int_{B_3} k(y,\mathrm{d}z)\mathrm{d}P_{1,2}(x,y)$, hence also:

$$\int_{y\in B_2}\iint_{x\in B_1,\, z\in B_3} k_{2,1}(y,\mathrm{d}x)k_{2,3}(y,\mathrm{d}z)\mu_2(\mathrm{d}y). \tag{6}$$

Concatenation is "reversed when time is reversed" in the sense that ${}^{\mathrm{t}}P_{2,3} \circ {}^{\mathrm{t}}P_{1,2} = {}^{\mathrm{t}}(P_{1,2} \circ P_{2,3})$; this is immediate after (6). Formula (5) gives immediately that $\circ$ is associative, leading to its following generalization:

$$P_{1,2} \circ \cdots \circ P_{n-1,n}\Big(\prod_{i=1}^n B_i\Big) = \int_{(x_i)_i\in\prod_i B_i} \mathrm{d}\mu_1(x_1)k_{1,2}(x_1,\mathrm{d}x_2)\ldots k_{n-1,n}(x_{n-1},\mathrm{d}x_n).$$

Remark 2.10

(a) In Section 2.1.2, one can also define $P . Q$ as $\mathrm{proj}_\#^{E,E''}(P \circ Q)$.

(b) The composition and concatenation of transport plans find an easy interpretation in terms of random variables. If (X_1,X_2,X_3) is a random vector with $\mathrm{Law}(X_1,X_2) = P$ and $\mathrm{Law}(X_2,X_3) = Q$ such that $(X_i)_{i\in\{1,2,3\}}$ is a Markov process (X_1 and X_3 are independent conditionally on X_2), then $P . Q$ is the law of (X_1,X_3) and $P \circ Q$ the law of (X_1,X_2,X_3), see (6).

2.1.5. *The Markov Property*. — We introduce the Markov property here in an alternative, equivalent way to the usual one.

Remark/Definition 2.11. — As recalled in (1), a process $(X_t)_{t\in\mathbb{R}}$ is said to be *Markov* if:

$$\forall s\in\mathbb{R}, \forall t>s,\quad \mathrm{Law}(X_t|(X_u)_{u\leqslant s}) = \mathrm{Law}(X_t|X_s).$$

Denoting $\mathrm{Law}((X_t)_t)\in\mathcal{P}(\mathbb{R}^{\mathbb{R}})$ by P, this is equivalent to the fact that for all finite subset $S=\{s_1,\dots,s_d\}$ of $\mathbb{R}$, $(\mathrm{proj}^S)_\# P$ is the concatenation $P^{s_1,s_2}\circ\cdots\circ P^{s_{d-1},s_d}$, where $P^{s_i,s_{i+1}}$ denotes $(\mathrm{proj}^{\{s_i,s_{i+1}\}})_\# P$. More generally, we say that any measure $P\in\mathcal{P}(\mathbb{R}^{\mathbb{R}})$ is Markov if it satisfies this property.

We extend these definition in the obvious way to processes indexed on subsets $R\subset\mathbb{R}$ and measures on $\mathcal{P}(\mathbb{R}^R)$.

We recall the Kolmogorov–Daniell theorem and its usual corollary on Markov measures.

Proposition 2.12 (Kolmogorov–Daniell theorem). — *Let E be a Polish space and, for each finite subset S of $\mathbb{R}$, μ_S be a probability measure on E^S. If $(\mathrm{proj}^{S'})_\#\mu_S=\mu_{S'}$ for any $S'\subset S$, there exists a unique $P\in\mathcal{P}(E^{\mathbb{R}})$ with $(\mathrm{proj}^S)_\# P=\mu_S$ for all S.*

One of the most usual applications of Proposition 2.12 is for measures μ_S of type $\mu_{s_1,s_2}\circ\cdots\circ\mu_{s_{d-1},s_d}$, where $S=\{s_1,\dots,s_d\}$.

Corollary 2.13. — *Let $(\mu_{s,t})_{s<t}$ be a family of transport plans in $\mathcal{P}(E\times E)$ such that:*

$$\mu_{s,u}=\mu_{s,t}\cdot\mu_{t,u}$$

for every $s<t<u$. Then there exists a unique Markov measure $P\in\mathcal{P}(E^{\mathbb{R}})$ with $P^{s,t}=\mu_{s,t}$ for every $s<t$.

Definition 2.14. — It is usual to call *consistent family* every family $(\mu_S)_S$ or $(\mu_{s,t})_{s<t}$ as in Proposition 2.12 and Corollary 2.13.

2.2. Kellerer's work. Our motivation and proof of Theorem C. — In [28] and [29], Kellerer proves the three results that we reproduce as Theorem 2.15, Lemma 2.19 and finally Theorem 2.21, which is a more precise version of Theorem 1.17 given in the introduction. He also introduces Definition 2.16. As we will see Theorem 2.15 extends Corollary 2.13: take $\mathsf{N}_{s,t}=\{\mu_{s,t}\}$.

Our goal here is to prove Theorem C. To put forward quickly both the background and our reasoning we postpone all the intermediate proofs, as well as the introduction of the technical tools they require to the next section.

Kellerer first proves the following statement—we give the sketch of proof on p. 21. It seems a bit stronger than in [28] but is what he actually shows.

Theorem 2.15 ([28, Th. 1]). — *Let $(\mu_t)_{t\in\mathbb{R}}$ be a family of probability measures on some Polish space E, and for every $s<t$ let $\mathsf{N}_{s,t}\subset\mathcal{P}(E^2)$ be a set of transport plans. Assume that:*

(i) *for every s,t, $\mathsf{N}_{s,t}$ is not empty,*

(ii) *for every* s,t, $\mathsf{N}_{s,t} \subset \mathrm{Marg}(\mu_s,\mu_t)$,

(iii) *for every* s,t, $\mathsf{N}_{s,t}$ *is closed for the weak topology,*

(iv) *for* $r<s<t$ *and any* $(P,P') \in \mathsf{N}_{r,s} \times \mathsf{N}_{s,t}$, $P.P' \in \mathsf{N}_{r,t}$,

(v) *for every* d *and* $t_1 < \cdots < t_d$, *if the sequences* $(Q^n_{t_i,t_{i+1}})_n \in \mathsf{N}_{t_i,t_{i+1}}$ *converge weakly to* $Q_{t_i,t_{i+1}}$, *then the sequence* $(Q^n_{t_1,t_2} \circ \cdots \circ Q^n_{t_{d-1},t_d})_n$ *tends weakly to* $Q_{t_1,t_2} \circ \cdots \circ Q_{t_{d-1},t_d}$.

Then, there exists a Markov measure $P \in \mathrm{Marg}((\mu_t)_t)$ *with* $(\mathrm{proj}^{s,t})_\# P \in \mathsf{N}_{s,t}$ *for every* $s<t$.

Definition 2.16 ([28, Def. 3]). — Let (E,μ) and (E',μ') be two measure metric spaces and P be in $\mathrm{Marg}(\mu,\mu')$. Then P *has Lipschitz kernel* if for every 1-Lipschitz map $h: E' \to [0,1]$, $P.h: E \to [0,1]$ is also 1-Lipschitz, i.e., more exactly, there is a 1-Lipschitz $\widetilde{h}: E \to [0,1]$ such that $\widetilde{h} = k_P . h$, μ-almost surely (see Notation 2.2 for k_P). For $(\mu_t)_{t\in\mathbb{R}} \in \mathcal{P}(\mathbb{R})^{\mathbb{R}}$, we denote $\{P \in \mathrm{Marg}(\mu_s,\mu_t) : P$ has Lipschitz kernel$\}$ by $\mathsf{N}^{\mathsf{LK}}_{s,t}$.

Remark 2.17 gives some comments, Remark 2.18 is used in the following.

Remark 2.17

(a) The terminology "Lipschitz property" was introduced in [35, Def. 4.1] for a Markov process with Lipschitz transition kernels. It is renamed as "Lipschitz-Markov property" by Hirsch, Roynette and Yor in [20]. The fact that this property is stable for finite dimensional convergence of processes is crucial in these papers and in [34] and appears as an avatar of Kellerer's Lemma 2.19 stating that the concatenation operator $\circ$ is continuous for the corresponding class of kernels. These kernels are called Lipschitz in [23, 5] and the present paper, and Lipschitz-Markov in [3].

(b) You may compare Definition 2.16 with that of transport plans with increasing kernel in Definition 3.11(b).

(c) [28, p. 115] In case the topology of E and E' is discrete, hence induced, e.g., by the distance $d(x,y) = 1 - \delta_{x,y}$, every $\widetilde{h}$ is 1-Lipschitz; hence any P has Lipschitz kernel.

Remark 2.18

(a) If some family $(\mathsf{N}_{s,t})_{s,t}$ satisfies the properties of Theorem 2.15, a family of subsets $(\mathsf{N}'_{s,t})_{s,t}$ with $\mathsf{N}'_{s,t} \subset \mathsf{N}_{s,t}$ satisfies them as soon as it satisfies (i), (iii) and (iv), (ii) and (v) being automatically true.

(b) For $(\mu_t)_t$ any family of measures on $\mathbb{R}$, it is easy to check that $(\mathsf{N}^{\mathsf{LK}}_{s,t})_{s<t}$ satisfies (i)–(iv) in Theorem 2.15 (for (iii), see [28, Satz 13]).

Lemma 2.19 (Continuity of $\circ$ when the kernels are Lipschitz [28, Sätze 14 & 15])

If (E_t,μ_t) *are complete and separable measure metric spaces,* $(\mathsf{N}^{\mathsf{LK}}_{s,t})_{s<t}$ *satisfies* (v) *in Theorem* 2.15.

Remark 2.20. — In fact [28, Sätze 14 & 15] proves (v) for sequences $(Q^n_{t_i,t_{i+1}})_n$ of Markov-Lipschitz transports, without the assumption that the $Q^n_{t_i,t_{i+1}}$ have the same marginals for all n, though this stronger result is not used further in [28]. In our Lemma 2.23 this assumption is crucial.

Finally Kellerer proves this more precise version of Theorem 1.17.

Theorem 2.21 ([28, Th. 3], [29]). — *If $(\mu_t)_t$ is an increasing family of measures on* R, *for* $_{C}$ *(or* $_{C,\mathrm{sto}}$*), there is a Markov measure $P \in \mathrm{Marg}((\mu_t)_t)$ such that P is a (sub)martingale and the couplings $P^{s,t}$ have Lipschitz kernel.*

Remark 2.22. — To prove Theorem 2.21, by both points of Remark 2.18, and Lemma 2.19, Kellerer has only to show that

$$(\mathrm{N}'^{\mathrm{LK}}_{s,t})_{s<t} := (\{P \in \mathrm{N}^{\mathrm{LK}}_{s,t} : P \text{ is a (sub)martingale transport}\})_{s<t}$$

—see Definition 1.15—satisfies Assumptions (i), (iii) and (iv) of Theorem 2.15 and to apply this theorem. Checking (iii) and (iv) being easy, we see that the two important facts enabling to use Theorem 2.15 and thereby getting Theorem 2.21 are:

(i) Lemma 2.19,

(ii) the proof of Property (i), i.e., the non-emptiness of the $\mathrm{N}'^{\mathrm{LK}}_{s,t}$.

Replacing point (i) by an alternative version (i'), consisting of Lemma 2.23 below, and proving a version of (ii) adapted to this change, we prove Theorem C. Namely we prove that increasing kernels, introduced in Definition 1.6 (see also Definition 3.11 for more details), satisfy Lemma 2.23, a counterpart of Lemma 2.19, as well as Property (iii), i.e., the little Lemma 2.24. They are proved respectively on p. 34 and on p. 32. Then we prove Theorem C.

Lemma 2.23 (Continuity of when the kernels are increasing). — *Take $(\mu_1, \dots, \mu_n) \in \mathrm{P}(\mathrm{R})^n$ and for all $i \in [\![1, n-1]\!]$ a closed set $\mathrm{I}_{i,i+1} \subset \mathrm{Marg}(\mu_i, \mu_{i+1})$ of transport plans with increasing kernel. The sets $\mathrm{I}_{i,i+1}$ satisfy Property* (v) *stated for the sets $\mathrm{N}_{t_i,t_{i+1}}$ in Theorem* 2.15.

Lemma 2.24. — *Take μ and μ' in* $\mathrm{P}(\mathrm{R})$. *The space of transport plans with increasing kernel in* $\mathrm{Marg}(\mu, \mu')$ *is closed for the weak topology.*

Notation 2.25. — If $(\mu_t)_t \in \mathrm{P}(\mathrm{R})^{\mathrm{R}}$ is given, we denote

$$\{P \in \mathrm{Marg}(\mu_s, \mu_t) : P \text{ has increasing kernel}\}$$

by $\mathrm{N}^{\mathrm{IK}}_{s,t}$.

Proof of Theorem C. — Take $(\mu_t)_t \in \mathrm{P}(\mathrm{R})^{\mathrm{R}}$, increasing for $_{C}$ (case (a)), $_{C,\mathrm{sto}}$ (case (b)) or for $_{\mathrm{sto}}$ (case (c)) to prove the corresponding cases of Theorem C. In the sketch of proof of Theorem 2.21 given in Remark 2.22, replace $\mathrm{N}^{\mathrm{LK}}_{s,t}$ by $\mathrm{N}^{\mathrm{IK}}_{s,t}$ and introduce, similarly as defined in Remark 2.22 for cases (a) and (b), the spaces $\mathrm{N}'^{\mathrm{IK}}_{s,t}$—equal to $\{P \in \mathrm{N}^{\mathrm{IK}}_{s,t} : P(\{(x,y) \in \mathrm{R}^2 : x \leq y\}) = 1\}$ in case (c).

Properties (ii)–(v) of Theorem 2.15 are satisfied by $(\mathrm{N}^{\mathrm{IK}}_{s,t})_{s<t}$: (ii) by definition, (iii) by Lemma 2.24, (iv) is straightforward and (v) by Lemma 2.23. By Remark 2.18(a) we are left with showing (i), (iii) and (iv) for $(\mathrm{N}'^{\mathrm{IK}}_{s,t})_{s<t}$. Plainly, the conditions

defining the $\mathsf{N}'^{\mathrm{IK}}_{s,t}$ as subspaces are closed and stable by composition, (iii) and (iv) follow. For (i), in our three cases:

(a) By [3, §3.1], if $P \in \mathrm{Marg}(\mu_s,\mu_t)$ is a martingale transport plan, then $P \in \mathsf{N}^{\mathrm{LK}}_{s,t} \Leftrightarrow P \in \mathsf{N}^{\mathrm{IK}}_{s,t}$. Therefore $\mathsf{N}'^{\mathrm{IK}}_{s,t} \neq \emptyset$ if and only if $\mathsf{N}'^{\mathrm{LK}}_{s,t} \neq \emptyset$, which Kellerer proved.

(b) The element of $\mathsf{N}'^{\mathrm{LK}}_{s,t}$ Kellerer built in [28, Def. 7 & Th. 2] is in $\mathsf{N}'^{\mathrm{IK}}_{s,t}$ when $\mu_s \preceq_C \mu_t$.

(c) As explained in Remark 3.24, $\mathrm{Q}(\mu_s,\mu_t) \in \mathsf{N}'^{\mathrm{IK}}_{s,t}$.

For the completeness of this exposition, we also provide the following.

Sketch of proof of Theorem 2.15, *in the manner of Kellerer*. — Take $S = \{s_1,\dots,s_d\}$ any finite subset of $\mathbb{R}$. We introduce

$$\mathsf{N}_{s_1,s_2}\circ\cdots\circ\mathsf{N}_{s_{d-1},s_d} := \{N_1\circ\cdots\circ N_{d-1} : \forall i, N_i \in \mathsf{N}_{t_i,t_{i+1}}\}$$

and

$$\mathsf{L}_S := (\mathrm{proj}^S)_\#^{-1}(\mathsf{N}_{s_1,s_2}\circ\cdots\circ\mathsf{N}_{s_{d-1},s_d}).$$

Since $\mathrm{proj}^S : \mathrm{Marg}((\mu_t)_t) \to \mathrm{Marg}((\mu_s)_{s\in S})$ is onto (since for any $\eta \in \mathrm{P}(E^S)$, $(\mathrm{proj}^S)_\#(\eta \otimes (\bigotimes_{s\in\mathbb{R}\smallsetminus S}\mu_s)) = \eta$), (i) and (ii) imply that $\mathsf{L}_S \neq \emptyset$; by Properties (iii) and (v), $\mathsf{N}_{s_1,s_2}\circ\cdots\circ\mathsf{N}_{s_{d-1},s_d}$ is weakly closed in $\mathrm{P}(E^S)$ hence so is L_S.

By [28, §1.2] or [3, §2], $\mathrm{Marg}((\mu_t)_{t\in\mathbb{R}})$ is weakly compact, so that $\mathsf{L}_{\mathbb{R}} := \bigcap_{S\subset\mathbb{R}\text{ finite}}\mathsf{L}_S \neq \emptyset$. Indeed, else, $\mathrm{Marg}((\mu_t)_{t\in\mathbb{R}})$ would be covered by the union of open sets $\bigcup_S(\mathbb{R}\smallsetminus\mathsf{L}_S)$, so that $\mathsf{L}_{S_1}\cap\cdots\cap\mathsf{L}_{S_N} = \emptyset$ for some N-tuple $(S_i)_{i\leq N}$ of finite sets. But by (iv), $S \subset S' \Rightarrow \mathsf{L}_{S'} \subset \mathsf{L}_S$, hence $\mathsf{L}_{S_1}\cap\cdots\cap\mathsf{L}_{S_N} \supset \mathsf{L}_{S_1\cup\cdots\cup S_N} \neq \emptyset$, a contradiction.

Finally take some $P \in \mathsf{L}_{\mathbb{R}}$. For every finite $S = \{s_1,\dots,s_d\}$, $(\mathrm{proj}^S)_\# P \in \mathsf{N}_{s_1,s_2}\circ\cdots\circ\mathsf{N}_{s_{d-1},s_d}$ hence P is a Markov measure.

2.3. Relation to the Markov-quantile process.

— What precedes provides also, through the application of Theorem 2.15, the following existence theorem for Markov processes being limits of products of transport plans taken out of a given process. When applied to the quantile measure $\mathrm{Q} \in \mathrm{Marg}((\mu_t)_t)$ introduced in Section 3.2, it provides the existence part of Theorem A, see below.

Theorem 2.26 (Markovinification). — *Let P be a measure of $\mathrm{P}(\mathbb{R}^{\mathbb{R}})$ with marginals $(\mu_t)_{t\in\mathbb{R}}$. If for each s and $t > s$, $P^{s,t}$ has increasing kernel, there exists a Markov measure P' in $\mathrm{Marg}((\mu_t)_t)$ such that each $P'^{s,t}$ is a limit of products $P^{s,r_1}\cdot\ldots\cdot P^{r_m,t}$ with $\{r_1,\dots,r_m\} \subset\,]s,t[$. One may take P' such that for each (s,t), the limit is obtained with a sequence $(\{r_1^n,\dots,r_{m(n)}^n\})_n$ such that $\max_{k=0}^{m(n)}|r_{k+1}-r_k| \to_{n\to\infty} 0$, where (r_0,r_{m+1}) stands for (s,t).*

Proof. — If $t > s$, setting $P^{s,t}_{[R]} := P^{s,r_1}\cdot\ldots\cdot P^{r_m,t}$ for all $R = \{r_1,\dots,r_m\} \subset\,]s,t[$, we introduce

$$\mathsf{N}^{(P)}_{s,t} = \bigcap_{\sigma>0}\overline{\{P^{s,t}_{[R]} : \max_{k=0}^{m}|r_{k+1}-r_k| \leq \sigma\}},$$

where (r_0,r_{m+1}) stands for (s,t). It is included in $\mathsf{N}^{\mathrm{IK}}_{s,t}$ by Lemma 2.24, and it satisfies Assumptions (i), (iii) and (iv) of Theorem 2.15. Indeed, for (i), $\mathsf{N}^{(P)}_{s,t} \neq \emptyset$ as an intersection of nested non-empty compact (closed in the compact space $\mathrm{Marg}(\mu_s,\mu_t)$) sets; (iii) is true by definition, and (iv) by Proposition 3.38. Thus by Remark 2.18(a), $\mathrm{N}^{(\mathrm{P})}_{s,t}$ satisfies all the assumptions of Theorem 2.15. We are done. (Notice that the alternative definition $\mathrm{N}^{(\mathrm{P})}_{s,t} = \overline{\{P^{s,t}_{[R]}\}}$ would have given the same result, except its last sentence.)

Note that if P is Markov the spaces $\mathrm{N}^{(\mathrm{P})}_{s,t}$ and $\overline{\{P^{s,t}_{[R]}\}}$ are both reduced to $\{P^{s,t}\}$, so that the Markov measure obtained from any of them is P itself. This conservation property also holds locally on intervals $I\subset\mathbb{R}$ if $(P_t)_{t\in I}$ is Markov. Notice also that Theorem 2.26 does not require $(\mu_t)_{t\in\mathbb{R}}$ to be increasing for $\preceq_{\mathrm{sto}}$.

Theorem 2.26 links Section 2 with the Markov-quantile process MQ built in Section 4. Indeed, taking $P=\mathrm{Q}$, $\mathrm{Q}^{s,t}$ is in $\mathrm{N}^{\mathrm{IK}}_{s,t}$ for all $s<t$ by Remark 3.25, so Theorem 2.26 gives the existence of a Markov process with 2-marginals in $\mathrm{N}^{(\mathrm{Q})}_{s,t}$ (here equal to $\overline{\{\mathrm{Q}^{s,t}_{[R]}\}}$). We prove in Section 4, by completely different means, that:
– this process is unique,
– it may be built using the order $\preceq_{\mathrm{sto}}$ (see also Remark 1.26), instead of being obtained by a non-constructive compactness argument.

This is Theorem A. See also Open question 5.5.1.

3. Three auxiliary notions, and postponed proofs of three lemmas

The next section introduces the notions needed to prove the results of Section 4 below. They are also necessary for the proofs of three lemmas that were therefore postponed: Lemma 1.19 on versions of increasing processes, the important Lemma 2.23 on the continuity of when the kernels are increasing, and Lemma 2.24.

3.1. Lower orthant and stochastic orders, related suprema, and increasing kernels

Notation 3.1

(a) Let us denote $(x_i)_{i=1}^d$ and $(y_i)_{i=1}^d$ in $\mathbb{R}^d$ by x and y. We endow $\mathbb{R}^d$ with the natural partial order defined by:

$$x\leqslant y \quad \text{if:}\quad \forall i, x_i\leqslant y_i.$$

We also set $[x,y] := \{z\in\mathbb{R}^d : x\leqslant z\leqslant y\} = \prod_i [x_i,y_i]$ and similarly $]x,y]$, etc. In particular $]-\infty,x] =]-\infty,x_1]\times\cdots\times]-\infty,x_d]$.

(b) Several times appear statements where some intervals have to be considered closed or open at some of their bounds, either arbitrarily or depending on possible cases. To alleviate the writing, we introduce the symbol "d" and place it at these bounds.

Definition 3.2. — If $\mu\in\mathcal{M}(\mathbb{R}^d)$, its cumulative distribution function F_μ, that we also denote by $F[\mu]$ to avoid multiple subscripts, is defined, using Notation 3.1, by:

$$F_\mu : x\in\mathbb{R}^d\longmapsto\mu(]-\infty,x]).$$

Reminder 3.3. — Recall for instance from [25, Th. 3.25] that such functions F are characterized by the fact that:

(a) for the natural partial order of $\mathbb{R}^d$ (see Notation 3.1), F is increasing and upper semi-continuous in the sense that for all $x \in \mathbb{R}^d$:

$$\forall \varepsilon > 0, \exists \eta > 0, \forall y \geqslant x, \quad \|y - x\|_\infty \leqslant \eta \implies F(x) \leqslant F(y) \leqslant F(x) + \varepsilon. \tag{7}$$

(b) $\lim_{\min_i(x_i) \to -\infty} F(x) = 0$ and $\lim_{\min_i(x_i) \to +\infty} F(x) = 1$,

(c) for every $h = (h_1, \ldots, h_d) \in [0, +\infty[^d$ and $x \in \mathbb{R}^d$, the quantity

$$\sum_\varepsilon \sigma(\varepsilon) F(x + \varepsilon h),$$

which is the measure of the rectangle $]x, x + h] \subset \mathbb{R}^d$, is non-negative. Here $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_d)$ ranges over $\{0, 1\}^d$, σ is 1 if $\sum \varepsilon_i$ is even, -1 otherwise, and εh means $(\varepsilon_1 h_1, \ldots, \varepsilon_d h_d)$.

Definition 3.4. — If $d \in \mathbb{N}^*$ and $m \in]0, +\infty[$, following [43, §6.G], we define the *lower orthant order* on $\{\mu \in \mathcal{M}(\mathbb{R}^d) : \mu(\mathbb{R}^d) = m\}$ by: $\mu \preceq_{\mathrm{lo}} \nu$ if $F_\mu \geqslant F_\nu$.

Definition 3.5. — We call lower orthant supremum of a family $(P_\tau)_{\tau \in \mathrm{T}}$ of measures of same mass m on $\mathbb{R}^d$, the smallest upper bound of $\{P_\tau\}_\tau$ for $\preceq_{\mathrm{lo}}$, if it exists, i.e., a measure P of mass m such that:

– for every τ, $P_\tau \preceq_{\mathrm{lo}} P$,
– $P \preceq_{\mathrm{lo}} Q$ as soon as $P_\tau \preceq_{\mathrm{lo}} Q$ for every τ.

By definition, if it exists it is unique. We denote it by $\operatorname{losup}_\tau P_\tau$. Similarly we define $\operatorname{loinf}_\tau P_\tau$.

Remark 3.6

(a) In Reminder 3.3, (a) and the first limit of (b) pass to the infimum of functions that are both monotone and upper semi-continuous. If moreover $(P_\tau)_\tau$ has an upper bound P, then the second limit of (b) holds. Indeed, for all τ, $F[P_\tau] \geqslant F[P]$, so that $\inf_\tau F[P_\tau] \geqslant F[P]$, and besides $\lim_{\min_i(x_i) \to +\infty} F[P](x) = 1$. Thus, if $(P_\tau)_\tau$ is bounded from above, $\inf_\tau F[P_\tau]$ satisfies (a)–(b) of Reminder 3.3, so is a cumulative distribution function if and only if it satisfies (c).

(b) If (c) of Reminder 3.3 is satisfied by the functions $F[P_\tau]$, then $(P_\tau)_\tau$ has a lower orthant supremum, and $F[\operatorname{losup}_\tau P_\tau] = \inf_\tau F[P_\tau]$.

Remark/Notation 3.7. — If $d = 1$ the order $\preceq_{\mathrm{lo}}$ is usually called stochastic order and denoted by $\preceq_{\mathrm{sto}}$; we will then call "stochastic supremum" the lower orthant supremum of Definition 3.5 and denote it by stosup.

Lemma 3.8 (Existence criteria for losup)

(a) *If, for $\preceq_{\mathrm{lo}}$, a sequence $(P_n)_{n \in \mathbb{N}} \in (\mathcal{P}(\mathbb{R}^d))^{\mathbb{N}}$ is bounded from above, and increasing, i.e., $n \leqslant m \Rightarrow P_n \preceq_{\mathrm{lo}} P_m$, then $\operatorname{losup}_n P_n$ exists and $(P_n)_n$ converges weakly to it.*

(b) *If a family $(P_\tau)_{\tau\in\mathrm{T}}$ is bounded from above for $\preceq_{\mathrm{lo}}$ and if for every $\tau,\tau'\in\mathrm{T}$ there exists $\sigma\in\mathrm{T}$ such that $P_\sigma \succeq_{\mathrm{lo}} P_\tau$ and $P_\sigma \succeq_{\mathrm{lo}} P_{\tau'}$, then* $\operatorname{losup}_\tau P_\tau$ *exists and there is an increasing sequence $(P_{\tau_n})_n$ that converges weakly to it.*

The results extend in an obvious way to measures of mass $m>0$ in $\mathcal{M}(\mathbb{R})$.

Proof

(a) Set $F := \inf_n F[P_n]$. After Remark 3.6, showing that F satisfies (c) of Reminder 3.3 ensures the existence of $\operatorname{losup}_n P_n$. Consider $M := \sum_\varepsilon \sigma(\varepsilon)F(x+\varepsilon h)$ as in Reminder 3.3(c). Since $F[P_n]$ is decreasing, F is its simple limit, so M is the simple limit of $\sum_\varepsilon \sigma(\varepsilon)F[P_n](x+\varepsilon h)$, hence $M\geq 0$. We are done. The weak convergence is given by the pointwise convergence of the cumulative distribution functions, see Reminder 3.26.

(b) This is a diagonal construction. Let $\mathrm{C}=\bigcup_{k\in\mathbb{N}}\{x_k\}$ be a countable dense set in $\mathbb{R}^d$. Set $F := \inf_{\tau\in\mathrm{T}} F_\tau$. Then for every $(k,n)\in\mathbb{N}^2$ we find $\tau_{k,n}\in\mathrm{T}$ such that $F[P_{\tau_{k,n}}](x_k)\leq F(x_k)+1/n$. For each n, by a finite induction using the assumption of (b) on the $P_{\tau_{k,n}}$, we find $\sigma_n\in\mathrm{T}$ such that: $\forall k\leq n,\ F[P_{\sigma_n}](x_k)\leq F(x_k)+1/n$ and $(P_{\sigma_n})_n$ is increasing. Hence:

$$\forall x\in\mathrm{C},\quad F[P_{\sigma_n}](x) \longrightarrow F(x). \tag{8}$$

By (a), $P=\operatorname{losup}_n P_{\sigma_n}$ exists and:

$$\forall x\in\mathbb{R}^d,\quad F[P_{\sigma_n}](x) \longrightarrow F[P](x). \tag{9}$$

Let us prove that (8) holds for any $x\in\mathbb{R}^d$, so that $F=F[P]$. Assume by contradiction that, for some x, $F(x)<\inf_n F[P_{\sigma_n}](x)$, i.e., by definition of F, that for some $\tau\in\mathrm{T}$, $F(x)<F[P_\tau](x)<\inf_n F[P_{\sigma_n}](x)=F[P](x)$. Now $F[P_\tau]$ is upper semi-continuous so in a neighbourhood U of x in $[x,+\infty[$, $F[P_\tau]<F[P]$. But on the dense set C, by (8) and (9), $F[P]=F$, hence on $\mathrm{U}\cap\mathrm{C}\neq\varnothing$, $F[P_\tau]<F$, a contradiction.

Finally $F=F[P]$, so F is a cumulative distribution function, so by Remark 3.6(b), $P=\operatorname{losup}_{\tau\in\mathrm{T}} P_\tau$. Moreover, $P_{\sigma_n}\to P$.

Remark 3.9

(a) (Case $d=1$) In this case, Reminder 3.3(c) is automatically true. Hence, in Lemma 3.8, (a) is true for any bounded $(P_n)_n$, increasing or not, hence (b) shows that any $\mathrm{S}\subset\mathcal{P}(I)$ bounded from above has a stochastic supremum (which has though not to be the weak limit of a sequence of elements of S, consider

$$\tfrac12(\delta_1+\delta_2)=\operatorname{stosup}\bigl\{\tfrac12(\delta_0+\delta_2),\delta_1\bigr\}.$$

Symmetrically, a family bounded from below has a stochastic infimum.

(b) Point (a) is false for $d>1$. Consider, e.g.,

$$\mathrm{S}=\{P_1,P_2\}=\bigl\{\tfrac12(\delta_{(1,0)}+\delta_{(0,1)}),\tfrac12(\delta_{(0,0)}+\delta_{(2,2)})\bigr\}\subset\mathcal{P}(\mathbb{R}^2),$$

then $\inf\{F[P_1],F[P_2]\}$ does not satisfy (c) of Reminder 3.3 and S has no lower orthant supremum: both $P:=\frac12(\delta_{(1,1)}+\delta_{(2,2)})$ and $P':=\frac12(\delta_{(0,2)}+\delta_{(2,0)})$ are upper bounds for it but no upper bound P'' satisfies $P''\preceq_{\mathrm{lo}} P$ and $P''\preceq_{\mathrm{lo}} P'$ (observe $F[P]$ and $F[P']$).

Remark 3.10. — In the following we use several times the Lebesgue differentiation theorem for Borel measures; a reference is, e.g., [12, §2.8–2.9].

Proposition/Definition 3.11. — *Take μ and ν in* $\mathrm{P}(\mathbb{R})$*. We say that a transport plan $P \in \mathrm{Marg}(\mu,\nu)$ has* increasing kernel *if one (and then any) of the following statements holds:*

(a) *Initial definition: if $\theta,\theta' \leqslant \mu$ and θ and θ' have he same mass, then $\theta \leqslant_{\mathrm{sto}} \theta'$ implies $\theta . P \leqslant_{\mathrm{sto}} \theta' . P$.*

(b) *For every increasing $h : \mathbb{R} \to [0,1]$, $P.h$ is μ-almost surely increasing, i.e., more exactly, there is an increasing $\widetilde{h} : \mathbb{R} \to [0,1]$ such that, for every bounded continuous function g:*

$$\int g(x)h(y)\mathrm{d}P(x,y) = \int g(x)\widetilde{h}(x)\mathrm{d}\mu(x).$$

(c) *There exists a kernel k in the μ-equivalence class of k_P such that $x \mapsto k(x,\cdot)$ is increasing from* $(\mathbb{R},\leqslant)$ *to* $(\mathrm{P}(\mathbb{R}),\leqslant_{\mathrm{sto}})$*.*

(d) *There exists a random vector (X,Y) with* $\mathrm{Law}(X,Y) = P$ *such that $x \in \mathbb{R} \mapsto$* $\mathrm{Law}(Y|X=x)$ *is increasing from* $(\mathbb{R},\leqslant)$ *to* $(\mathrm{P}(\mathbb{R}),\leqslant_{\mathrm{sto}})$ *(in the sense of the μ-equivalence classes of* increasing *functions: it is increasing on a set of full measure $F \subset \mathbb{R}$).*

Remark 3.12. — Be cautious that having increasing kernel is distinct from being an increasing coupling, a notion defined in Definition 1.15(c).

Proof of the equivalence in Proposition 3.11. — Statements (c) and (d) are essentially a change of notation. (To get (d) $\Rightarrow$ (c), notice that for y in the μ-null set $\mathbb{R} \smallsetminus E$ of (d), $k(y,\cdot)$ can be defined as $\mathrm{stosup}_{x\in E,\, x<y}\, k(x,\cdot)$; (c) $\Rightarrow$ (b) follows from the definition of $P.h$.) Let us show (b) $\Rightarrow$ (a) $\Rightarrow$ (c).

(b) $\Rightarrow$ (a): if $\theta \leqslant_{\mathrm{sto}} \theta'$, take $h : \mathbb{R} \to [0,1]$ increasing, then $(\theta . P) . h = \theta . \widetilde{h} \leqslant \theta' . \widetilde{h}$ since $\theta \leqslant_{\mathrm{sto}} \theta'$ and $\widetilde{h}$ is increasing by (b). Then $\theta' . \widetilde{h} = (\theta' . P) . h$ yields (a).

(a) $\Rightarrow$ (c): Suppose (a) and set $I_q := \,]-\infty, q]$ for all q. We will build $R \subset \mathbb{R}$, with $\mu(R) = 1$, on which $x \leqslant x' \Rightarrow k(x,I_q) \geqslant k(x',I_q)$ for all $q \in \mathbb{Q}$, hence all $q \in \mathbb{R}$, ensuring (c). By definition of the kernel k_P, $k_P(\,.\,,I_q)$ is the density with respect to μ of the measure $B \mapsto P(B \times I_q)$, thus by the Lebesgue differentiation theorem, setting $r(x,\varepsilon) := P([x-\varepsilon,x+\varepsilon] \times I_q)/\mu([x-\varepsilon,x+\varepsilon])$, the function $x \mapsto \lim_{\varepsilon\to 0} r(x,\varepsilon)$ is μ-almost everywhere defined and equal to it. Now

$$r(x,\varepsilon) = \Big(\Big(\frac{1}{\mu([x-\varepsilon,x+\varepsilon])}\mu|_{[x-\varepsilon,x+\varepsilon]}\Big) . P\Big)(I_q),$$

thus by (a), $x \leqslant x' \Rightarrow r(x,\varepsilon) \geqslant r(x',\varepsilon)$. Hence for all $q \in \mathbb{Q}$ there is some $R_q \subset \mathbb{R}$ with $\mu(R_q) = 1$ on which $x \leqslant x' \Rightarrow k(x,I_q) \geqslant k(x',I_q)$. Finally set $R := \bigcap_{q\in\mathbb{Q}} R_q$.

Notation 3.13. — If $\mu \in \mathrm{M}(\mathbb{R})$ we denote by $\mathrm{M}(\mu)$ and $\mathrm{P}(\mu)$ the sets of positive measures, respectively probability measures, absolutely continuous with respect to μ, and $\mathrm{M}^{\searrow}(\mu)$ and $\mathrm{P}^{\searrow}(\mu)$, or $\mathrm{M}^{\searrow}$ and $\mathrm{P}^{\searrow}$ if there is no ambiguity, their subsets of measures with bounded and decreasing density.

Remark 3.14. — A direct consequence of the first point of Proposition 3.11 is that, if $P \in \mathrm{Marg}(\mu_1, \mu_2)$ and $Q \in \mathrm{Marg}(\mu_2, \mu_3)$ have increasing kernel, so has the result $P \cdot Q$ of their composition.

Remark 3.15 ($\mathrm{M}^{\searrow}(\mu)$ is closed for the weak topology). — A measure θ belongs to $\mathrm{M}^{\searrow}(\mu)$ if and only if its density is bounded and, for all $(a, b, c, d) \in \mathbb{R}^4$:

$$(a < b < c < d \text{ and } \mu([a, b]) \cdot \mu([c, d]) \neq 0) \implies \frac{\theta([a, b])}{\mu([a, b])} \geqslant \frac{\theta([c, d])}{\mu([c, d])}. \tag{10}$$

"Only if" is clear. For the "if" part, by the Lebesgue differentiation theorem,

$$x \longmapsto \lim_{\varepsilon \to 0} \frac{\theta([x - \varepsilon, x + \varepsilon])}{\mu([x - \varepsilon, x + \varepsilon])}$$

provides a representative of the density. Now if a sequence $(\theta_n)_n$ satisfies (10) and weakly tends (see Reminder 3.26) to $\theta \in \mathrm{M}(\mu)$, θ satisfies it also (if a, b, c or d is an atom of θ, re-obtain (10) by limit of larger intervals).

Remark 3.16 (P has increasing kernel if and only if ${}^{t}P$ maps $\mathrm{M}^{\searrow}(\nu)$ to $\mathrm{M}^{\searrow}(\mu)$)

In Proposition 3.11, (b) is equivalent to the same statement with decreasing functions; in turn, transposing, this means that, for all decreasing $h : \mathbb{R} \to [0, 1]$, $(h\nu) \cdot {}^{t}P$, which is equal to $(P \cdot h) \cdot \mu$, has decreasing density.

3.2. Quantile measures and minimal couplings. — We define the quantile coupling (Definition 3.18) and quantile process law (Definition 3.20) through a minimality property that is crucial in our paper. We also state the direct and more classical Definition 3.23 of the quantile measure. Further characterizations of these coupling and process are given throughout the paper and in [8] where the approach is optimal transport. The reader can refer to [41, 45] for more background. See also the papers [23, 40].

Reminder 3.17. — Take μ and ν in $\mathrm{M}(\mathbb{R})$. Every P of $\mathrm{Marg}(\mu, \nu)$ satisfies the Fréchet–Hoeffding bound:

$$\forall (x, y) \in \mathbb{R}^2, \quad F_P(x, y) \leqslant \min(F_\mu(x), F_\nu(y)). \tag{11}$$

Proposition/Definition 3.18. — *There is a unique P such that* (11) *is an equality. We call it the Fréchet–Hoeffding, comonotonic or* quantile coupling *and denote it by* $\mathrm{Q}(\mu, \nu)$. *Said briefly:* $\mathrm{Q}(\mu, \nu) = \operatorname{loinf} \mathrm{Marg}(\mu, \nu)$.

The proof follows from Definition 3.23, see below.

Remark 3.19 (Minimality in the language of transport plans)

(a) The quantile process Q may also be defined by the fact that its transitions are *minimal* among those of all transport plans of $\mathrm{Marg}(\mu, \nu)$: for any $P \in \mathrm{Marg}(\mu, \nu)$ and fixed x, $F[\mu\llcorner_{]-\infty,x]} \cdot k_P](y) = F_P(x, y)$ for every $y \in \mathbb{R}$, hence, after Definition 3.18 and the characterization of stochastic order of Remark 3.9:

$$\forall P \in \mathrm{Marg}(\mu, \nu), \quad \mu\llcorner_{]-\infty,x]} \cdot \mathrm{Q}(\mu, \nu) \preceq_{\mathrm{sto}} \mu\llcorner_{]-\infty,x]} \cdot k_P \tag{12}$$

(in fact, $\mu|_{]-\infty,x]}.\mathrm{Q}(\mu,\nu)$ is of the type $\nu|_{]-\infty,y[}+a\delta_y$), i.e., the quantile coupling maps the measures $\mu|_{]-\infty,x]}$ to the stochastically smallest possible measures, (12) being also true for any $\theta\in\mathrm{M}^{\leqslant}(\mu)$ in place of $\mu|_{]-\infty,x]}$.

Notice that a property similar to (12), with $\preceq_C$ in place of $\preceq_{\mathrm{sto}}$, defines the (left-)curtain coupling in [4].

(b) As $\mu|_{]x,+\infty[}.k_P=\nu-\mu|_{]-\infty,x]}.k_P$, $\mu|_{]x,+\infty[}.\mathrm{Q}(\mu,\nu)$ is paradoxically *maximal* for $\preceq_{\mathrm{sto}}$, hence minimal transitions mean that the mass of μ is mixed as less as possible when transported on that of ν.

Proposition/Definition 3.20. — *If $(\mu_\tau)_{\tau\in\mathrm{T}}$ is a family of measures, there is a unique measure $\mathrm{Q}\in\mathrm{Marg}((\mu_\tau)_{\tau\in\mathrm{T}})$ such that for every $\tau\neq\sigma$, the transport plan $\mathrm{Q}^{\tau,\sigma}=(\mathrm{proj}^{\{\tau,\sigma\}})_\#\mathrm{Q}$ is the quantile coupling $\mathrm{Q}(\mu_\tau,\mu_\sigma)$. We call it the* quantile measure *of $(\mu_\tau)_{\tau\in\mathrm{T}}$.*

Proof of Propositions 3.18 *and* 3.20. — The existence parts follows from Definition 3.23 below and are proved just after it; in Proposition 3.18 uniqueness is clear; let us prove it in Proposition 3.20. It is rather easy to prove that the equality in (11) for every pair of measures implies the equality in the Fréchet–Hoeffding bound of general dimension:

$$\forall (x_i)_{i=1}^d\in\mathbb{R}^d,\quad F_{P^S}(x_1,\dots,x_d)\leqslant\min_{i\in\{1\dots,d\}}(F_{\mu_{s_i}}(x_i)),\tag{13}$$

where $S=\{s_1,\dots,s_d\}$ is any finite subset of $\mathbb{R}$. (Take j such that $F_{\mu_{s_j}}(x_j)=\min_{1\leqslant i\leqslant d}(F_{\mu_{s_i}}(x_i))$. For every $k\neq j$,

$$F[\mathrm{proj}^{\{s_j,s_k\}}_\# P^S](x_j,x_k)=F[\mu_{s_j}](x_j)$$

so that finally $P^S(]-\infty,x])=F[\mu_{s_j}](x_j)$.) Since equalities (13) for all finite $S\subset\mathbb{R}$ are a compatible set of conditions, this proves with Proposition 2.12 that there exists at most one quantile measure P in $\mathrm{Marg}((\mu_\tau)_{\tau\in\mathrm{T}})$.

Notation 3.21. — We denote by $\mathrm{Q}(\mu_{s_1},\dots,\mu_{s_d})$ and $\mathrm{Q}((\mu_t)_{t\in\mathbb{R}})$ the multidimensional quantile coupling and the quantile measure.

Remark 3.22. — Definition 3.20 uses no order on the set T of indices to define a quantile measure. So if $\mathrm{T}=\mathbb{R}$, the order of the marginals does not matter; bijections of $\mathbb{R}$ act naturally on the quantile measures and their marginals. But because the Markov property is based on the order on $\mathrm{T}=\mathbb{R}$, it will be different for our Markov-quantile measure. Only monotone bijections act naturally, see Example 5.12.

Now here is the definition of the quantile measure through the *quantile function*. It ensures the existence of $\mathrm{Q}(\mu,\nu)$ in Definition 3.18.

Definition 3.23. — Take $\mu\in\mathrm{P}(\mathbb{R})$. The *quantile function*

$$G_\mu:[0,1]\longrightarrow\mathbb{R}\cup\{-\infty,+\infty\}$$

is the increasing left-continuous function such that $(G_\mu)_\# \lambda|_{[0,1]} = \mu$, i.e., the generalized inverse of the cumulative distribution function F_μ:

$$G_\mu(q) = \inf\{x \in \mathrm{R} : F_\mu(x) \geq q\}.$$

The quantile measure $\mathrm{Q}((\mu_\tau)_{\tau\in\mathrm{T}})$ is obtained by pushing forward $\lambda|_{[0,1]}$ on $\mathrm{P}(\mathrm{R}^\mathrm{T})$ by the map $G = (G_{\mu_\tau})_{\tau\in\mathrm{T}} : x \in [0,1] \mapsto (G_{\mu_\tau}(x))_{\tau\in\mathrm{T}} \in \mathrm{R}^\mathrm{T}$. In other words, the functions G_{μ_τ} can be seen as random variables defined on $([0,1],\lambda|_{[0,1]})$ and $\mathrm{Q}((\mu_\tau)_{\tau\in\mathrm{T}})$ is the law of the process $(G_{\mu_\tau})_{\tau\in\mathrm{T}}$. One can check from this definition the equality in (11) and (13).

Remark 3.24. — When $(\mu_\tau)_{\tau\in\mathrm{T}} = (\mu,\nu)$, $\mathrm{Law}(G_\mu, G_\nu) = \mathrm{Q}(\mu,\nu)$. Now

$$\mu \preceq_{\mathrm{sto}} \nu \iff F_\mu \geq F_\nu \iff G_\mu \leq G_\nu,$$

hence $\mu \preceq_{\mathrm{sto}} \nu$ if and only if $\mathrm{Q}(\mu,\nu)$ is an increasing coupling, i.e., concentrated on $\{(x,y)\in\mathrm{R}^2 : x \leq y\}$.

Remark 3.25 (Products of quantile couplings have increasing kernel and map $\mathrm{M}^{\nearrow}$ on $\mathrm{M}^{\nearrow}$)

We will see in Section 4.1 that $\mathrm{Q}(\mu,\nu)$ is a composition of two transport plans with increasing kernel. Therefore by Remark 3.14 it has increasing kernel. Since ${}^t\mathrm{Q}(\mu,\nu) = \mathrm{Q}(\nu,\mu)$, Remark 3.16 ensures that $\mathrm{Q}(\nu,\mu)$ maps $\mathrm{M}^{\nearrow}(\mu)$ to $\mathrm{M}^{\nearrow}(\nu)$.

All this also ensure both properties for products of quantile couplings.

3.3. Distances ρ and $\widetilde{\rho}$

3.3.1. *A distance that metricizes* $\mathrm{Marg}(\mu,\nu)$

Reminder 3.26. — We remind the reader of the "Portmanteau theorem" (see, e.g., [7, Th. 2.1]): the weak convergence on some metric space E endowed with its Borel σ-algebra is (equivalently) defined by:

(14) $$P_n \underset{n\to\infty}{\longrightarrow} P \text{ if } P_n(R) \underset{n\to\infty}{\longrightarrow} P(R) \text{ for all } R \text{ such that } P(\partial R) = 0.$$

In R^n, it is equivalent to consider in (14) only sets R of the form $\prod_{i=1}^n]-\infty, x_i]$, see [7, Exam. 2.3 p. 18].

Proposition/Notation 3.27. — *Let ρ be the function defined on $\mathrm{P}(\mathrm{R}^d)^2$ by:*

$$\rho(P,Q) = \|F_P - F_Q\|_\infty.$$

(See Definition 3.2 for F_P and F_Q.) This is a distance. If $\mu_1,\dots,\mu_d$ are probability measures on R*, it induces the weak topology on* $\mathrm{Marg}((\mu_i)_{1\leq i\leq d})$*. More precisely, for any $P \in \mathrm{Marg}((\mu_i)_{1\leq i\leq d})$ and any sequence $(P_n)_{n\in\mathrm{N}}$ of elements of* $\mathrm{Marg}((\mu_i)_{1\leq i\leq d})$*, the following are equivalent:*

(i) *For all $x = (x_i)_{i=1}^d$, $P\big(\partial\big(\prod_{i=1}^d]-\infty, x_i]\big)\big) = 0 \Rightarrow \lim_n F_{P_n}(x) = F_P(x)$,*
(ii) *For all $x = (x_i)_{i=1}^d$, $\lim_n F_{P_n}(x) = F_P(x)$,*
(ii') *For all $x = (x_i)_{i=1}^d$, $\lim_n\big(P_n\big(\prod_i]-\infty, x_i[\big)\big) = P\big(\prod_i]-\infty, x_i[\big)$,*
(iii) *$(F_{P_n})_{n\in\mathrm{N}}$ converges uniformly to F_P.*

Proof. — The fact that ρ is a distance is immediate. To get (iii) $\Rightarrow$ (ii') use that $]-\infty,x[=\bigcup_{y<x}]-\infty,y]$. Now it suffices to prove (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii).

(i) $\Rightarrow$ (ii). Take $Q\in\mathrm{Marg}((\mu_i)_{1\leqslant i\leqslant d})$ and $b=(b_i)_{i=1}^d$. For $b'=(b'_i)_{i=1}^d\geqslant b$:

$$F_Q(b')-F_Q(b)=Q(R')-Q(R)\leqslant\mu_1(]b_1,b'_1])+\cdots+\mu_d(]b_d,b'_d]),\tag{15}$$

where $R'=]-\infty,b']$ and $R=]-\infty,b]$, which shows that $(F_Q)_Q$ is "equicontinuous on the right". Take $\varepsilon>0$. Since each $b_i\in\mathbb{R}$ can be approached from the right by a sequence of non-atomic points for μ_i, there exists $b'\geqslant b$ such that

$$\mu_1(]b_1,b'_1])+\cdots+\mu_d(]b_d,b'_d])<\varepsilon\quad\text{and}\quad\mu_1(b'_1)+\cdots+\mu_d(b'_d)=0,$$

hence in particular $Q(\partial R')=0$ for every $Q\in\mathrm{Marg}((\mu_i)_{1\leqslant i\leqslant d})$ so that (i) applies to R' for P and all the P_n. Thus:

$$\begin{aligned}|P_n(R)-P(R)|&\leqslant|P_n(R)-P_n(R')|+|P_n(R')-P(R')|+|P(R')-P(R)|\\&\leqslant 2\varepsilon+|P_n(R')-P(R')|,\end{aligned}$$

and if (i) holds, $|P_n(R')-P(R')|\leqslant\varepsilon$ for n great enough, hence (ii) follows.

(ii) $\Rightarrow$ (iii). We apply an adapted version of the prior argument. Suppose (ii) and take $\varepsilon>0$. For every i there exists a finite sequence

$$-\infty=b_0^{(i)}(\varepsilon)<\cdots<b_{N(\varepsilon,i)}^{(i)}(\varepsilon)=+\infty$$

avoiding the big atoms of μ_i, i.e., so that $\mu_i(]b_k^{(i)},b_{k+1}^{(i)}[)<\varepsilon$ (this is classical and is proved, e.g., in [7, §12], which deals with the modulus of continuity of càdlàg paths). Every $R=]-\infty,b]$ contains some rectangle R_ε^- and is included in the interior of some rectangle R_ε^+, both bounded by consecutive points $(b_k^{(i)})_{i,k}$. Using again (15) for the first and last terms:

$$|P_n(R)-P(R)|\leqslant\underbrace{|P_n(R)-P_n(R_\varepsilon^-)|}_{\leqslant d\varepsilon}+\underbrace{|P_n(R_\varepsilon^-)-P(R_\varepsilon^-)|}_{\xrightarrow[n\to\infty]{}0\text{ by (ii)}\ (*)}+\underbrace{|P(R_\varepsilon^-)-P(R)|}_{\leqslant d\varepsilon},$$

where $(*)$ is uniform as the rectangles R_ε^- are finitely many. We get (iii).

3.3.2. *Another expression for ρ; an alternative distance $\widetilde{\rho}$*. — Let $\mu_1,\dots,\mu_d$ be probability measures on $\mathbb{R}$ and P, Q stand for elements of $\mathrm{Marg}((\mu_i)_{i=1}^d)$.

Notation 3.28. — When f, g and $(f_i)_{i=1}^d$ are functions, by a slight abuse in this subsection, $f\cdot g$ denotes the function $(x,y)\mapsto f(x)g(y)$ and $\prod_{i=1}^d f_i$ the function $(x_1,\dots,x_d)\mapsto\prod_{i=1}^d f_i(x_i)$

Remark 3.29. — By definition, $\rho(P,Q)=\sup\big|\int\prod_{i=1}^d f_i\,\mathrm{d}P-\int\prod_{i=1}^d f_i\,\mathrm{d}Q\big|$, where each f_i ranges over $\{\mathbb{1}_{]-\infty,x]}:x\in\mathbb{R}\}$. But in fact:

Proposition/Definition 3.30

(a) *For any P and Q in* $\mathrm{Marg}((\mu_i)_{i=1}^d)$*:*

$$\rho(P,Q)=\sup\Big|\int\prod_{i=1}^d f_i\,\mathrm{d}P-\int\prod_{i=1}^d f_i\,\mathrm{d}Q\Big|,\tag{16}$$

where each f_i ranges over $\{f:\mathbb{R}\to[0,1]:f$ is decreasing$\}$.

(b) *Proposition* 3.27 *may be stated with a distance* $\widetilde{\rho}$ *based on* $x \in \mathbb{R}^n \mapsto P([x,+\infty[)$ *in place of* $F[P]$. *Then* $\widetilde{\rho}$ *satisfies* (16) *with* increasing *functions* f_i.

Proof. — Once (a) is shown, (b) is clear. To show (a), by Remark 3.29, we have only to prove $\geqslant$. For $f:\mathbb{R}\to[0,1]$ a decreasing function and $I(t) := f^{-1}([t,+\infty[)$, $f(x) = \int_0^1 \mathbb{1}_{I(t)}(x)\mathrm{d}t$, thus for $(f_i)_{i=1}^d$ such functions:

$$\Big(\prod_{i=1}^d f_i\Big)(x_1,\dots,x_d) = \int_{t\in[0,1]^d} \mathbb{1}_{R(t)}(x_1,\dots,x_d)\mathrm{d}t,$$

where $R(t) = I(t_1)\times\cdots\times I(t_d)$. Therefore:

$$\begin{aligned}\Big|\int\prod_{i=1}^d f_i\,\mathrm{d}P - \int\prod_{i=1}^d f_i\,\mathrm{d}Q\Big| &= \Big|\int_{t\in[0,1]^d} P(R(t)) - Q(R(t))\mathrm{d}t\Big|\\ &\leqslant \int_{t\in[0,1]^d} |P(R(t)) - Q(R(t))|\mathrm{d}t\\ &\leqslant \rho(P,Q).\end{aligned}$$

Proposition 3.30 has a corollary in the case of transport plans ($d=2$).

Proposition 3.31. — *Take* μ, μ' *in* $\mathrm{P}(\mathbb{R})$ *and* P, Q *in* $\mathrm{Marg}(\mu,\mu')$. *Then:*

$$\rho(P,Q) = \sup\{\rho(\theta . P, \theta . Q) : \theta \in \mathrm{M}^{\&}(\mu) \text{ and } \theta \text{ has density bounded by } 1\},$$

where the distance ρ *on the right of the equality is that on* $\mathrm{P}(\mathbb{R})$.

Hence if, for some $\nu \in \mathrm{P}(\mathbb{R})$, $R \in \mathrm{Marg}(\nu,\mu)$ *is a transport preserving* $\mathrm{M}^{\&}$, *i.e.,* $\forall\eta \in \mathrm{M}^{\&}(\nu)$, $\eta . R \in \mathrm{M}^{\&}(\mu)$, *then* $\rho(R . P, R . Q) \leqslant \rho(P,Q)$.

Proof. — For the first equality, as $\int f . g\,\mathrm{d}P = ((f\mu) . P) . g$, by Proposition 3.30: $\rho(P,Q) = \sup_{f,g}|((f\mu).P).g - ((f\mu).Q).g|$ and $\rho(\theta.P,\theta.Q) = \sup_g|(\theta.P).g - (\theta.Q).g|$. Now θ is as in the proposition if and only if $\theta = f\mu$ with $f:\mathbb{R}\to[0,1]$, decreasing. The result follows. Then if $R \in \mathrm{Marg}(\nu,\mu)$ is as claimed,

$$\rho(RP,RQ) = \sup\{\rho(\theta . P, \theta . Q) \mid \theta = \overline{\theta} . R,$$

where $\overline{\theta} \in \mathrm{M}^{\&}(\mu)$ and $\overline{\theta}$ has density bounded by $1\}$. This set is included in that of the proposition, since the action of R on $\overline{\theta}$ does not increase the maximum of its density.

3.4. Proof of Lemma 1.19 and remarks about it. — We prove Lemma 1.19 on the existence of a process consisting exclusively of increasing paths. Our proof requires the use of stoinf and stosup introduced in Section 3.1.

Proof of Lemma 1.19. — Let $(\mu_t)_{t\in\mathbb{R}}$ be an increasing family of probability measures for $\leqslant_{\mathrm{sto}}$ and P be a Markov measure in $\mathrm{Marg}((\mu_t)_{t\in\mathbb{R}})$ such that for every $S = \{s_1,\dots,s_d\}$, the measure $(\mathrm{proj}^S)_\# P$ is concentrated on

$$\{(x_1,\dots,x_d)\in\mathbb{R}^d : x_1 \leqslant \cdots \leqslant x_d\}.$$

Set $\mu_{t^-} := \mathrm{stosup}_{s<t}\,\mu_s$ and $\mu_{t^+} := \mathrm{stoinf}_{s>t}\,\mu_s$, that are also the left and right limits of $(\mu_t)_t$ for the weak topology, see Remark 3.9. Since $(\mu_t)_{t\in\mathbb{R}}$ is increasing for $\preceq_{\mathrm{sto}}$, we have $\mu_{t^-} \preceq_{\mathrm{sto}} \mu_t \preceq_{\mathrm{sto}} \mu_{t^+}$ and $t \in \mathbb{R}$ is a discontinuity time of $(\mu_t)_{t\in\mathbb{R}}$ for the weak topology if and only if $\mu_{t^-} \neq \mu_{t^+}$. Such points are at most countably many. Indeed, the regions in $\mathbb{R}^2$ between the graphs of $F[\mu_{t^-}]$ and $F[\mu_{t^+}]$, for the discontinuity times t, are disjoint and of positive Lebesgue measure, hence are at most countably many, by σ-additivity of the measure.

Let C be a countable dense subset of $\mathbb{R}$ containing the discontinuity points. Introduce:

$$N = \{x \in \mathbb{R}^{\mathbb{R}} : \exists\,(s,t) \in C^2, s<t \text{ and } x(s) > x(t)\}.$$

Being a countable union of P-null sets, N is P-null. Now take $(X_t)_{t\in\mathbb{R}}$ with law P, e.g., take the canonical process $\Omega := \mathbb{R}^{\mathbb{R}}$ and $X_t = \mathrm{proj}^t : x \in \mathbb{R}^{\mathbb{R}} \mapsto x(t)$. We define $(\widetilde{X}_t)_t$ as null functions on N, and as follows on $\Omega \smallsetminus N$:

– for $t \in C$, $\widetilde{X}_t(\omega) = X_t(\omega)$,

– for $t \notin C$, $\widetilde{X}_t(\omega) = \lim_{s<t,\, s\in C} X_s(\omega)$.

Hence for every $\omega \in \Omega$, the curve $t \in C \mapsto X_t(\omega)$ is increasing and, even better, $t \in \mathbb{R} \mapsto X_t(\omega)$ is increasing. We are left to prove that $X_t = \widetilde{X}_t$ almost surely. This is clear for $t \in C$. For each $t \in \mathbb{R} \smallsetminus C$, $\{\omega \in \Omega : \exists\, s \in C, s<t \text{ and } X_s(\omega) > X_t(\omega)\}$ is a union of null sets, hence is null, so almost surely, $X_t \geqslant \sup_{s<t,\, s\in C} X_s = \widetilde{X}_t$. Besides $X_s \to_{s<t,\, s\in C, s\to t} \widetilde{X}_t$ almost surely and thus in law, so that $\mathrm{Law}(\widetilde{X}_t) = \mu_{t^-}$. Moreover, t is a continuity point of $(\mu_t)_t$, so that $\mathrm{Law}(\widetilde{X}_t) = \mu_t = \mathrm{Law}(X_t)$. Thus, $X_t = \widetilde{X}_t$ almost surely.

Remark 3.32

(a) If $(\mu_t)_t$ is moreover left-continuous for the weak topology, we can adapt the proof of Lemma 1.19 so that $s \mapsto \widetilde{X}_s(\omega)$ is increasing and left continuous by using the formula:

$$\widetilde{X}_t(\omega) = \lim_{s<t,\, s\in C} X_s(\omega) \text{ if } \omega \in \Omega \smallsetminus N \text{ and } X_t(\omega) = 0 \text{ otherwise.}$$

(b) If $(\mu_t)_t$ is this time moreover right-continuous, using a symmetric construction all the curves can be chosen càdlàg (right-continuous, with limit on the left at any point).

(c) However, for a continuous $\mu = (\mu_t)_t$, it is false that some choice of $(s,\omega) \mapsto \widetilde{X}_s(\omega)$ with $\widetilde{X}_s = X_s$ almost surely (fore every s), can make $s \mapsto \widetilde{X}_s(\omega)$ almost surely continuous. A simple (counter)example is: $t \in [0,1] \mapsto (1-t)\delta_0 + t\delta_1$. In this example, it is even impossible with the looser constraint that $\widetilde{X}$ only satisfies $\widetilde{X}_s = \mu_s$. With this looser constraint, it is shown in [8] that an assumption ensuring the continuity of the curve $s \mapsto \widetilde{X}_s(\omega)$ (and even the finiteness of its energy), when $\widetilde{X}$ is the quantile or the Markov-quantile process, is that $\mu = (\mu_t)_t$ has *finite energy* in the sense of Section 1.4.

Carrying on with the similarity between $\preceq_{\mathrm{sto}}$ and $\preceq_C$ established in Theorem C we mention that Kellerer's theorem has also been revisited under continuity assumptions, see [35, 20, 3]. In particular, it has been proved that if $(\mu_t)_{t\in\mathbb{R}}$ is right-continuous the associated martingale can be defined in the space of càdlàg paths. Moreover, Lowther proved [34, Th. 1.5] that there exists a unique function that associates a strongly Markovian martingale with every continuous increasing in convex order $t \in \mathbb{R} \mapsto \mu_t$ provided this function is moreover continuous for the pointwise convergence.

3.5. Some remarks following from Proposition 3.27; proofs of Lemmas 2.23, and 2.24

3.5.1. *The remarks on Proposition* 3.27*; proof of Lemma* 2.24

Remark 3.33 ($\{P \in \mathrm{Marg}(\mu,\nu) : P$ maps $\mathcal{M}^{\nearrow}(\mu)$ to $\mathcal{M}^{\nearrow}(\nu)\}$ is closed for the weak topology)

Any $\theta \in \mathcal{M}^{\nearrow}(\mu)$ is an increasing limit of positive combinations of characteristic functions $\mathbb{1}_{]-\infty,x]}$, so $P \in \mathrm{Marg}(\mu,\nu)$ maps $\mathcal{M}^{\nearrow}(\mu)$ in $\mathcal{M}^{\nearrow}(\nu)$ if and only if it maps $\{\mu\lfloor_{]-\infty,x]} : x \in \mathbb{R}\}$ in it. Now take a sequence $(P_n)_n \in \mathrm{Marg}(\mu,\nu)^{\mathbb{N}}$ of transport plans having this property and converging weakly to $P \in \mathrm{Marg}(\mu,\nu)$. After Proposition 3.27, $\|F[P_n] - F[P]\|_\infty \to 0$; in particular, for any $x \in \mathbb{R}$, $\|F[P_n](x,\cdot) - F[P](x,\cdot)\|_\infty \to 0$, i.e., $\|F[\mu\lfloor_{]-\infty,x]}.P_n] - F[\mu\lfloor_{]-\infty,x]}.P]\|_\infty \to 0$, i.e., $\mu\lfloor_{]-\infty,x]}.P_n$ converges weakly to $\mu\lfloor_{]-\infty,x]}.P$. By Remark 3.15, $\mu\lfloor_{]-\infty,x]}.P \in \mathcal{M}^{\nearrow}(\nu)$, we are done.

Now the little Lemma 2.24, which we use several times, is immediate.

Proof of Lemma 2.24. — Apply Remarks 3.33 and 3.16.

Remark 3.34. — Take $(H_i)_{i=1}^d$ increasing real functions, and $H = H_1 \otimes \cdots \otimes H_d : \mathbb{R}^d \to \mathbb{R}^d$. Then $H_\# : \mathrm{Marg}((\mu_i)_{i=1}^d) \to \mathrm{Marg}((H_i(\mu_i))_{i=1}^d)$ is contracting for ρ. To check it, think that $(F[H_\# P] - F[H_\# Q])(x_1,\dots,x_d) = (H_\# P - H_\# Q)(\prod_i]-\infty, x_i])$ equals a term of the type $(H_\# P - H_\# Q)(\prod_i]-\infty, y_i \rrbracket)$, and use Proposition 3.27. Thus if $(P_n)_n$ converges to P in $\mathrm{Marg}(\mu_1,\dots,\mu_d)$, $(H_\# P_n)_n$ converges to $H_\# P$ in $\mathrm{Marg}(H_1(\mu_1),\dots,H_d(\mu_d))$.

The next remark is neither related to Proposition 3.27 nor to Lemma 2.24 but is an analogue of Remark 3.34 with $\preceq_{\mathrm{sto}}$ in place of ρ.

Remark 3.35. — If P,Q are in $\mathcal{P}(\mathbb{R}^d)$, the $(H_i)_{i=1}^d$ are increasing functions, and $H := H_1 \otimes \cdots \otimes H_d$ then $H_\# : \mathcal{P}(\mathbb{R}^d) \to \mathcal{P}(\mathbb{R}^d)$ is increasing for $\preceq_{\mathrm{lo}}$, i.e., $P \preceq_{\mathrm{lo}} Q \Rightarrow H_\# P \preceq_{\mathrm{lo}} H_\# Q$. To check it, think that

$$(F[H_\# P] - F[H_\# Q])(x_1,\dots,x_d) = (H_\# P - H_\# Q)(\textstyle\prod_i]-\infty, x_i])$$

equals a term of the type $(H_\# P - H_\# Q)(\prod_i]-\infty, y_i \rrbracket)$, and, for the indices i such that "$\rrbracket$" stands for "[", use that $]-\infty, x_i[= \bigcup_n]-\infty, x_i - 1/n]$.

Remark 3.36. — Notice then that $\rho(T,T') = \rho({}^{\mathrm{t}}T, {}^{\mathrm{t}}T')$ (see Definition 2.7 for ${}^{\mathrm{t}}T$). Thus, in Proposition 3.31, if $R \in \mathrm{Marg}(\mu',\nu)$ and if R has increasing kernel (i.e., by Remark 3.16, ${}^{\mathrm{t}}R$ maps $\mathcal{M}^{\nearrow}(\nu)$ into $\mathcal{M}^{\nearrow}(\mu')$):

$$\rho(PR,QR) = \rho({}^{\mathrm{t}}(PR), {}^{\mathrm{t}}(QR)) = \rho({}^{\mathrm{t}}R\,{}^{\mathrm{t}}P, {}^{\mathrm{t}}R\,{}^{\mathrm{t}}Q)) \leqslant \rho({}^{\mathrm{t}}P, {}^{\mathrm{t}}Q) = \rho(P,Q).$$

3.5.2. *Proof of Lemma* 2.23. — First we prove Lemma 3.37, then its consequence Proposition 3.38, and finally Lemma 2.23.

LEMMA 3.37. — *For every $n \in \mathbb{N}$, let $P_n \in \mathrm{Marg}(\mu,\nu)$ have increasing kernel. Suppose moreover that $(P_n)_{n\in\mathbb{N}}$ converges to P_0. Let $h : [0,1] \to [0,1]$ be an increasing function and for every $n \in \mathbb{N}$, $\widetilde{h}_n = P_n . h$. Then $(\widetilde{h}_n)_{n\in\mathbb{N}}$ converges to $\widetilde{h}_0$, μ-almost surely.*

Proof. — By the equivalence proved in Proposition 3.11, the sequence $(\widetilde{h}_n)_{n\in\mathbb{N}}$ is increasing. By Propositions 3.30(b) and 3.27, $\widetilde{\rho}(P_n,P_0) \to_{n\to\infty} 0$, i.e., for every increasing g with values in $[0,1]$, $\int g(y)h(z)\mathrm{d}P_n(y,z) \to \int g(y)h(z)\mathrm{d}P_0(y,z)$. Hence:

$$\int g(x)\widetilde{h}_n(x)\,\mathrm{d}\mu(x) \quad\to\quad \int g(x)\widetilde{h}_0(x)\,\mathrm{d}\mu(x), \tag{17}$$

by definition of $\widetilde{h}_n$. This also holds if g is the difference of two increasing functions, in particular $g = \mathbb{1}_{[a,b]} = \mathbb{1}_{[a,+\infty[} - \mathbb{1}_{[b,+\infty[}$.

Let A be the set of the elements x of $[0,1]$ such that $\mu[x,1]\mu[0,x] > 0$ and for every $n \in \mathbb{N}$, μ-esssup$_{[0,x]}\widetilde{h}_n = \mu$-essinf$_{[x,1]}\widetilde{h}_n = \widetilde{h}_n(x)$. In case $\mu = \lambda$, since $\widetilde{h}_n$ is increasing, hence has at most a countable number of discontinuity points, $\mu(A)$ is 1. This also holds in the general case and is given by the increasing functions $\widetilde{h}_n \circ G_\mu$, we leave the details to the reader. Now we take any $x \in A$ and prove $\widetilde{h}_n(x) \to \widetilde{h}_0(x)$. It is enough to prove $\limsup_n \widetilde{h}_n(x) \leq \widetilde{h}_0(x)$, since $\liminf_n \widetilde{h}_n(x) \leq \widetilde{h}_0(x)$ can be proved symmetrically. Suppose, for contradiction:

$$\limsup_n \widetilde{h}_n(x) > \widetilde{h}_0(x) + \varepsilon, \quad \text{for some } \varepsilon > 0.$$

As $\widetilde{h}_0$ is increasing and $\widetilde{h}_0(x) = \mu$-essinf$_{[x,1]}\widetilde{h}_0$ there exists $y > x$ such that $\mu[x,y] > 0$ and:

$$\widetilde{h}_0(x) \leq \frac{1}{\mu[x,y]}\int_{[x,y]} \widetilde{h}_0\,\mathrm{d}\mu \leq \widetilde{h}_0(x) + \varepsilon/2.$$

This contradicts the facts that $\frac{1}{\mu[x,y]}\int_{[x,y]}\widetilde{h}_n\,\mathrm{d}\mu \to_{n\to\infty} \frac{1}{\mu[x,y]}\int_{[x,y]}\widetilde{h}_0\,\mathrm{d}\mu$, obtained with $g = \mathbb{1}_{[x,y]}$ in (17), and that $\widetilde{h}_n$ is increasing.

PROPOSITION 3.38. — *Let P_n tend to P_0 in $\mathrm{Marg}(\mu_1,\dots,\mu_d,\eta)$ and P'_n tend to P'_0 in $\mathrm{Marg}(\eta,\nu)$. Assume moreover that P'_n has increasing kernel for every n. Then $P'_n \circ P_n$ tends to $P'_0 \circ P_0$.*

Proof. — By Propositions 3.30(b) and 3.27, we must show that

$$\int f(x)g(y)h(z)\mathrm{d}(P_n \circ P'_n)(x,y,z) \quad\to\quad \int f(x)g(y)h(z)\mathrm{d}(P_0 \circ P'_0)(x,y,z),$$

where $f((x_i)_{i=1}^d) = \prod_i f_i(x_i)$ and g, h and the f_i are any increasing functions from $\mathbb{R}$ to $[0,1]$. For all $n \geq 0$:

$$\begin{aligned}\int f(x)g(y)h(z)\mathrm{d}(P_n \circ P'_n)(x,y,z) &= \int f(x)g(y)\int h(z)k_{P'_n}(y,\mathrm{d}z)\mathrm{d}P_n(x,y)\\ &= \int f(x)(g\widetilde{h}_n)(y)\mathrm{d}P_n,\end{aligned}$$

where $\widetilde{h}_n = P_n . h$, by Definition 2.8 and Section 2.1.3.

By Lemma 3.37, $(g\widetilde{h}_n)_n$ pointwise converges to $g\widetilde{h}_0$, almost surely. Hence by the dominated convergence theorem:

$$\left|\int f(x)(g\widetilde{h}_n)(y)\mathrm{d}P_0 - \int f(x)(g\widetilde{h}_0)(y)\mathrm{d}P_0\right| \longrightarrow 0.$$

Moreover, $\widetilde{\rho}(P_n, P_0)$ tends to zero and for every n, $g\widetilde{h}_n$ is an increasing functions taking values in $[0,1]$. Hence, again by Propositions 3.30(b) and 3.27:

$$\left|\int f(x)(g\widetilde{h}_n)(y)\mathrm{d}P_n - \int f(x)(g\widetilde{h}_n)(y)\mathrm{d}P_0\right| \longrightarrow 0. \tag{18}$$

Therefore with the triangle inequality, (18) holds with $\widetilde{h}_0$ in place of $\widetilde{h}_n$ on the right, which is exactly what we claimed.

Remark 3.39 (Counterexamples to Proposition 3.38)

(a) Case where the marginals are not fixed. Take $d=1$, $\mu=\nu=\frac{1}{2}(\delta_{-1}+\delta_1)$ and $\eta=\frac{1}{2}(\delta_{-1/n}+\delta_{1/n})$ and its limit δ_0 in the parameter n. We consider the quantile couplings $P_n=\frac{1}{2}(\delta_{-1,-1/n}+\delta_{1,1/n})\in\mathrm{Marg}(\mu_0,\mu_1)$ and $P'_n={}^{t}P_n\in\mathrm{Marg}(\mu,\eta)$. They have increasing kernel but not fixed marginals. On the one hand,

$$P_n\otimes P'_n=\frac{1}{2}(\delta_{-1,-1/n,-1}+\delta_{1,1/n,1})\xrightarrow[n\to\infty]{}\frac{1}{2}(\delta_{-1,0,-1}+\delta_{1,0,1}).$$

On the other hand,

$$(\lim_n(P_n)_n)\otimes(\lim_n(P'_n)_n)=\frac{1}{4}(\delta_{1,0,1}+\delta_{1,0,-1}+\delta_{-1,0,1}+\delta_{-1,0,-1}).$$

(b) Case where the transitions are not increasing. See the example in [28] after the proof of Satz 14. Here the marginals are fixed but the transport plans neither have Lipschitz nor increasing kernel.

We can now prove Lemma 2.23. It follows directly from Proposition 3.38.

Proof of Lemma 2.23. — Consider a sequence $P^n_{1,2}\otimes\cdots\otimes P^n_{d,d+1}$ with limit $P\in\mathrm{Marg}(\mu_1,\ldots,\mu_{d+1})$ and denote $\mathrm{proj}^{i,i+1}_{\#}P$ by $P_{i,i+i}$. Since $(\mathrm{proj}^{i,i+1})_{\#}$ is continuous, $P_{i,i+1}$ is the limit of $(P^n_{i,i+1})_n$, hence $P_{i,i+1}\in\mathrm{I}_{i,i+1}$. Proposition 3.38 used by induction shows that $P=P_{1,2}\otimes\cdots\otimes P_{n,n+1}$.

4. Construction and characterization of the Markov-quantile process

Now $(\mu_t)_{t\in\mathbb{R}}$ is a family of probability measures on $\mathbb{R}$, F_t *denotes the cumulative distribution function F_{μ_t} of μ_t and G_t its quantile function, see Section* 3.2. In this section we build the Markov-quantile measure MQ and prove Theorems A and B. Our proof of Theorems A–B is based on transport plans $L_{[s,t]}\in\mathrm{Marg}(\lambda,\lambda)$, defined in Section 4.1, that are the 2-marginals of an important auxiliary process law called the quantile level measure Lev $\in\mathrm{Marg}((\lambda_t)_{t\in\mathbb{R}})$, where each λ_t is a copy indexed by t of $\lambda=\lambda_{[0,1]}$. Here is the link between Lev and MQ: G_t maps $([0,1],\lambda_t)$ to $(\mathbb{R}\cup\{\pm\infty\},\mu_t)$;

set $G = \prod_{t\in\mathbb{R}} G_t$, so that $G_\#$ maps $\mathrm{Marg}((\lambda_t)_{t\in\mathbb{R}})$ to $\mathrm{Marg}((\mu_t)_{t\in\mathbb{R}})$. Then, as proved in the proof Theorem 4.21 p. 52:

$$G_\# \mathrm{Lev} = \mathrm{MQ}. \tag{19}$$

Remark 4.1. — Readers familiar with Mathematical Statistics may compare $L_{[s,t]}$ with the notion of *copula* and (19) with Sklar's theorem. In particular, the (bivariate) cumulative distribution function of $\mathrm{Lev}^{s,t} = L_{[s,t]}$ is a copula associated with $\mathrm{MQ}^{s,t}$. It is defined as the pointwise maximum of the family of copulas made of the cumulative distribution functions associated with $(L_{[s,t]\cap R})_{R\subset\mathbb{R}}$, where R ranges over finite sets, see Definitions 3.4 and 3.5, Lemma 3.8, Proposition 3.27 and Notation 4.10. Although the link is immediate we will not use further the terminology of Copula Theory.

This section is divided in three. In Section 4.1 we define the coupling L_R for all $R\subset\mathbb{R}$, using the key monotonicity Lemma 4.9. In Section 4.2 we define MQ and prove Theorem A, purely via the 2-marginals $(\mathrm{MQ}^{s,t})_{s<t} := ((G_s\times G_t)_\# L_{]s,t[})_{s<t} = ((G_s\times G_t)_\# L_{[s,t]})_{s<t}$, i.e., without introducing Lev. Finally in Section 4.3 we prove Theorem B, i.e., a refinement of (19) and the approximation of Lev and MQ by sequences $(\mathrm{Lev}_{R_n})_n$ and $(\mathrm{Q}_{[R_n]})_n$, respectively.

4.1. Transitions kernels to and from the space of quantiles $[0,1]$.

— In this paper we will need to consider the quantile couplings of the measures μ_t with the reference measure $\lambda = \lambda|_{[0,1]}$.

Notation 4.2 (q_r, k_r, ${}^tk_r$). — For all $r\in R$ we set $q_r = \mathrm{Q}(\lambda,\mu_r)$; thus ${}^tq_r$ (see Definition 2.7) is $\mathrm{Q}(\mu_r,\lambda)$. Those couplings admit the respective disintegration kernels:

$$k_r : (\alpha, B) \longmapsto \delta_{G_r(\alpha)}(B),\quad {}^tk_r(x,\cdot) = \begin{cases}\delta_{F_r(x)} & \text{if } \mu_r(x)=0\\ (\mu_r(x))^{-1}\lambda|_{]F_r(x^-),F_r(x)[} & \text{if } \mu_r(x)>0.\end{cases}$$

Remark 4.3

(a) Notice that ${}^tq_s\,.\,q_t = \mathrm{Joint}(\mu_s, {}^tk_s\,.\,k_t) = \mathrm{Q}^{s,t}$, where $\mathrm{Q}^{s,t} = \mathrm{Q}(\mu_s,\mu_t)$, see Definition 3.20. Indeed, ${}^tq_s\,.\,q_t = {}^tq_s\,.\,\mathrm{Id}_{\lambda,2}\,.\,q_t$, then apply (b) below, which will also be useful farther.

(b) If some ordered pair (U,V) of variables has law $T\in\mathrm{Marg}(\lambda,\lambda)$, then

$$\mathrm{Marg}(\mu_s,\mu_t)\ni {}^tq_s\,.\,T\,.\,q_t = \mathrm{Law}(G_s(U),G_t(V)) = (G_s\times G_t)_\# T.$$

Indeed, ${}^tq_s\,.\,T\,.\,q_t = \mathrm{proj}^{1,4}_\#({}^tq_s\otimes T\otimes q_t)$ and the 4-times process $(G_{\mu_s}(U),U,V,G_{\mu_t}(V))$ has law ${}^tq_s\otimes T\otimes q_t$ and is Markov, since the σ-fields spanned by U and $\{U, G_{\mu_s}(U)\}$ are the same.

(c) From (a) we get ${}^tq_r\,.\,q_r = \mathrm{Id}_{\mu_r,2}$, so that ${}^tq_r\,.\,q_r\,.\,{}^tq_r = {}^tq_r$ and $q_r\,.\,{}^tq_r\,.\,q_r = q_r$. *However*, $q_r\,.\,{}^tq_r \neq \mathrm{Id}_{\lambda,2}$. Indeed, $k_r\,.\,{}^tk_r$ maps any quantile level $\alpha\in\,]0,1[$ on itself *except when* $G_r(\alpha)$ is an atom of μ_r. Actually, μ_r-almost surely:

$$k_r\,.\,{}^tk_r(\alpha) = \begin{cases}\delta_\alpha & \text{if } \mu_r(G_r(\alpha))=0\\ (\alpha^+-\alpha^-)^{-1}\lambda|_{]\alpha^-,\alpha^+[} & \text{if } \alpha\in\,]\alpha^-,\alpha^+[,\end{cases}$$

where $]\alpha^-,\alpha^+[$ denotes any set $A_{r,x}$ as follows.

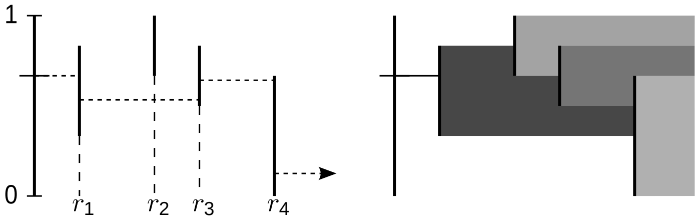

Figure 2. Composition of kernels ℓ_r

Notation 4.4 (Atomic levels, ℓ_r)

(a) We denote by $A_{r,x}$ the interval $]F_r(x-), F_r(x)[$ of quantile levels merged by G_r on some atom x of μ_r, and by A_r the set $\bigcup_{\mu_r(x)>0} A_{r,x} \subset [0,1]$ of "atomic levels" of μ_r.

(b) We denote $k_r \cdot {}^{\mathrm{t}}k_r$ described in Remark 4.3(c) by ℓ_r.

Remark 4.5. — Measures $\theta \ll \lambda$ are transported by ℓ_r as follows: $\theta \cdot \ell_r$ coincides with θ on $]0,1[\smallsetminus A_r$, and on each $A_{r,x}$ it has constant density and mass $\theta(A_{r,x})$, i.e., equals $(\alpha^+ - \alpha^-)^{-1}\theta(A_{r,x})$. Equivalently, $F[\theta \cdot \ell_r]$ is continuous, equal to $F[\theta]$ on $]0,1[\smallsetminus A_{r,x}$, and affine on each connected component $A_{r,x}$ of A_r.

The product $\mathrm{Q}^{s,r_1} \cdot \mathrm{Q}^{r_1,r_2} \cdots \mathrm{Q}^{r_{m-1},r_m} \cdot \mathrm{Q}^{r_m,t}$ appearing in Theorem A is more deeply analyzed in Theorem B. Its kernel reads:

$$\begin{aligned}
({}^{\mathrm{t}}k_s \cdot k_{r_1}) \cdot ({}^{\mathrm{t}}k_{r_1} \cdot k_{r_2}) \cdots ({}^{\mathrm{t}}k_{r_{m-1}} \cdot k_{r_m}) \cdot ({}^{\mathrm{t}}k_{r_m} \cdot k_t) \\
= {}^{\mathrm{t}}k_s \cdot (k_{r_1} \cdot {}^{\mathrm{t}}k_{r_1}) \cdots (k_{r_m} \cdot {}^{\mathrm{t}}k_{r_m}) \cdot k_t \\
= {}^{\mathrm{t}}k_s \cdot \ell_{r_1} \cdots \ell_{r_m} \cdot k_t.
\end{aligned} \tag{20}$$

In Remark 4.17, (20) is further commented and re-expressed for transports in place of kernels. For now, it leads to introduce the following kernel.

Notation 4.6. — Take $R = \{r_1, \dots, r_m\} \subset \mathrm{R}$. We denote the kernel $\ell_{r_1} \cdot \ell_{r_2} \cdots \ell_{r_m}$ from $]0,1[$ to itself by ℓ_R, and $\mathrm{Joint}(\lambda; \ell_R) \in \mathrm{Marg}(\lambda, \lambda)$ by L_R. If $R = \varnothing$, $\ell_\varnothing$ is the identity kernel and $L_\varnothing$ the identity transport.

Notice that $\ell_{\{r\}} = \ell_r$ and that for any R, $\lambda \cdot \ell_R = \lambda$. Moreover, ℓ_R only depends on $(A_r)_{r \in R}$. The following lemma is particularly simple.

Lemma 4.7. — *Let $R \subset \mathrm{R}$ be a finite set and μ, ν be in $\mathrm{M}(\lambda)$.*

(a) *If $\mu \leqslant_{\mathrm{sto}} \nu$ then $\mu \cdot \ell_R \leqslant_{\mathrm{sto}} \nu \cdot \ell_R$, i.e., $\mathrm{Joint}(\lambda; \ell_R)$ has increasing kernel.*

(b) *If μ is in $\mathrm{M}^{\searrow}(\lambda)$, then so is $\mu \cdot \ell_R$, and $\mu \leqslant_{\mathrm{sto}} \mu \cdot \ell_R$.*

Proof. — For $r \in \mathrm{R}$, one easily checks that ℓ_r is increasing and stabilizes $\mathrm{M}^{\searrow}$; (a) and the first point of (b) follow. If $\mu \in \mathrm{M}^{\searrow}(\lambda)$, to see that $\mu \leqslant_{\mathrm{sto}} \mu \cdot \ell_R$, look at the cumulative distribution functions. They coincide off the components of A_r. Now on each of those, $F_{\mu \cdot \ell_R}$ is affine whereas F_μ is concave since μ has decreasing density, so necessarily $F_\mu \geqslant F_{\mu \cdot \ell_R}$.

We add a remark, linked with Figures 2–4, on the principle of Theorem A's proof. A reader only looking for the formal proof itself may skip it.

Remark 4.8. — We will not (directly) obtain the couplings $\mathsf{MQ}^{s,t}$ as a limit of products $\mathsf{Q}^{s,t}_{[R_n]}$, for some finite sets R_n with dense union, as suggested in Section 1.5 p. 11—to show this does not work. We aim at obtaining $\mathsf{MQ}^{s,t}$ as a *supremum*, of the set $\{\mathsf{Q}^{s,t}_{[R]} : R$ finite and $R \subset\,]s,t[\}$ (see Theorem A(iv) for the notation), and actually we do it on the space of quantile levels, i.e., we look for a supremum of $\{\ell_R : R$ finite and $R \subset\,]s,t[\}$. The question is to find the adequate quantity, or order relation, for which a supremum (and hopefully then a maximum) shall be sought.

First, Figure 2 makes us observe how kernels of the type ℓ_R act on measures of $\mathsf{P}(]0,1[)$. It displays the action of ℓ_R with $R = \{r_1, r_2, r_3, r_4\}$ on some Dirac measure δ. The vertical segment on the left is the space $]0,1[$ of quantile levels. We suppose that each μ_{r_i} has a single atom and draw vertically, at abscissa r_i, the interval A_{r_i} (see Notation 4.4(a)). The drawing is in the case where $\delta = \delta_x$ with $x \in A_{r_1}$. Then, see Remark 4.5: $\ell_{\{r_1\}}$ maps δ on the uniform probability measure on A_{r_1}; in turn, $\ell_{\{r_2\}}$ leaves the latter unchanged outside of A_{r_2} and makes it uniform on A_{r_2}, etc. The first drawing shows a "possible trajectory of an element of mass at x" transported by the discrete Markov chain with transition kernels $(\ell_{r_i})_{i=1}^4$. Since we take $x \in A_{r_1}$, it is displaced by ℓ_{r_1} to x', picked uniformly at random in A_{r_1}. In case $x' \notin A_{r_2}$, as in the figure, it is unchanged by ℓ_{r_2}; then in case $x' \in A_{r_3}$ (figure), it is displaced by ℓ_{r_3} to a random $x'' \in A_{r_3}$, and finally, in case $x'' \in A_{r_4}$, displaced by ℓ_{r_4} to a random $x''' \in A_{r_4}$. The second drawing shows the successive measures δ_x, $\delta_x . \ell_{r_1}$, $\delta_x . \ell_{r_1} . \ell_{r_2}$ etc., the level of grey being proportional to the value of their density.

So each ℓ_{r_i} "spreads" a little more the mass of $\delta.\ell_{r_1}.\cdots.\ell_{r_{i-1}}$, replacing it by its mean (measure of constant density) on each connected component of A_{r_i}. If $\theta \ll \lambda$, this averaging process lowers the total variation of the density at each step: at most, you get the measure with constant density one, i.e., λ itself, on which all the transports ℓ_R act trivially. Thus a natural idea is to consider that, if $R' \subsetneq R$ are finite and $\theta \ll \lambda$, the density of $\theta . \ell_R$ will be closer to $\mathbb{1}_{]0,1[}$, for some adequate distance, than that of $\theta . \ell_{R'}$.

Unfortunately, this is the case if $R' = R \cap\,]-\infty, t]$ for some t but not in general. Figure 3 shows, from top to bottom, the Dirac measure δ_x and the graph of the densities of:

- $\delta_x . \ell_{r_1}$, $\delta_x . \ell_{r_1} . \ell_{r_2}$, $\delta_x . \ell_{r_1} . \ell_{r_2} . \ell_{r_3}$, and $\delta_x . \ell_{r_1} . \ell_{r_2} . \ell_{r_3} . \ell_{r_4}$ (on the left),
- $\delta_x . \ell_{r_1}$, $\delta_x . \ell_{r_1} . \ell_{r_2}$, and $\delta_x . \ell_{r_1} . \ell_{r_2} . \ell_{r_4}$ (on the right),

in the case $x = \frac{2}{3}$, $A_{r_1} =]\frac{1}{3}, \frac{5}{6}[$, $A_{r_2} =]\frac{2}{3}, 1[$, $A_{r_3} =]\frac{1}{2}, \frac{5}{6}[$, $A_{r_4} =]0, \frac{2}{3}[$. Then $\delta . \ell_{\{r_1,r_2,r_4\}} = \lambda$, i.e., has exactly density $\mathbb{1}_{]0,1[}$, whereas $\delta . \ell_R \neq \lambda$.

A remedy is to look the kernels ℓ_R act *on measures of density* $\frac{1}{x}\mathbb{1}_{]0,x[}$. For the latter, and more generally any element of $\mathsf{P}^{\ll}(]0,1[)$, which is stable by the action of the couplings ℓ_R, the idea above works, with the stochastic order. This is provided by Lemmas 4.9 and 4.10 below; see also Remark 4.13. Figure 4 gives the example of the

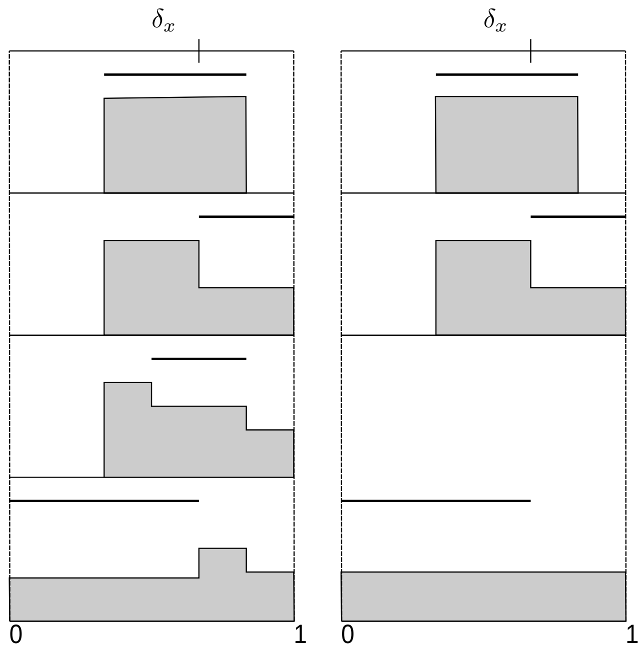

Figure 3. Composition of kernels ℓ_r, acting on some Dirac measure δ_x. On each column, the horizontal interval is the space $]0,1[$ of quantile levels. From top to bottom, the horizontal bars represent successive atomic intervals A_{r_i} and, below each one, the density of δ_x transported by the composition of the successive corresponding kernels $\ell_{r_1}, \ldots, \ell_{r_i}$. On the right, A_{r_3} is omitted.

kernels ℓ_{r_i} of Figure 3 acting on the measure ν of density $\frac{1}{x}\mathbb{1}_{]0,x[}$ with $x = 2/3$: one gets $\nu\,.\,\ell_{r_1}\,.\,\ell_{r_2}\,.\,\ell_{r_3}\,.\,\ell_{r_4}$ $_{\mathrm{sto}}$ $\nu\,.\,\ell_{r_1}\,.\,\ell_{r_2}\,.\,\ell_{r_4}$.

Though simple, the next lemma is a key of our construction of MQ.

Lemma 4.9. — *Let R R^0 be two finite subsets of* R *and μ 2* M$^{\&}(\lambda)$. *Then $\mu.\ell_R$* $_{\mathrm{sto}}$ *$\mu\,.\,\ell_{R^0}$.*

Proof. — Using an induction on the cardinal di erence, it is enough to prove this if R^0 has one more element than R, say r^0. We order it with the elements r_i of R: $r_1 <$ $< r_k < r^0 < r_{k+1} <$ $< r_m$. By Lemma 4.7(b), if μ is in M$^{\&}(\lambda)$, so is $\mu_k := \mu\,.\,\ell_{r_1}\,.$ $.\,\ell_{r_k}$ and μ_k $_{\mathrm{sto}}$ $\mu_k\,.\,\ell_{r^0}$. We apply $\ell_{r_{k+1}}\,.$ $.\,\ell_{r_m}$ to each term of this inequality. Lemma 4.7(a) concludes.

The choice of the suitable set M$^{\&}$ in our monotonicity Lemma 4.9 (see also Remark 4.8) enables us to extend the definition of L_R to infinite sets R. Recall that $_{\mathrm{lo}}$ is interpreted in terms of $_{\mathrm{sto}}$ in Remark 3.19.

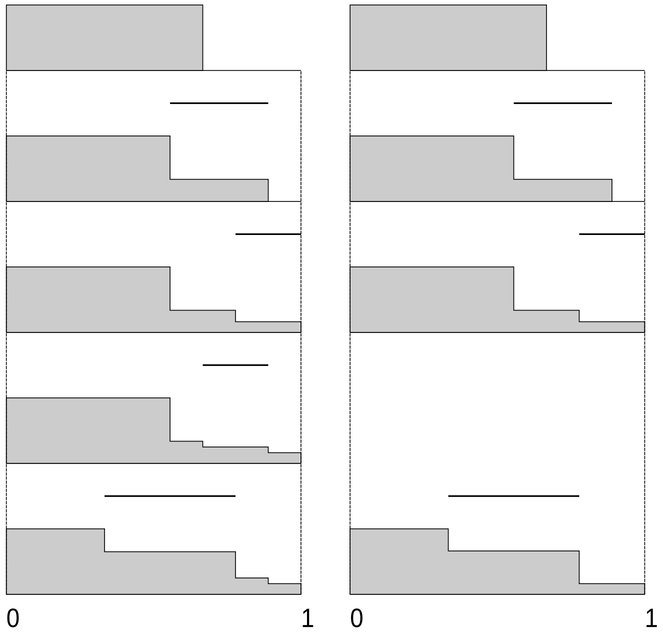

Figure 4. The kernels ℓ_r of Figure 3, acting on the measure of density $\frac{1}{x}\mathbb{1}_{]0,x[}$. The global setup of this figure is the same as that of Figure 3.

Lemma/Notation 4.10

(a) *For all finite subsets R, R' of* $\mathbb{R}$, $L_R \leq_{\mathrm{lo}} \lambda\otimes\lambda$, *and* $R' \subseteq R \Rightarrow L_{R'} \leq_{\mathrm{lo}} L_R$.

(b) *For any* $R \subseteq \mathbb{R}$, $\operatorname{losup}\{L_{R'} : R' \subseteq R$ *and* R' *finite*$\}$ *exists. We denote it by* L_R *(which is consistent with Notation* 4.6 *when* R *is finite).*

(c) *For any* $R \subseteq \mathbb{R}$, *there is a nested sequence* $(R_n)_n$ *of finite subsets of* R *such that* $(L_{R_n})_n$ *converges weakly to* L_R*; if* $(R_n)_n$ *satisfies this property, all sequence* $(R'_n)_n$ *such that* $R'_n \supseteq R_n$ *also does.*

Proof. — Let us prove (a). For all $(x,y) \in [0,1]^2$, $F[L_R](x,y) = (\lambda\llcorner_{[0,x]}).L_R([0,y])$. After Lemma 4.7(b), $(\lambda\llcorner_{[0,x]}).L_R \in \mathcal{M}^{\searrow}$ so $f : y \mapsto \frac{1}{y}((\lambda\llcorner_{[0,x]}).L_R)([0,y])$ is decreasing, thus $f(y) \geq f(1) = x$. Hence:

$$\forall y \in [0,1],\quad ((\lambda\llcorner_{[0,x]}).L_R)([0,y]) \geq xy = F[\lambda\otimes\lambda](x,y),\quad \text{i.e., } L_R \leq_{\mathrm{lo}} \lambda\otimes\lambda.$$

Now if $R' \subseteq R$ for every x we can apply Lemma 4.9 to $\lambda\llcorner_{[0,x]} \in \mathcal{M}^{\searrow}$. This yields $F[L_R](x,y) = (\lambda\llcorner_{[0,x]}).L_R([0,y]) \geq (\lambda\llcorner_{[0,x]}).L_{R'}([0,y]) = F[L_{R'}](x,y)$, which is the expected relation for $\leq_{\mathrm{lo}}$. Then, (b) and (c) follow from criterion (b) of Lemma 3.8. Indeed, by (a), $S = \{L_{R'} : R' \subseteq R$ and R' finite$\}$ is bounded from above by $\lambda\otimes\lambda$, and if $\{L_{R'_1}, L_{R'_2}\} \subseteq S$, then $L_{R'_1\cup R'_2} \in S$ and $L_{R'_1\cup R'_2} \geq_{\mathrm{lo}} L_{R'_i}$ for $i = 1,2$. Finally, the assertion about $(R'_n)_n$ follows from (a) and the interpretation of both ρ and $\leq_{\mathrm{sto}}$ with cumulative distribution functions.

Notation 4.11. — For all $R \subset \mathbb{R}$, we denote by ℓ_R the kernel associated with L_R. This is again consistent with Notation 4.4(b) when R is finite.

Remark 4.12. — For every $\alpha \in\,]0,1[$, due to the definition of $\preceq_{\mathrm{lo}}$,

$$\lambda b_{[0,\alpha]} . \ell_R = \mathrm{stosup}_{T \subset R,\ T \text{ finite}} \lambda b_{[0,\alpha]} . \ell_T .$$

Indeed, Lemma 4.10(c) actually proves that there exists a nested sequence $(R_n)_n$ of finite subsets of $\mathbb{R}$ such that $F[L_{R_n}]$ pointwise converges to $F[L_R]$. Therefore $F[\lambda b_{[0,\alpha]} . \ell_{R_n}](\cdot) = F[R_n](\alpha, \cdot)$ pointwise converges to $F[\lambda b_{[0,\alpha]} . \ell_R]$. Moreover, $F[\lambda b_{[0,\alpha]} . \ell_T] \geq F[\lambda b_{[0,\alpha]} . \ell_R]$ for every finite $T \subset R$. These two facts give the remark.

Remark 4.13 (complement to Remark 4.8). — Lemma 4.10 means that, for $R \subset \mathbb{R}$ and all interval $]a,b[\ \subset\ ]0,1[$, $\big(\lambda b_{]a,b[}\big) . L_R$ is not obtained as a supremum, but as the difference of two:

$$\begin{aligned}\big(\lambda b_{[0,\alpha]}\big) . L_R = \mathrm{stosup}\big\{\big(\lambda b_{[0,\alpha]}\big) . L_{R'} : R' \subset R \text{ and } R' \text{ finite}\big\}\\ - \mathrm{stosup}\big\{\big(\lambda b_{[0,\alpha]}\big) . L_{R'} : R' \subset R \text{ and } R' \text{ finite}\big\}.\end{aligned}$$

The following result is crucial to define processes on $(]0,1[,\lambda)$ with Corollary 2.13, as is done in particular in Definition 4.19 for Lev_R and Lev.

Proposition 4.14. — *If R and R' are subsets of $\mathbb{R}$ such that $r \leq r'$ for all $(r,r') \in R \times R'$, then : $L_{R \cup R'} = L_R . L_{R'}$. In particular, for $s < t < u$, $L_{[s,u]} = L_{[s,t]} . L_{[t,u]}$.*

Proof. — Using Lemma 4.10(b) and (c) we find sequences $(R_n)_n$ and $(R'_n)_n$ of finite subsets of R and R' respectively, such that L_{R_n} converges weakly to L_R, $L_{R'_n}$ to $L_{R'}$, and $L_{R_n \cup R'_n}$ to $L_{R \cup R'}$. Besides, since $r \leq r'$ for all $(r,r') \in R \times R'$, since the R_n are finite and since $L_{\{r\}}$ is idempotent (so that in case $R \cap R' \neq \emptyset$ and $R_n \cap R'_n = R \cap R' = \{r\}$, a repetition of $L_{\{r\}}$ does not matter), $L_{R_n \cup R'_n} = L_{R_n} . L_{R'_n}$. Then, using Proposition 3.27 and the distance ρ introduced in it:

$$\begin{aligned}\rho(L_R . L_{R'}, L_{R\cup R'}) &\leq \rho(L_R . L_{R'}, L_R . L_{R'_n})\\ &\quad + \rho(L_R . L_{R'_n}, L_{R_n} . L_{R'_n}) + \rho(L_{R_n\cup R'_n}, L_{R\cup R'})\\ &\leq \rho(L_{R'}, L_{R'_n}) + \rho(L_R, L_{R_n}) + \rho(L_{R_n\cup R'_n}, L_{R\cup R'})\end{aligned}$$

by Proposition 3.31 and Remark 3.36, since ${}^t L_{R'_n}$ also preserves $\mathrm{M}^{\&}(\lambda)$. All terms tend to zero when n tends to infinity. The desired equality follows.

4.2. The Markov-quantile process MQ; proof of Theorem A

Reminder 4.15. — For $(s,t) \in \mathbb{R}^2$ with $t > s$, and $R = \{r_1, \dots, r_m\} \subset [s,t]$, $\mathrm{Q}^{s,t}_{[R]}$ denotes the coupling $\mathrm{Q}(\mu_s, \mu_{r_1}).\mathrm{Q}(\mu_{r_2}, \mu_{r_3}).\cdots.\mathrm{Q}(\mu_{r_m}, \mu_t) \in \mathrm{Marg}(\mu_s, \mu_t)$, see Theorem A(iv).

Proposition/Definition 4.16

(a) *The set $\{\mathrm{Q}^{s,t}_{[R]} \mid R \subset \mathbb{R}$ and $\sharp R < \infty\}$ has a lower orthant supremum and there is a nested sequence $(R_n)_{n\in\mathbb{N}}$ of finite sets such that $\mathrm{Q}^{s,t}_{[R_n]}$ tends to it. We denote it by $\mu_{s,t}$.*

(b) *The family* $(\mu_{s,t})_{s<t}$ *is consistent in the sense of Definition* 2.14, *giving rise to a Markov measure* $\mathsf{MQ} \in \mathrm{Marg}((\mu_t)_{t\in\mathbb{R}})$, *that we call the* Markov-quantile measure *attached to* $(\mu_t)_{t\in\mathbb{R}}$.

(c) *For all* s *and* $t > s$, $\mathsf{MQ}^{s,t} = (G_s \otimes G_t)_\# L_{]s,t[} = (G_s \otimes G_t)_\# L_{[s,t]}$.

To show Proposition 4.16 we first state the following crucial relation between $\mathsf{Q}^{s,t}_{[R]} \in \mathrm{Marg}(\mu_s, \mu_t)$ and $L_R \in \mathrm{Marg}(\lambda, \lambda)$ defined in Section 4.1.

Remark 4.17. — Take any finite subset R of $\mathbb{R}$ and (s,t) with $s < t$, then:

$$\mathsf{Q}^{s,t}_{[R]} = (G_s \otimes G_t)_\# L_{]s,t[\,\cap R} = (G_s \otimes G_t)_\# L_{[s,t]\,\cap R}.$$

Indeed, $\mathsf{Q}^{s,t} = {}^{\mathrm{t}}q_s \cdot q_t = \mathrm{Law}(G_s, G_t)$, see Remark 4.3(a). It also equals $(G_s \otimes G_t)_\#(\mathrm{Id}_2)$, where $\mathrm{Id}_2 = \mathrm{Id}_{\lambda,2}$ is the identity transport form λ to itself, see Notation 2.6, since by Remark 4.3(b), $\mathsf{Q}(\mu_s, \mu_t) = {}^{\mathrm{t}}q_s \cdot q_t = {}^{\mathrm{t}}q_s . \mathrm{Id}_2 \cdot q_t = (G_s \otimes G_t)_\#(\mathrm{Id}_2)$. More generally, for any $\{r_1, \ldots, r_m\} \subset\,]s,t[$:

$$\begin{aligned}
\mathsf{Q}(\mu_s, \mu_{r_1}) \cdot \cdots \cdot \mathsf{Q}(\mu_{r_m}, \mu_t)
&= {}^{\mathrm{t}}q_s \cdot q_{r_1} \cdot {}^{\mathrm{t}}q_{r_1} \cdot q_{r_2} \cdot \cdots \cdot {}^{\mathrm{t}}q_{r_{m-1}} \cdot q_{r_m} \cdot {}^{\mathrm{t}}q_{r_m} \cdot q_t \\
&= (G_s \otimes G_t)_\# (q_{r_1} \cdot {}^{\mathrm{t}}q_{r_1} \cdot q_{r_2} \cdot \cdots \cdot {}^{\mathrm{t}}q_{r_{m-1}} \cdot q_{r_m} \cdot {}^{\mathrm{t}}q_{r_m}) \\
&\qquad\qquad \text{(notice that this writing involves neither } {}^{\mathrm{t}}q_s \text{ nor } q_t\text{)} \\
&= (G_s \otimes G_t)_\# L_{\{r_1, \ldots, r_m\}}.
\end{aligned}$$

Besides recall that $q_s \cdot {}^{\mathrm{t}}q_s \cdot q_s = q_s$ and ${}^{\mathrm{t}}q_t \cdot q_t \cdot {}^{\mathrm{t}}q_t = {}^{\mathrm{t}}q_t$, so that:

$$\begin{aligned}
\mathsf{Q}(\mu_s, \mu_{r_1}) \cdot \cdots \cdot \mathsf{Q}(\mu_{r_m}, \mu_t) &= (G_s \otimes G_t)_\# (q_s \cdot {}^{\mathrm{t}}q_s \cdot q_{r_1} \cdot \cdots \cdot {}^{\mathrm{t}}q_{r_m} \cdot q_t \cdot {}^{\mathrm{t}}q_t) \\
&= (G_s \otimes G_t)_\# L_{\{s, r_1, \ldots, r_m, t\}}.
\end{aligned}$$

Proof of Proposition 4.16. — We have only to gather our results. We set

$$\mathsf{S} := \{R \subset\,]s,t[\; : R \text{ is finite}\}.$$

By Remark 3.35 and Lemma 4.10(a),

$$\begin{aligned}
&\forall R \in \mathsf{S}, \quad (G_s \otimes G_t)_\# L_R \preceq_{\mathrm{lo}} (G_s \otimes G_t)_\# (\lambda \otimes \lambda), \\
\text{and} \qquad & R' \supset R \implies (G_s \otimes G_t)_\# L_{R'} \preceq_{\mathrm{lo}} (G_s \otimes G_t)_\# L_R.
\end{aligned}$$

Thus, by Lemma 3.8(b), $\mathrm{losup}_{R' \in \mathsf{S}} (G_s \otimes G_t)_\# L_{R'}$ exists and is the limit of some sequence $((G_s \otimes G_t)_\# L_{R_n})_n$. Hence the limit in (a) is given by Remark 4.17:

$$\mathsf{Q}^{s,r_1} \cdot \mathsf{Q}^{r_1,r_2} \cdot \cdots \cdot \mathsf{Q}^{r_{m(n)-1}, r_{m(n)}} \cdot \mathsf{Q}^{r_{m(n)},t} = (G_s \otimes G_t)_\# L_{R_n}.$$

We prove now the first part of (c), i.e., $\mu_{s,t} = (G_s \otimes G_t)_\# L_{]s,t[}$ and, at the same time, that the sequence $(R_n)_n$ in (a) can be chosen to be nested. Let $(R_n)_n$ be a nested sequence of S such that $L_{R_n} \to L_{]s,t[}$, i.e., $F[L_{R_n}]$ pointwise converges to $F[L_{]s,t[}]$. If $M \in \mathrm{Marg}(\mu_s, \mu_t)$ satisfies $(G_s \otimes G_t)_\# L_{R'} \preceq_{\mathrm{lo}} M$ for every $T \in \mathsf{S}$, this also holds for $R' = R_n$ for all n. Now by Remark 3.34, $(G_s \otimes G_t)_\# L_{R_n} \to (G_s \otimes G_t)_\# L_{]s,t[}$.

Therefore, going to the limits at the level of the cumulative distribution functions we get $(G_s \times G_t)_\# L_{]s,t[} \preceq_{\mathrm{lo}} M$. Then Remark 3.35 gives

$$(G_s \times G_t)_\# L_{R'} \preceq_{\mathrm{lo}} (G_s \times G_t)_\# L_{]s,t[},$$

hence:

$$(G_s \times G_t)_\# L_{]s,t[} = \mathrm{lo}\sup_{R' \subset R \text{ and } R' \text{ finite}} \big\{ \mathsf{Q}^{s,t}_{[R']} = (G_s \times G_t)_\# L_{R'} \big\}.$$

Corollary 2.13 gives (b). Indeed, Proposition 4.14 on the composition of transports L_R gives the consistency of $(\mu_{s,t})_{s,t}$ (see Definition 2.14):

$$(G_s \times G_t)_\# L_{]s,t[} = {}^{\mathrm{t}}q_s \cdot L_{]s,t[} \cdot q_t \quad \text{and} \quad (G_t \times G_u)_\# L_{]t,u[} = {}^{\mathrm{t}}q_t L_{]t,u[} q_u.$$

Since $q_t \cdot {}^{\mathrm{t}}q_t = L_{\{t\}}$, $\mu_{s,u} = \mu_{s,t} \cdot \mu_{t,u}$.

For the second equality of (c), proceed as at the end of Remark 4.17.

We now prove Theorem A.

Proof of Theorem A

(a) Recall that MQ is Markov and defined in Definition 4.16. By construction, $\mathsf{MQ} \in \mathrm{Marg}((\mu_t)_{t\in\mathbb{R}})$ and satisfies (iv). Then MQ satisfies (ii), i.e., has increasing kernel as quantile couplings have, see Remark 3.25, and since this property is stable by composition and weak limit, see Remarks 3.14 and Lemma 2.24. The last claim of (iii) reads:

$\mathrm{Law}(X_t | X_s \leq x) = \mathrm{sto}\inf\{\mathrm{Law}(Y_t | Y_s \leq x) : \mathrm{Law}(Y) \in \mathrm{Marg}(\mu)$ satisfies (i) and (ii)$\}$,

where $(X_t)_{t\in\mathbb{R}}$ has law MQ. An alternative writing is that for all $P = \mathrm{Law}(Y)$ as above and all $s < t$, $F[\mathsf{MQ}^{s,t}] \geq F[P^{s,t}]$, i.e., $\mathsf{MQ}^{s,t} \preceq_{\mathrm{lo}} P^{s,t}$. To show it, it is sufficient to show that for any strictly increasing m-tuple $(r_i)_{i=1}^m$:

(21) $$\mathsf{Q}^{r_1,r_2} \cdot \cdots \cdot \mathsf{Q}^{r_{m-1},r_m} \preceq_{\mathrm{lo}} P^{r_1,r_m}.$$

Indeed $\mathsf{MQ}^{s,t} = \mathrm{lo}\sup\{\mathsf{Q}^{r_1,r_2} \cdot \cdots \cdot \mathsf{Q}^{r_{m-1},r_m} : s = r_1 < \cdots < r_m = t\}$, by definition of P in Proposition 4.16. We write (21) in the following equivalent form, in which we $\overline{r}_i$ stands for r_{m+1-i}:

(H_m) $$\underbrace{{}^{\mathrm{t}}\mathsf{Q}^{\overline{r}_2,\overline{r}_1} \cdot \cdots \cdot {}^{\mathrm{t}}\mathsf{Q}^{\overline{r}_m,\overline{r}_{m-1}}}_{\text{denoted by } C \text{ below}} \preceq_{\mathrm{lo}} {}^{\mathrm{t}}P^{\overline{r}_m,\overline{r}_1}.$$

For $A, B \in \mathrm{Marg}(\mu,\nu)$, from the definitions, $A \preceq_{\mathrm{lo}} B$ is equivalent to

$$\forall x, \quad \mu\lfloor_{]-\infty,x]} \cdot A \preceq_{\mathrm{sto}} \mu\lfloor_{]-\infty,x]} \cdot B.$$

Notice also that ${}^{\mathrm{t}}\mathsf{Q}_{\overline{r}_i,\overline{r}_j} = \mathsf{Q}_{\overline{r}_j,\overline{r}_i}$. We prove ($\mathsf{H}_m$) by induction on m. (H_2) is true by definition of Q. Suppose (H_m). Take $x \in \mathbb{R}$. Then (see the justifications below):

$$\mu_{\overline{r}_1}\lfloor_{]-\infty,x]} \cdot \mathsf{Q}^{\overline{r}_1,\overline{r}_2} \cdot \cdots \cdot \mathsf{Q}^{\overline{r}_m,\overline{r}_{m+1}} = \mu_{\overline{r}_1}\lfloor_{]-\infty,x]} \cdot C \cdot \mathsf{Q}^{\overline{r}_k,\overline{r}_{m+1}}$$

(22) $$\preceq_{\mathrm{sto}} \mu_{\overline{r}_1}\lfloor_{]-\infty,x]} \cdot {}^{\mathrm{t}}P^{\overline{r}_m,\overline{r}_1} \cdot \mathsf{Q}^{\overline{r}_m,\overline{r}_{m+1}}$$

(23) $$\preceq_{\mathrm{sto}} \mu_{\overline{r}_1}\lfloor_{]-\infty,x]} \cdot {}^{\mathrm{t}}P^{\overline{r}_m,\overline{r}_1} \cdot {}^{\mathrm{t}}P^{\overline{r}_{m+1},\overline{r}_m}$$

(24) $$\preceq_{\mathrm{sto}} \mu_{\overline{r}_1}\lfloor_{]-\infty,x]} \cdot {}^{\mathrm{t}}P^{\overline{r}_{m+1},\overline{r}_1}.$$

Above, (22) holds since $\mu_{\bar r_1}\lfloor_{]-\infty,x]}\cdot C \preceq_{\mathrm{sto}} \mu_{\bar r_1}\lfloor_{]-\infty,x]}\cdot {}^{\mathrm{t}}P^{\bar r_m,\bar r_1}$ by (H_m) and since $\mathsf{Q}_{\bar r_m,\bar r_{m+1}}$ has increasing kernel, i.e., respects $\preceq_{\mathrm{sto}}$. For (23), since $P^{\bar r_1,\bar r_m}$ has increasing kernel, ${}^{\mathrm{t}}P^{\bar r_1,\bar r_m}$ maps $\mathcal{M}^{\nearrow}(\mu_{\bar r_m})$ on $\mathcal{M}^{\nearrow}(\mu_{\bar r_1})$ by Remark 3.16, so $\mu_{\bar r_1}\lfloor_{]-\infty,x]}\cdot {}^{\mathrm{t}}P^{\bar r_m,\bar r_1}\in \mathcal{M}^{\nearrow}(\mu_{\bar r_m})$, hence it is an increasing limit of positive combinations of measures of the type $\mu_{\bar r_m}\lfloor_{]-\infty,y]}$. Then, for these measures, $\mu_{\bar r_m}\lfloor_{]-\infty,y]}\cdot \mathsf{Q}^{\bar r_m,\bar r_{m+1}} \preceq_{\mathrm{sto}} \mu_{\bar r_m}\lfloor_{]-\infty,y]}\cdot {}^{\mathrm{t}}P^{\bar r_{m+1},\bar r_m}$ by definition of Q. Finally $(Y_t)_t$ is Markov, which gives (24), i.e., (H_{m+1}). We are done.

(b) By Remark 3.24, $(\mu_t)_{t\in\mathbb{R}}$ is increasing for $\preceq_{\mathrm{sto}}$ if and only if every $\mathsf{Q}^{s,t}$ is an increasing coupling, i.e., concentrated on $\{(x,y) : x\leqslant y\}$. This implies the same for their products and the limits of those, so for MQ. Then apply Lemma 1.19.

4.3. Proof of Theorem B: convergence of $\mathsf{Q}_{[R_n]}$ to MQ. — For all finite subset R of $\mathbb{R}$, $P\in \mathrm{Marg}((\mu_t)_{t\in\mathbb{R}})$ and (s,t) with $s<t$ we introduced the couplings $P_{[R]}^{s,t}\in \mathrm{Marg}(\mu_s,\mu_t)$ in Theorem 2.26—and actually in Theorem A(iv) in the case $P=\mathsf{Q}$. We used them in Section 4.2. Now we introduce the measure $P_{[R]}\in\mathrm{Marg}((\mu_t)_t)$ that was announced in Section 1.5 in Notation 1.10. The notation is consistent, i.e., for all $s<t$, $\mathrm{proj}^{s,t}_{\#}P_{[R]}$ is the previously defined $P_{[R]}^{s,t}$. Then we prove Theorem B, which means than we implement the tentative program introduced p. 11 sq. in Section 1.5, in a way that avoids the problems explained there.

Definition 4.18. — If $M\in\mathrm{Marg}((\mu_t)_t)$ and if $R=\{r_1,\dots,r_m\}\subset\mathbb{R}$, we denote by $M_{[R]}\in\mathrm{Marg}((\mu_t)_{\in\mathbb{R}})$ the measure M *made Markov at the points of* R defined by the data of its finite marginals $(\mathrm{proj}^S)_{\#}M_{[R]}$, for all finite S containing R, as follows:

$$(\mathrm{proj}^S)_{\#}M_{[R]} = M^{s_1^0,\dots,s_{n_0}^0,r_1}\otimes M^{r_1,s_1^1,\dots,s_{n_1}^1,r_2}\otimes\cdots\otimes M^{r_m,s_1^m,\dots,s_{n_m}^m},$$

where $S=\{s_1^0,\dots,s_{n_0}^0,r_1,s_1^1,\dots,s_{n_1}^1,r_2,\dots,r_m,s_1^m,\dots,s_{n_m}^m\}$ and where the first or last term disappears if n_0 or n_m is null, respectively. These marginals are consistent in the sense of Definition 2.14. So by Proposition 2.12 this defines $M_{[R]}$. We also commit an abuse of language: $M_{[R]}$ is rather the "law of a process X of law M, made Markov at the points of R".

By Proposition 4.14, for any $R\subset\mathbb{R}$, $(L_{R\setminus[s,t]})_{s<t}\in\mathrm{Marg}(\lambda,\lambda)$ is also a consistent family, thus again Proposition 2.12 enables us to define the following processes on the set of quantile levels.

Definition 4.19. — For all $R\subset\mathbb{R}$ we denote by $\mathrm{Lev}_R\in\mathrm{Marg}((\lambda_t)_{t\in\mathbb{R}})$ (λ_t denotes λ at each t) the Markov process with 2-marginals $\mathrm{Lev}_R^{s,t}=L_{R\setminus[s,t]}$. We call $\mathrm{Lev}_{\mathbb{R}}$ the *level process* attached to $(\mu_t)_{t\in\mathbb{R}}$ and denote it by Lev.

Remark 4.20. — In this subsection, using Definition 4.18 with $M=\mathsf{Q}$ we obtain measures $\mathsf{Q}_{[R]}$ linked with Lev_R as follows. After Remark 4.17, $\mathsf{Q}_{[R]}^{s,t}=(G_s\otimes G_t)_{\#}L_{[s,t]\setminus R}$, thus, with $G=(\bigotimes_{t\in\mathbb{R}}G_t)$, the measures $G_{\#}\mathrm{Lev}_R$ and $\mathsf{Q}_{[R]}$ have the same 2-marginals. They are actually equal, as we prove in the "claim" at the beginning of the proof of Theorem 4.21(b).

The goal of the remaining part of this section is to prove the following statement that is a more precise and technical version of Theorem B. After some preparation its part (a) is proved on p. 47. Its parts (b) and (c) are proved on p. 50 after some more auxiliary results. In the statement below, see Notation 4.10(b) for L, Definition 1.22 for "atomic times" and Definition 4.25 for "essential atomic" intervals or times.

Theorem 4.21. — *Let $(\mu_t)_{t\in\mathbb{R}}$ be a family of probability measures on $\mathbb{R}$. In points* (b) *and* (c)*, $(R_n)_n$ stands for a nested sequence of finite subsets of $\mathbb{R}$ and R for $\bigcup_n R_n$.*

(a) *There is a countable set $R\subset\mathbb{R}$ satisfying:*

$$\text{for all } (s,t) \text{ with } s<t,\ L_{R\cap]s,t[} = L_{]s,t[}. \tag{25}$$

(b) *If R satisfies* (25) *then (see Reminder* 1.11 *for the weak convergence):*

$$(\mathsf{Q}_{[R_n]})_n \text{ converges weakly to } \mathsf{MQ}. \tag{26}$$

(c) *Conversely, if $(R_n)_n$ satisfies* (26)*, then:*

(i) *For any nested finite sets $(R'_n)_n$ such that $R=\bigcup_n R'_n$, $(\mathsf{Q}_{[R'_n]})_n \xrightarrow[n\to\infty]{} \mathsf{MQ}$. In other words,* (26) *is a property of R. Moreover,* (26) *is also satisfied by any countable $R''\supset R$.*

(ii) *Let $E\subset R$ be the set of non-atomic times of R, then $R\smallsetminus E$ satisfies* (26)*. Moreover, for any finite set E' of non-essential atomic times, there is a set R' satisfying* (26) *and such that $R'\cap E'=\emptyset$.*

(iii) *The set R meets each essential atomic interval of $(\mu_t)_t$, hence in particular, it contains all its essential atomic times (which are at most countably many, by Proposition* 4.26*).*

Remark 4.22

(a) We see no simple condition on R that, added to Condition (c)(iii) above, is necessary and sufficient to ensure $\mathsf{Q}_{[R_n]}\to\mathsf{MQ}$ in Theorem 4.21. For example, density in the set of the atomic times is neither necessary, see Example 4.29, nor sufficient: take, e.g., $\mu_t=\lambda\llcorner_{[0,1/2]}+\frac12\delta_1$ if $t\in\mathbb{Q}$ and $\mu_t=\frac12\delta_0+\lambda\llcorner_{[1/2,1]}$ otherwise, then there is no essential atomic interval, all time is atomic, and R suits if and only if $R\cap\mathbb{Q}$ is dense in $\mathbb{Q}$ and $R\cap(\mathbb{R}\smallsetminus\mathbb{Q})$ is dense in $\mathbb{R}\smallsetminus\mathbb{Q}$, so that *any* set dense in $\mathbb{R}$ is not suitable. Even the condition "R is the projection on $\mathbb{R}$ of a set dense in the set of atomic levels $A=\bigcup_{t\in\mathbb{R}}\big(\{t\}\times A_t\big)\subset\mathbb{R}\times[0,1]$" (see Notation 4.4) is not sufficient. Take $\mu_t=\frac12\delta_0+\frac12\delta_1$ if $t\in\mathbb{Q}$, otherwise $\mu_t=\delta_0$, then $E=(\mathbb{Q}\cap([0,1]\smallsetminus\{1/2\}))\times\mathbb{Q}$ is dense in A but $R=\mathrm{proj}_{\mathbb{R}}(E)$ does not suit.

(b) In point (c)(ii) of Theorem 4.21, any non-essential atomic time t may be avoided by a set R satisfying (26), but it is not true that if R satisfies (26), any such $t\in R$ may be removed from R without making (26) false: see Example 4.29 where R consists of a single non-essential atomic time.

(c) Condition (25) implies (26), but notice that it is not necessary. Take, e.g., $\mu_t=\lambda\llcorner_{[0,1]}$ for $t<0$ and $\mu_t=\delta_0$ otherwise. Then R satisfies (25) if and only if $R\cap\mathbb{R}^+\neq\emptyset$, but Q is Markov, so that (26) is true with $R=\emptyset$.

Lemma 4.23. — *Let T denote a totally ordered set of indices (in practice, $\mathsf{T} = \mathbb{R}^+$ or $\mathsf{T} = \mathbb{N}$). If $(R_\tau)_{\tau\in\mathsf{T}}$ is a family of subsets of $\mathbb{R}$, increasing for the inclusion, setting $R = \bigcup_\tau R_\tau$, $(L_{R_\tau})_{\tau\in\mathsf{T}}$ is increasing for $\preceq_{\mathrm{lo}}$ and tends weakly to L_R when τ tends to infinity.*

Proof. — It rests on this remark: by definition of $\preceq_{\mathrm{lo}}$ in Definition 3.4 and of ρ in Proposition 3.27, if A, A', A'' are measures of $\mathcal{P}(\mathbb{R}^d)$ with the same marginals and $A \preceq_{\mathrm{lo}} A' \preceq_{\mathrm{lo}} A''$, then $\rho(A', A'') \leqslant \rho(A, A'')$. Now we prove the lemma. By definition of L_{R_τ} in Lemma 4.10(b) the sequence $(L_{R_\tau})_{\tau\in\mathsf{T}}$ is increasing, be the sets R_τ finite or not. By Lemma 4.10(c), and as ρ metrizes the weak topology (Proposition 3.27), for all $\varepsilon > 0$ we find a finite $R' \subset R$ such that $\rho(L_R, L_{R'}) \leqslant \varepsilon$. As $R = \bigcup_\tau R_\tau$ there is a τ_0 such that $R_{\tau_0} \supset R'$. Then: $\tau \geqslant \tau_0 \Rightarrow L_{R'} \preceq_{\mathrm{lo}} L_{R_\tau} \preceq_{\mathrm{lo}} L_R \Rightarrow \rho(L_{R'}, L_{R_\tau}) \leqslant \varepsilon$ by the remark.

Lemma 4.24. — *There is a nested sequence $(R_n)_{n\in\mathbb{N}}$ of finite subsets of $\mathbb{R}$ such that $s < t \Rightarrow L_{R_n\cap]s,t[} \to L_{]s,t[}$*

Proof. — Let $((u_k, u'_k))_{k\in\mathbb{N}}$ be a dense sequence in

$$\{(x, y) \in \mathbb{R}^2 : x < y\}$$

and for all $n \geqslant 1$, $(R^n_k)_{k\in\mathbb{N}}$ a nested sequence of finite subsets of $\mathbb{R}$ such that $\rho(L_{R^n_k\cap]u_k,u'_k[}, L_{]u_k,u'_k[}) \leqslant 1/n$ for every $k \in \mathbb{N}$. Denote $\bigcup_{k=1}^n R^n_k$ by R_n. Then for any s and $t > s$, $L_{R_n\cap]s,t[} \xrightarrow[n\to\infty]{} L_{]s,t[}$. Let us prove it. Fix $\varepsilon > 0$. By Lemma 4.23 (use a subsequence $(k_n)_n$ such that $(]u_{k_n}, u'_{k_n}[)_n$ is an exhaustion of $]s, t[$), there exists k such that $\rho(L_{]s,t[}, L_{]u_k,u'_k[}) < \varepsilon/2$. Then $L_{]u_k,u'_k[\cap R_n} \preceq_{\mathrm{lo}} L' \preceq_{\mathrm{lo}} L_{]s,t[}$, where L' stands for $L_{]s,t[\cap R_n}$ or $L_{]u_k,u'_k[}$ for any k. For $n \geqslant \max\{k, 2/\varepsilon\}$:

$$\rho(L_{]s,t[}, L_{]u_k,u'_k[\cap R_n}) \leqslant \rho(L_{]s,t[}, L_{]u_k,u'_k[}) + \rho(L_{]u_k,u'_k[}, L_{]u_k,u'_k[\cap R_n}) < \varepsilon.$$

Therefore with $L' = L_{]s,t[\cap R_n}$ we obtain the desired convergence to $L_{]s,t[}$.

Definition 4.25. — Let I be an interval. If, for some interval $J \supset I$ such that $J \smallsetminus I$ is disconnected (then for all such smaller J'), $L_J \neq L_{J\smallsetminus I}$, we call I an *essential atomic interval of* $(\mu_t)_{t\in\mathbb{R}}$. If $I = \{t\}$ is essential, we call t an *essential atomic time of* $(\mu_t)_{t\in\mathbb{R}}$.

To check the parenthesis in the definition, suppose that $L_{J'} = L_{J'\smallsetminus I}$ and get $L_J = L_{J\smallsetminus I}$ by Proposition 4.14.

Proposition 4.26. — *If a nested sequence $(R_n)_n$ is as in Lemma 4.24 then $\bigcup_n R_n$ contains all the essential atomic times of $(\mu_t)_t$. In particular, these times are at most countably many.*

Proof. — We show a contrapositive result. Suppose that t is an essential atomic time, that $s < t$ and $s' > t$ are such that $L_{]s,s'[\smallsetminus\{t\}} \neq L_{]s,s'[}$, and that $(R_n)_n$ is a nested sequence of finite sets such that $t \notin \bigcup_n R_n$, that $L_{R_n \cap]s,t[} \to L_{]s,t[}$ and that $L_{R_n \cap]t,s'[} \to L_{]t,s'[}$. Then:

$$\begin{aligned} L_{R_n \cap]s,s'[} &= L_{R_n \cap]s,t[} . L_{R_n \cap]s,t[} && \text{by Proposition 4.14} \\ &\to L_{]s,t[} . L_{]t,s'[} && \text{by assumption and by Proposition 3.38} \\ &= L_{]s,s'[\setminus \{t\}} \neq L_{]s,s'[} && \text{by Proposition 4.14 and the assumption.} \end{aligned}$$

Therefore, $(R_n)_n$ cannot be as in Lemma 4.24.

Remark 4.27. — Be careful, the property to be an essential atomic interval is true in general neither for a union of two such (intersecting) intervals, nor for their intersection, nor for an interval containing a such interval or included in it.

Remark 4.28 and Example 4.29 are qualitative comments on the notions introduced in Definition 4.25. They are independent of the proof of Theorem 4.21 and may be skipped in a first reading.

Remark 4.28

(a) Essential atomic times are of course atomic. Indeed, else, $L_{\{t\}}$ is the identity transport so $L_I . L_{\{t\}} . L_{I'} = L_I . L_{I'}$ for all intervals with $\sup I \leq t \leq \inf I'$ and t cannot be essential.

(b) Suppose, to simplify, that some $\mu \in \mathrm{P}(\mathbb{R})$ has exactly one atom x and consider a family $(\mu_t)_{t\in\mathbb{R}}$ such that $\mu = \mu_0$. An obvious sufficient condition for 0 to be unessential is to choose μ_t such that for a certain sequence $t_n \to 0$, μ_{t_n} has an atom x_n such that $A_{t_n,x_n} \supset A_{t,x}$ (see Notation 4.4), or even only

$$\forall \varepsilon > 0, \exists n_0 : \quad n \geq n_0 \implies A_{t_n,x_n} +]-\varepsilon,\varepsilon[\supset A_{t,x} :$$

"atoms merging the same levels of quantile as x merges at $t = 0$, accumulate on 0". Indeed, taking possibly a subsequence, we may suppose that $(t_n)_n$ tends to zero from the right or the left (say, from the right). For any $s > t$, the function $x \mapsto k_{L_{[t_n,s[}}(x,\cdot)$ (when it is defined i.e., $t_n < s$) is constant on A_{t_n,x_n}, hence since $L_{]0,s[} = \lim L_{[t_n,s[}$ (see Lemma 4.23), $x \mapsto k_{L_{]0,s[}}(x,\cdot)$ is constant on $A_{t,x}$ (use Proposition 3.27). Considering $k_{L_{\{0\}}}$ described in Remark 4.3(c), the composition formula in Section 2.1.1 yields $L_{]0,s[} = L_{\{0\}} . L_{]0,s[}$; hence $L_{]s',s[\setminus\{t\}} = L_{]s',s[}$ for all $s' < t$.

If μ_0 has several atoms $(x_i)_{i\in I}$, a similar statement can be shown, the condition being that each of the intervals A_{0,x_i} has the property above.

(c) A necessary condition for an atomic time t to be unessential is immediate: $\{t\} \times A_t$ shall be included in $E = \overline{A_{]-\infty,t[}} \cap \overline{A_{]t,+\infty[}}$, where $A_I := \bigcup_{r\in I}(\{r\} \times A_r)$. There is some nonempty open interval J such that $\overline{J} \subset A_t \setminus E$. Then for ε small enough, $L_{]t-\varepsilon,t+\varepsilon[}$ restricted to J^2 is the identity transport, which prevents t from being essential (see Remark 4.34).

(d) The condition of point (b) is not necessary, nor that of point (c) sufficient. For instance take $\mu_0 = \delta_0$ and for $t \neq 0$, $(\mu_t)_t = a(t)\delta_0 + (1 - a(t))\delta_1$. If a has unbounded total variation on any interval $]0,r[$, then any $L_{]0,r[}$ is the uniform measure on $[0,1]^2$, see Example 5.9, thus 0 is not essential (in fact, even not right-essential in the sense given in Remark 4.34). On the contrary if a has bounded total variation, none of the measures $L_{]0,r[}$ or $L_{]r,0[}$ is uniform, hence 0 is essential (see Remark 4.34).

Example 4.29. — A family $(\mu_t)_{t\in\mathbb{R}}$ may have an essential atomic interval and no essential atomic time. Let I be any interval (but not a singleton) and take, e.g., $\mu_t = \delta_0$ for $t \in I$ and $\mu_t = \lambda_{[0,1]}$ otherwise, then I is the only essential atomic interval of $(\mu_t)_t$. Here, $\mathsf{Q}_{[\{t_0\}]} = \mathsf{MQ}$ for any $t_0 \in I$.

We can now prove Theorem 4.21(a).

Proof of Theorem 4.21(a). — Take $(R_n)_n$ a sequence as given by Lemma 4.24. Then, for any s and $t > s$:

$$
\begin{aligned}
L_{R\cap]s,t[} &= \mathrm{losup}\{L_{R_n\cap]s,t[}\} && \text{by Lemma 4.23}\\
&\leqslant_{\mathrm{lo}} \mathrm{losup}\{L_{R'_n\cap]s,t[}\} && \text{as } R'_n \supset R_n \text{ and by Lemma 4.10(a)}\\
&= \lim_{n\to\infty} L_{R'_n\cap]s,t[} = L_{]s,t[} && \text{by Lemma 4.23 and}\\
& && \text{by property of the sets } R'_n.
\end{aligned}
$$

But by definition, $L_{R\cap]s,t[} \geqslant_{\mathrm{lo}} L_{]s,t[}$, thus $L_{R\cap]s,t[} = L_{]s,t[}$.

To provide a clear proof of Theorem 4.21(b) we introduce an auxiliary notion in Definition 4.30 below. Suppose that some $P \in \mathrm{Marg}(\mu, \mu_1, \dots, \mu_k)$ is disintegrated as $P = \mathrm{Joint}(\mu, k_P)$ and that $g : \mathbb{R} \to \mathbb{R}$ and $h : \mathbb{R}^k \to \mathbb{R}^k$ are measurable maps. May we disintegrate $(g \times h)_\# P$? In case g is into one easily checks $(g \times \mathrm{Id})_\# P = \mathrm{Joint}(g_\# \mu, k_P^g)$, where k_P^g is defined by $k_P^g(y, \cdot) = k_P^g(g^{-1}(y), \cdot)$, $g_\#\mu$-almost surely. Otherwise, the next notion and lemma will enable us to obtain a similar disintegration, and associated properties.

Definition 4.30. — We say that $g : \mathbb{R} \to \mathbb{R}$ *fits* $P \in \mathrm{Marg}(\mu, \mu_1, \dots, \mu_k)$ if there exists a kernel k_P^g such that $k_P^g(g(x), \cdot) = k_P(x, \cdot)$, μ-almost surely.

Remark 4.31

(a) If g fits $P = \mathrm{Joint}(\mu, k)$ and h is a measurable map from $\mathbb{R}^k$ into itself we can disintegrate $(g \times h)_\# P$ as $\mathrm{Joint}(g_\#\mu, h_\# k_P^g)$, where $h_\# k_P^g(y, \cdot) = k_P^g(y, h^{-1}(\cdot))$.

(b) If g fits P, it also fits $P \,.\, P'$ and $P \circ P'$.

(c) If g fits $P \in \mathrm{Marg}(\mu, \nu_1, \dots, \nu_k)$ or $Q \in \mathrm{Marg}(\mu, \nu'_1, \dots, \nu'_{k'})$, and if $f : \mathbb{R}^k \to \mathbb{R}^k$ and $h : \mathbb{R}^{k'} \to \mathbb{R}^{k'}$ are measurable maps, then:

$$(f \times g \times h)_\# ({}^tP \circ Q) = (f \times g)_\# {}^tP \circ (g \times h)_\# Q.$$

The proofs are direct, using the definitions of a kernel, composition and concatenation in Section 2.1.4. Notice that, in case $\mu = \lambda_{[0,1]}$ and g is a quantile function (which is the only case in which we will use the remark), point (c) is a particular case of Lemma 4.33(b) below. In the language of this lemma, "g fits ${}^tP$ or Q" means that ${}^tP \,.\, L = {}^tP$ or $L \,.\, Q = Q$, which both imply in particular that ${}^tP \,.\, L \,.\, Q = {}^tP \,.\, Q$.

We will need the following little technical result.

Lemma 4.32. — *Take $a<b$ in $\mathbb{R}$ and f and g two positive bounded increasing functions from $]a,b[$ to $\mathbb{R}$. Then*

$$\frac{1}{b-a}\int_a^b fg\,\mathrm{d}\lambda \geqslant \Big(\frac{1}{b-a}\int_a^b f\mathrm{d}\lambda\Big)\Big(\frac{1}{b-a}\int_a^b g\mathrm{d}\lambda\Big),$$

with equality if and only if f or g is constant.

Proof. — The measures $\mu = f\lambda$ and $\nu = \frac{1}{b-a}\int_a^b f\,\mathrm{d}\lambda$ have the same mass $\int_a^b f\,\mathrm{d}\lambda$. Besides, $\nu \preceq_{\mathrm{sto}} \mu$; indeed, for any $c \in\,]a,b[$, since f is increasing:

$$\frac{1}{c-a}\mu(]-\infty,c]) = \frac{1}{c-a}\int_a^c f\,\mathrm{d}\lambda \leqslant \frac{1}{b-a}\int_a^b f\,\mathrm{d}\lambda = \frac{1}{c-a}\nu(]-\infty,c]).$$

Hence, as g is increasing, $\int g\,\mathrm{d}\mu \geqslant \int g\,\mathrm{d}\nu$, equivalent to the wished inequality.

Plainly, equality occurs if f or g is constant. If f is not constant, we in fact proved that equality in $\int g_c\mathrm{d}\mu \leqslant \int g_c\mathrm{d}\nu$ holds for $g_c = \mathbb{1}_{[c,b]}$ if and only if $c \in \{a,b\}$. This remains true for positive combinations of functions g_c and, being a little careful, for limits of them.

Lemma 4.33. — *Take $\mu\in\mathrm{P}(\mathbb{R})$, denote by $g=G_\mu$ its quantile function and $F=F_\mu$ its cumulative distribution function. Take $L = \mathrm{Q}(\lambda,\mu).\mathrm{Q}(\mu,\lambda)$ and, similarly as in Notation 4.4, let $(b_i)_{i\in I}$ be the atoms of μ and for each i, A_i be the interval $]F(b_i^-),F(b_i)[$ of quantile levels merged by g on b_i. Finally set $A = \bigcup_{i\in I} A_i$. Moreover, take*

$$((\nu_i)_i,(\nu'_i)_i) \in \mathrm{P}(\mathbb{R})^{k+k'},$$
$$P \in \mathrm{Marg}(\lambda_{[0,1]},\nu_1,\dots,\nu_k) \quad\textit{and}\quad Q\in\mathrm{Marg}(\lambda_{[0,1]},\nu'_1,\dots,\nu'_{k'}).$$

Suppose that P and Q have increasing kernel (for $\preceq_{\mathrm{lo}}$ in place of $\preceq_{\mathrm{sto}}$ if $k\geqslant 2$ or $k'\geqslant 2$). Then:

(a) *${}^tP.L.Q$ equals ${}^tP.Q$ if and only if for each $i\in I$, at least one of the two kernel functions $x\mapsto k_P(x,\cdot)$ or $x\mapsto k_Q(x,\cdot)$ is constant on A_i,*

(b) *For any measurable maps $f:\mathbb{R}^k\to\mathbb{R}^k$ and $h:\mathbb{R}^{k'}\to\mathbb{R}^{k'}$,*

$${}^tP.L.Q = {}^tP.Q \implies (f\times g\times h)_\#({}^tP\circ Q) = (f\times g)_\#{}^tP\circ(g\times h)_\#Q.$$

Proof. — We recall particularly here that we have throughout in mind the analogy between composition of transport plans, or of their kernels, and product of matrices, hinted at in Remark 2.3. In the following, $B_1\subset\mathbb{R}^k$ and $B_3\subset\mathbb{R}^{k'}$ stand for any sets of the type $\prod_j]-\infty,d_j]$ and B_2 for any interval $]-\infty,b]$; (a,b,c) stand for variables in the target of (f,g,h), and (x,y,z) for variables in their sources. The coupling L, equal to ${}^tq_r\,.\,q_r$ (with $\mu=\mu_r$) introduced in Remark 4.3, is described in this remark. It is such that ${}^tP.L.Q = ({}^tP.L).(L.Q) = {}^t((g\times\mathrm{Id}_k)_\#P).((g\times\mathrm{Id}_{k'})_\#Q)$. In turn, for any measurable functions $f:\mathbb{R}^k\to\mathbb{R}^k$ and $h:\mathbb{R}^{k'}\to\mathbb{R}^{k'}$:

$$
\begin{aligned}
&\int_{B_2} k_{(g\circ f)_\# P}(b,B_1)\cdot k_{(g\circ h)_\# Q}(b,B_3)\,\mathrm{d}\mu(b)\\
&\qquad=\sum_{i\in I,\,b_i\in B_2}\mu(b_i)[k_{(g\circ f)_\# P}(b_i,B_1)\cdot k_{(g\circ h)_\# Q}(b_i,B_3)]\\
&\qquad\quad+\int_{B_2\setminus\bigcup_i\{b_i\}} k_P(g^{-1}(b),f^{-1}(B_1))\cdot k_Q(g^{-1}(b),h^{-1}(B_3))\,\mathrm{d}\mu(b)\\
&\hspace{20em}\text{as } g \text{ is injective outside of } A\\
&\qquad=\sum_{i\in I,\,b_i\in B_2}\frac{1}{\lambda(A_i)}\int_{A_i}k_P(y,f^{-1}(B_1))\,\mathrm{d}\lambda(y)\int_{A_i}k_Q(y,h^{-1}(B_3))\,\mathrm{d}\lambda(y)\\
&\qquad\quad+\int_{g^{-1}(B_2)\setminus A}k_P(y,f^{-1}(B_1))\cdot k_Q(y,h^{-1}(B_3))\,\mathrm{d}\lambda(y).
\end{aligned}
\tag{27}
$$

Then, using for instance the expression (6) given in Definition 2.8 of the concatenation, and the fact that for any transport plans R and R', $R\cdot R'=\mathrm{proj}^{1,3}_\#(R\otimes R')$ we get:

$$
\begin{aligned}
&({}^tP\cdot Q-{}^t((g\circ\mathrm{Id})_\# P)\cdot(g\circ\mathrm{Id})_\# Q)(B_1\times B_3)\\
&\qquad=\int_{[0,1]}k_P(y,B_1)k_Q(y,B_3)\mathrm{d}y-\int_{\mathbb{R}}k_{(g\circ\mathrm{Id})_\# P}(b,B_1)k_{(g\circ\mathrm{Id})_\# Q}(b,B_3)\mathrm{d}\mu(b)\\
&\qquad=\sum_{i\in I}\int_{A_i}k_P(y,B_1)k_Q(y,B_3)\mathrm{d}y-\frac{1}{\lambda(A_i)}\int_{A_i}k_P(y,B_1)\mathrm{d}y\int_{A_i}k_Q(y,B_3)\mathrm{d}y
\end{aligned}
$$

by (27) with $B_2=\mathbb{R}$, $f=\mathrm{Id}$, $h=\mathrm{Id}$—notice in particular that the two terms of the difference, viewed as integrals on $[0,1]$, differ possibly only on $A\subset[0,1]$. Now, after Lemma 4.32 applied on each A_i with $f_i:y\mapsto k_P(y,B_1)$ and $g_i:y\mapsto k_Q(y,B_3)$, which are increasing since P and Q have increasing kernel, each term of the sum is non-negative. Thus, equality holds if and only if each vanishes, which again by Lemma 4.32 means that for each i, f_i or g_i is constant. This, holding for any B_1 and B_3, means point (a). For point (b):

$$
\begin{aligned}
&{}^t((g\circ f)_\# P)\otimes((g\circ h)_\# Q)(B_1\times B_2\times B_3)\\
&\qquad=\int_{B_2}k_{(f\circ g)_\# P}(b,B_1)k_{(g\circ h)_\# Q}(b,B_3)\,\mathrm{d}\mu(b)\quad\text{by (6)}\\
&\qquad=\int_{g^{-1}(B_2)\setminus A}k_P(y,f^{-1}(B_1))k_Q(y,h^{-1}(B_3))\mathrm{d}y\\
&\qquad\quad+\sum_{i\in I,\,b_i\in B_2}\frac{1}{\lambda(A_i)}\int_{A_i}k_P(y,f^{-1}(B_1))\mathrm{d}y\int_{A_i}k_Q(y,h^{-1}(B_3))\mathrm{d}y\quad\text{by (27),}
\end{aligned}
$$

but on each A_i, by (a), $k_P(\,.\,,f^{-1}(B_1))$ or $k_Q(\,.\,,h^{-1}(B_3))$ is a constant function, so the last term may be written:

$$
\sum_{i\in I,\,b_i\in B_2}\int_{A_i}k_P(y,f^{-1}(B_1))k_Q(y,h^{-1}(B_3))\mathrm{d}y,
$$

and therefore the initial expression equals $\int_{g^{-1}(B_2)}k_P(y,f^{-1}(B_1))k_Q(y,h^{-1}(B_3))\mathrm{d}y$, that is to say $(f\times g\times h)_\#({}^tP\otimes Q)(B_1\times B_2\times B_3)$.

Remark 4.34. — Lemma 4.33(a) provides a more precise characterization of essential atomic times: t is such a time if and only if μ_t has at least one atom $b \in \mathbb{R}$ that is both "left-essential" and "right-essential", as follows.

– The atom b is said to be left-essential if there is some $\varepsilon > 0$ such that if $s \in]t-\varepsilon, t[$, the kernel $x \mapsto {}^t\ell_{]s,t[}(x,\cdot)$ (see Notation 4.4) is not constant on $A_{t,b}$,

– the atom b is said to be right-essential if there is some $\varepsilon > 0$ such that if $s \in]t, t+\varepsilon[$, the kernel $x \mapsto \ell_{]t,s[}(x,\cdot)$ is not constant on $A_{t,b}$,

Let us call "left-" or "right-essential" an atomic time t such that μ_t has at least one left, respectively right essential atom. Left- or right-essential atomic times must not be essential. However, they are also at most countably many. Here is a sketch of proof, written for right-essential times: show that any such time is a discontinuity point of some function $\varphi_s :]-\infty, s] \ni u \mapsto L_{[u,s]}$. Now if some φ_s is discontinuous at t, all $\varphi_{s'}$ are, for $s' \in]t, s]$ (use Remark 3.36). Hence, the union of the sets of discontinuity points of the functions φ_s, for $s \in \mathbb{R}$, is the same as their union for $s \in \mathbb{Q}$. Finally the claim below, left to the reader, implies that for each φ_s, this set is at most countable, which gives the result.

Claim. — *If $\psi : u \mapsto M_u \in \mathrm{Marg}(\lambda, \lambda)$ is increasing or decreasing for $\leq_{\mathrm{lo}}$, it has at most countably many (weak) discontinuity points.*

Now we can prove the end of Theorem 4.21, i.e., its parts (b) and (c).

Proof of Theorem 4.21(b)–(c). — Let us prove point (b). Take $R \subset \mathbb{R}$ and Lev_R and Lev in $\mathrm{Marg}((\lambda_t)_{t\in\mathbb{R}})$ given by Definition 4.19. We need the following claim, that extends Remark 4.20, and relies eventually on Remark 4.3(b):

Claim. — *If R is finite, $G_\# \mathrm{Lev}_R = \mathrm{Q}_{[R]}$, where $G = (\otimes_{t\in\mathbb{R}} G_t)$.*

Let us prove it. Take $R \subset \mathbb{R}$ finite and n its cardinal. We must prove that for any finite $S = \bigcup_{i=1}^N \{s_i\}$, $G_\# \mathrm{Lev}_R^{s_1,\dots,s_N} = \mathrm{Q}_{[R]}^{s_1,\dots,s_N}$, where, by an abuse of notation we will often make use of, G stands for $\otimes_{i=1}^N G_{s_i}$. It suffices to prove it in the case $S \supset R$, which we suppose now. We introduce the cardinals $\ell_0, \dots, \ell_n$ of the subsets of $S \setminus R$ situated between to consecutive elements r_k and r_{k+1} of $R \cup \{\pm\infty\}$. We re-index these subsets as $\{s_1^k, \dots, s_{\ell_k}^k\}$. Since Lev_R is Markov, $(\mathrm{proj}^S)_\# \mathrm{Lev}_R = B_0 \otimes \cdots \otimes B_m$, where:

$$B_k = L_{r_k} \otimes \mathrm{Id}_{\ell_k} \otimes L_{r_{k+1}} = \begin{cases} \mathrm{Id}_{\ell_0} \otimes L_{r_1} & \text{if } k = 0 \\ L_{r_m} \otimes \mathrm{Id}_{\ell_m} & \text{if } k = m \\ L_{[r_k, r_{k+1}]} = L_{r_k}.L_{r_{k+1}} & \text{if } \ell_k = 0 \\ L_{[r_k, r_{k+1}]} = L_{r_k} \otimes L_{r_{k+1}} & \text{if } \ell_k = 1 \\ L_{r_k} \otimes \underbrace{\mathrm{Id}_2 \otimes \cdots \otimes \mathrm{Id}_2}_{\ell_k - 1} \otimes L_{r_{k+1}} & \text{otherwise,} \\ & \quad \text{i.e., if } \ell_k \geq 2, \end{cases}$$

hence we must prove that: $G_\# (B_1 \otimes \cdots \otimes B_m) = \mathrm{Q}_{[R]}^{s_1,\dots,s_N}$. For all $r \in R$, the quantile function $G_r = G_{\mu_r}$ fits $q_r = \mathrm{Joint}(\mu_r, x \mapsto \delta_{G_r(x)})$, so we may take $k_{q_r}^{G_r} : x \mapsto \delta_x$ as given by Definition 4.30. Now, by Remark 4.31(b), it also fits $L_{\{r\}} = {}^{\mathrm{t}}L_{\{r\}} = q_r \cdot {}^{\mathrm{t}}q_r$. Therefore, by Remark 4.31(b) and (c), $G_\#(B_1 \otimes \cdots \otimes B_m) = B'_1 \otimes \cdots \otimes B'_m$, where:

$$B'_k = G_\# B_k = \begin{cases} (\otimes_{s\leqslant r_1} G_s)_\# \mathrm{Id}_{\ell_0} \otimes L_{r_1} & \text{if } k=0\\ (\otimes_{s\geqslant r_m} G_s)_\# L_{r_m} \otimes \mathrm{Id}_{\ell_m} & \text{if } k=m\\ (G_{r_k}\otimes G_{r_{k+1}})_\#(L_{r_k}\cdot L_{r_{k+1}}) & \text{if } \ell_k = 0\\ (G_{r_k}\otimes G_{s_1^k}\otimes G_{r_{k+1}})_\#(L_{r_k}\otimes L_{r_{k+1}}) & \text{if } \ell_k = 1\\ (\otimes_{r_k\leqslant s\leqslant r_{k+1}} G_s)_\#(L_{r_k}\otimes \mathrm{Id}_{\ell_k}\otimes L_{r_{k+1}}) & \text{otherwise, i.e., } \ell_k \geqslant 2. \end{cases}$$

Proving that $B'_k = \mathrm{Q}(\mu_{r_k}, \mu_{s_1^k}, \dots, \mu_{s^k_{\ell_k-1}}, \mu_{r_{k+1}})$ will now prove the claim. For simplicity we assume $k \notin \{0, m\}$ and $\ell_k \geqslant 1$ but the other cases, which are simpler, can be proved similarly. Observe that:

– by definition of Q and G,

$$\mathrm{Q}(\mu_{r_k}, \mu_{s_1^k}, \dots, \mu_{s^k_{\ell_k}}, \mu_{r_{k+1}}) = \mathrm{Law}((G_{r_k}(U), \dots, G_{r_{k+1}}(U)),$$

where U is a variable of law λ on $[0,1]$,

– $(\otimes_{r_k\leqslant s\leqslant r_{k+1}} G_s)_\#(L_{r_k}\otimes \mathrm{Id}_{\ell_k}\otimes L_{r_{k+1}})$ is the law of

$$(G_{r_k}(U_1), G_{s_1^k}(U_2), \dots, G_{s^k_{\ell_k}}(U_2), G_{r_{k+1}}(U_3)),$$

where $\mathrm{Law}(U_1, U_2, \dots, U_2, U_3) = B_k$. In particular, $\mathrm{Law}(U_i) = \lambda$ for $i = 1, 2, 3$, $\mathrm{Law}(U_1, U_2) = L_{r_k}$, and $\mathrm{Law}(U_2, U_3) = L_{r_{k+1}}$.

So only the first and last variables may differ. Now, notice that by Remark 4.3(b) applied to $T = L_{\{r\}} = q_r \cdot {}^{\mathrm{t}}q_r$, if two variables (U, V) satisfy $\mathrm{Law}(U, V) = L_{\{r\}}$, the law of $(G_r(U), G_r(V))$ is

$${}^{\mathrm{t}}q_r \cdot L_{\{r\}} \cdot q_r = {}^{\mathrm{t}}q_r \cdot q_r \cdot {}^{\mathrm{t}}q_r \cdot q_r = {}^{\mathrm{t}}q_r \cdot q_r = \mathrm{Id}_{\mu_r,2},$$

therefore $G_r(U) = G_r(V)$ almost surely. Using this with $r = r_k$ and $(U, V) = (U_1, U_2)$, respectively $r = r_{k+1}$ and $(U, V) = (U_2, U_3)$, we get respectively $G_{r_k}(U_1) = G_{r_k}(U_2)$ and $G_{r_{k+1}}(U_2) = G_{r_{k+1}}(U_3)$ almost surely. The claim is proved.

Now take $R \subset \mathbb{R}$ satisfying (25) and $(R_n)_n$ a nested exhaustion of it by finite sets. Let $\widetilde{\mathrm{Lev}} \in \mathrm{Marg}((\lambda_t)_{t\in\mathbb{R}})$ be the Markov process having $\widetilde{\mathrm{Lev}}^{s,t} = L_{R\cap[s,t]}$ as 2-marginals. Recall that $L_{R\cap[s,t]} = L_{R\cap\{s\}} \cdot L_{R\cap]s,t[} \cdot L_{R\cap\{t\}} = L_{R\cap\{s\}} \cdot L_{]s,t[} \cdot L_{R\cap\{t\}} = \lim_{n\to\infty} \mathrm{Lev}_{R_n\cap[s,t]} = \lim_{n\to\infty} \mathrm{Lev}^{s,t}_{R_n}$. It exists by Corollary 2.13, whose assumption is satisfied by Proposition 4.14. All these transport plans have increasing kernel, hence by Lemma 2.23, for every $S = \bigcup_{i=1}^k \{s_i\}$:

$$\begin{aligned}(28)\quad \mathrm{Lev}^{s_1,\dots,s_k}_{R_n} = \mathrm{Lev}^{s_1,s_2}_{R_n} \circ \cdots \circ \mathrm{Lev}^{s_{k-1},s_k}_{R_n} \\ \to \widetilde{\mathrm{Lev}}^{s_1,s_2} \circ \cdots \circ \widetilde{\mathrm{Lev}}^{s_{k-1},s_k} = \widetilde{\mathrm{Lev}}^{s_1,\dots,s_k}.\end{aligned}$$

Thus Lev_{R_n} converges weakly to $\widetilde{\mathrm{Lev}}$, and then by Remark 3.34, $G_\# \mathrm{Lev}_{R_n} \to G_\# \widetilde{\mathrm{Lev}}$. We are left with the tasks to prove $G_\# \mathrm{Lev}_{R_n} = \mathrm{Q}_{[R_n]}$ and $G_\# \widetilde{\mathrm{Lev}} = \mathrm{MQ}$. The former is our claim above. Let us prove the latter.

Note. — At the beginning of Section 4 we announced (19), i.e., $G_\#\mathrm{Lev} = \mathsf{MQ}$. In fact, we prove $G_\#\widetilde{\mathrm{Lev}} = \mathsf{MQ}$, which is a bit more difficult. To get (19) the same arguments work, the final reasoning with Lemma 4.33(b) being replaced by a direct use of Remark 4.31(c), as for all $\{r_1,\dots,r_\ell\}$, G_{r_ℓ} fits $\mathrm{Lev}^{r_1,\dots,r_\ell}$.

We recall that MQ was defined as the unique Markov law with the same marginals of dimension 2 as $G_\#\mathrm{Lev}$. But by (25) and Proposition 4.16(c), the 2-marginals of $G_\#\widetilde{\mathrm{Lev}}$ and $G_\#\mathrm{Lev}$ are equal. Hence it is sufficient to prove that $G_\#\widetilde{\mathrm{Lev}}$ is Markov, i.e., that for all $(s_1,\dots,s_k)$, $(G_\#\widetilde{\mathrm{Lev}})^{s_1,\dots,s_k} = (G_\#\widetilde{\mathrm{Lev}})^{s_1,s_2} \circ \cdots \circ (G_\#\widetilde{\mathrm{Lev}})^{s_{k-1},s_k}$. Since $\widetilde{\mathrm{Lev}}$ is Markov, $(\widetilde{\mathrm{Lev}})^{s_2,\dots,s_k} = \widetilde{\mathrm{Lev}}^{s_2,s_3} \circ \cdots \circ \widetilde{\mathrm{Lev}}^{s_{k-1},s_k}$; besides, notice the following fact, that we will prove a bit below:

Fact. — For all $(s_i)_{i=1}^k \in \mathbb{R}^k$, $\mathrm{Lev}^{s_1,\dots,s_k}$ viewed as a transport plan from $(\mathbb{R},\lambda)$ to $(\mathbb{R}^{k-1},\lambda^{\otimes k-1})$ has increasing kernel (for the order $\preceq_{\mathrm{lo}}$ instead of $\preceq_{\mathrm{sto}}$).

Hence we may conclude by using $k-1$ times Lemma 4.33(b). Let us check the first step. Since the measures ${}^t(\widetilde{\mathrm{Lev}}^{s_1,s_2})$ and $\widetilde{\mathrm{Lev}}^{s_2,\dots,s_k}$ have increasing kernel, we have only to show that $\widetilde{\mathrm{Lev}}^{s_1,s_2}.\widetilde{\mathrm{Lev}}^{s_2,s_3} = \widetilde{\mathrm{Lev}}^{s_1,s_2}.L_{\{s_2\}}.\widetilde{\mathrm{Lev}}^{s_2,s_3}$, i.e., by definition of $\widetilde{\mathrm{Lev}}$, that $L_{R\cap[s_1,s_2]}.L_{R\cap[s_2,s_3]} = L_{R\cap[s_1,s_2]}.L_{s_2}.L_{R\cap[s_2,s_3]}$. This amounts to checking that:

$$\begin{cases} L_{\llbracket s_1,s_2[}\,.\,L_{]s_2,s_3\rrbracket} = L_{\llbracket s_1,s_2[}\,.\,L_{s_2}\,.\,L_{]s_2,s_3\rrbracket} & \text{if } s_2 \notin R\\ L_{\llbracket s_1,s_2]}\,.\,L_{[s_2,s_3\rrbracket} = L_{\llbracket s_1,s_2]}\,.\,L_{s_2}\,.\,L_{[s_2,s_3\rrbracket} & \text{otherwise.}\end{cases}$$

The second point is true by Proposition 4.14, and if $s_2 \notin R$, by Proposition 4.26, s_2 is not an essential atomic time, which is the wanted equality. We finally must prove the fact stated above. Actually it is true for any concatenation $P_1 \circ \cdots \circ P_k$ of couplings P_i from $\mathbb{R}$ to $\mathbb{R}$ with increasing kernel —and $\mathrm{Lev}^{s_1,\dots,s_k}$ is of this type, see (28). We check this for $k=2$; the same argument, applied by induction, gives the general case. Take B_2 and B_3 two intervals of the type $]-\infty,a]$ and $x \leqslant x'$; we must show that: $k_{P_1\circ P_2}(x, B_2\times B_3) \geqslant k_{P_1\circ P_2}(x', B_2\times B_3)$. As P_2 has increasing kernel, the function $k_{P_2}(\cdot,B_3)$ is decreasing (thus also $\mathbb{1}_{B_2}k_{P_2}(\cdot,B_3)$). As P_1 has increasing kernel, $k_{P_1}(x,\cdot) \preceq_{\mathrm{sto}} k_{P_1}(x',\cdot)$. Thus:

$$k_{P_1\circ P_2}(x,B_2\times B_3) = \int k_{P_1}(x,\mathrm{d}y)\mathbb{1}_{B_2}k_{P_2}(y,B_3) \geqslant \int k_{P_1}(x',\mathrm{d}y)\mathbb{1}_{B_2}k_{P_2}(y,B_3),$$

the desired result. This proves part (b) of the theorem.

Now we prove part (c). For point (i), each 2-margin $\mathsf{Q}^{s,t}_{[R_n]}$ and $\mathsf{Q}^{s,t}_{[R'_n]}$ tends respectively to $\mathrm{losup}_n\{\mathsf{Q}^{s,t}_{[R_n]}\}$ and $\mathrm{losup}_n\{\mathsf{Q}^{s,t}_{[R'_n]}\}$, which are equal by Lemma 4.23 since $\bigcup_n R_n = \bigcup_n R'_n$. Besides, if $R'' \subset R$, taking a nested sequence $(R''_n)_n$ of finite sets such that $\bigcup_n R''_n = R''$ and $R''_n \subset R_n$ for all n, we get that, for any s and $t>s$, $\lim_n \mathsf{Q}^{s,t}_{[R''_n]} \preceq_{\mathrm{lo}} \lim_n \mathsf{Q}^{s,t}_{[R_n]} = \mathsf{MQ}^{s,t}$, but by the minimality property of Theorem A(iii), for all n, $\mathsf{MQ}^{s,t} \preceq_{\mathrm{lo}} \mathsf{Q}^{s,t}_{[R''_n]}$. Therefore, $\lim_n \mathsf{Q}^{s,t}_{[R''_n]} = \mathsf{MQ}^{s,t}$ and the result follows. For (ii), it is sufficient to show that for any finite subsets R and E of $\mathbb{R}$ with μ_t diffuse for all $t\in E$, $\mathsf{Q}^{s,t}_{[R\cup E]} = \mathsf{Q}^{s,t}_{[R]}$. This follows plainly from the definitions. After Definition 4.18 it amounts to showing that for all $t \in E$, all $\{s_1,\dots,s_k\} \subset\ ]-\infty,t[$ and all $\{s'_1,\dots,s'_{k'}\} \subset\ ]t,+\infty[$, $\mathrm{Q}^{s_1,\dots,s_k,t} \circ \mathrm{Q}^{t,s'_1,\dots,s'_{k'}} = \mathrm{Q}^{s_1,\dots,s_k,t,s'_1,\dots,s'_{k'}}$. This comes from a trivial case of Remark 4.31(c) applied with all the measures equal to $\lambda_{[0,1]}$, $P = {}^t\mathrm{Id}_{\lambda,k+1}$ (i.e., $P = \mathrm{Id}_{\lambda,k+1}$), $Q = \mathrm{Id}_{\lambda,k'+1}$ and $(f,g,h) = (\bigotimes_{i=1}^k G_{s_i}, G_t, \bigotimes_{i=1}^{k'} G_{s'_i})$. Indeed, as μ_t is diffuse, G_t is injective, hence trivially fits P and Q. Then:

$$
\begin{aligned}
&\mathrm{Q}^{s_1,\dots,s_k,t} \circ \mathrm{Q}^{t,s'_1,\dots,s'_{k'}}\\
&\qquad = (\textstyle\bigotimes_{i=1}^k G_{s_i} \otimes G_t)_\# \mathrm{Id}_{\lambda,k+1} \circ (G_t \otimes (\bigotimes_{i=1}^{k'} G_{s'_i}))_\# \mathrm{Id}_{\lambda,k'+1} \quad \text{by definition of Q}\\
&\qquad = ((\textstyle\bigotimes_{i=1}^k G_{s_i}) \otimes G_t \otimes (\bigotimes_{i=1}^{k'} G_{s'_i}))_\# \mathrm{Id}_{\lambda,k+1} \circ \mathrm{Id}_{\lambda,k'+1} \quad \text{by the remark}\\
&\qquad = ((\textstyle\bigotimes_{i=1}^k G_{s_i}) \otimes G_t \otimes (\bigotimes_{i=1}^{k'} G_{s'_i}))_\# \mathrm{Id}_{\lambda,k+k'+1}\\
&\qquad = \mathrm{Q}^{s_1,\dots,s_k,t,s'_1,\dots,s'_{k'}}.
\end{aligned}
$$

To alleviate the writing, we prove the rest of (ii), with $\sharp E' = 1$ (the general proof is alike). Take R given by point (a) and $t \in\ ]s,s'[$ some unessential atomic time of $(\mu_t)_t$, then $L_{R_n \cap]s,s'[\cup \{t\}} = L_{R_n \cap]s,t[} . L_{R_n \cap]t,s'[} \to L_{]s,t[} . L_{]t,s'[}$ by Lemma 2.23. Now $L_{]s,t[} . L_{]t,s'[} = L_{]s,s'[}$ since t is unessential. Thus $R' = R \cup \{t\}$ satisfies (25), hence (26), by point (b).

Let us prove (iii). Suppose that I is some essential atomic interval, i.e., there is an interval $J \supset I$ such that $J \setminus I$ is disconnected and $L_J \neq L_{J \setminus I}$, and assume that $I \cap R = \varnothing$. Since $L_J \preceq_{\mathrm{lo}} L_{J\setminus I}$, this means that there is some $(a,a') \in\ ]0,1[^2$ such that $L_J([0,a] \times [0,a']) < L_{J\setminus I}([0,a] \times [0,a'])$. Denote $(\inf J, \sup J) \in \mathbb{R}^2$ by (s,s'). If $a \notin A_s$ and $a' \notin A_{s'}$ (see Notation 4.4), i.e., if $G([0,a^{(\prime)}]) = [0, G(a^{(\prime)})]$, then, pushing the inequality by G, and reminding that, since $I \cap R = \varnothing$, $L_{J\setminus I} \preceq_{\mathrm{lo}} \lim_n L_{R_n \cap J}$, so that $L_{J\setminus I}([0,a] \times [0,a']) \leqslant \lim_n L_{R_n \cap J}([0,a] \times [0,a'])$, we get:

$$
\mathrm{MQ}^{s,s'}(]-\infty, G(a)] \times]-\infty, G(a')]) < \lim_{n\to\infty} \mathrm{Q}^{s,s'}_{[R_n]}(]-\infty, G(a)] \times]-\infty, G(a')]). \tag{29}
$$

Therefore $\lim_{n\to\infty} \mathrm{Q}_{[R_n]}$ cannot be equal to MQ and we are done. If $a \in A_s$, let $A =]a_0, a_1[$ be the connected component of A_s containing a; notice that $\{a_0, a_1\} \cap A = \varnothing$. We prove that (29) holds with a_0 or a_1 in place of a. By construction of L_E for any set E having s as minimum, the functions $b \mapsto k_{L_E}(b,.)$, the values of which are measures on $[0,1]$, are constant on A. Applying this for $E = L_{\{s\}\cup J}$ and $E = L_{\{s\}\cup J\setminus I}$, we get that either (i) below is true, or the restrictions of L_J and $L_{J\setminus I}$ to $[0,a_0] \times [0,a']$ coincide, thus necessarily (ii) is true (we let $\{s\} \cup J$ appear instead of J but this does not matter by Remark 4.17):

(i) $L_{\{s\}\cup J}([0,a_0] \times [0,a']) < L_{\{s\}\cup J\setminus I}([0,a_0] \times [0,a'])$,

(ii) on A, $k_{L_{\{s\}\cup J}}(b,[0,a']) < k_{L_{\{s\}\cup J\setminus I}}(b,[0,a'])$,

and if (i) is false and (ii) is true then $L_J([0,a_1] \times [0,a']) < L_{J\setminus I}([0,a_1] \times [0,a'])$. Hence anyway (29) holds with a_0 or a_1 in place of a. Proceed symmetrically for a' if $a' \in A_{s'}$. This shows point (iii).

Remark 4.35. — The proof of Theorem 4.21(c)(ii) shows directly the last sentence of Remark 1.8(a). Indeed it shows that if every μ_t is diffuse, then for all finite $R \subset \mathbb{R}$, $\mathrm{Q}_{[R]} = \mathrm{Q}$, which gives an expression of the Markov property, see Definition 2.11.

5. Examples and open questions

5.1. Example of Markov-quantile processes attached to discrete measures on $\mathbb{N}$

In this section x_+ is the positive part $\max\{0, x\}$ of any $x \in \mathbb{R}$.

Example 5.1 (Discrete measures). — Let $(\mu_t)_{t\in[0,1]}$ be concentrated on $\mathbb{N}$ for every t and assume that for every $k \in \mathbb{N}$ the map $A_k : t \mapsto \sum_{i=0}^{k} \mu_t(i)$ is in $\mathrm{C}^1([0,1])$ and piecewise monotone (e.g., A_k is analytic). Let moreover A_{-1} be the zero constant function. We assume that:

$$\max\Big(-\frac{A_k'(t)}{\mu_t(k)}, \frac{A_{k-1}'(t)}{\mu_t(k)}\Big)$$

is bounded from above for $(t, k) \in [0,1] \times \mathbb{N}$. Then, using the characterization of the Markov-quantile process as a limit of quantile couplings, namely Theorem A(iv), it can be proved that the Markov-quantile process $(X_t)_{t\in[0,1]}$ is the time continuous Markov chain with jump rate $q_{k,k+1} = (-A_k'(t))_+/\mu_t(k)$ from k to $k+1$, and $q_{k,k-1} = (A_{k-1}'(t))_+/\mu_t(k)$ from k to $k-1$ and $q_{k,j} = 0$ for $|j-k| \neq 1$. Denoting $\mathrm{P}(X_t = k)$ by p_k it means that the so-called forward Kolmogorov–Chapman system is satisfied:

$$\frac{\mathrm{d}p_k}{\mathrm{d}t}(t) = \begin{cases} p_1 q_{1,0} - p_0 q_{0,1} & \text{if } k = 0,\\ p_{k+1}q_{k+1,k} + p_{k-1}q_{k-1,k} - p_k(q_{k,k-1} + q_{k,k+1}) & \text{if } k \in \mathbb{N}^*, \end{cases}$$

where the derivative is a right derivative. Recall that the jump rate is defined for $i \neq j$ by:

$$q_{i,j}(t) = \lim_{h\to 0^+} \frac{\mathrm{P}(X_{t+h} = j \mid X_t = i)}{h}.$$

The classical theory that can be read in Feller's book [13, Chap. XVII, §9] and the references therein (see also [10]) ensures that our process is solution of the forward Kolmogorov–Chapman system. The uniqueness of the solution for a Markov process is obtained from the uniform bound on the rates $q_{i,j}(t)$.

In place of a complete proof let us compute the jump rate in a typical case. Notice before that similar computations can be found in [24, §4]. We are looking for the jump rate $q_{k,k+1}(t)$ in the case of $A := A_{k-1}$ and $B := A_k$ locally decreasing on the right of t. At every time t the atomic measure μ_t is completely described by the partition of the interval $[0,1]$ of quantile levels through the sequence $(A_k(t))_{k\in\mathbb{N}}$. Indeed, $]A_{k-1}(t), A_k(t)[\subset [0,1]$ is the interval of the quantile levels of the atom $\mu_t(k)\delta_k$. Recall that both A and B are in $\mathrm{C}^1([0,1])$. We can assume that h is so small that $B(t+h) > A(t)$. For $\{r', r''\} \subset [t, t+h]$ with $r' < r''$ and r'' close to r', the quantile coupling between $\mu_{r'}$ and $\mu_{r''}$ transports the main part of the mass of the atom $\mu_{r'}(\{k\})\delta_k$ on itself and the rest on the atoms $\mu_{r''}(\{k'\})\delta_{k'}$ with $k' > k$. We aim at proving that the conditional probability to be still in k at time $t+h$ is:

$$1 - \frac{B'(t)}{(B-A)(t)}\,h + O(h^2). \tag{30}$$

Since the probability to jump more than twice is $O(h^2)$, (30) furnishes the announced jump rate $q_{k,k+1}(t) = (-A_k'(t))_+/\mu_t(k)$ in the case of decreasing functions. So let us prove (30).

We consider a partition $R=\{r_0,\dots,r_m\}$ of $[t,t+h]$ with $(r_0,r_m)=(t,t+h)$ and the discrete quantile Markov chain associated with it. As A and B are decreasing, note that no mass can leave the quantile level interval $[A,B]$ and come back on it in the same time interval $[t,t+h]$. On $[r_n,r_{n+1}]$ the probability to stay in the interval is

$$\frac{B(r_{n+1})-A(r_n)}{B(r_n)-A(r_n)} = 1-\frac{B(r_n)-B(r_{n+1})}{B(r_n)-A(r_n)}$$

as one can easily convince oneself with a picture similar to the left part of Figure 2. We let the proof of the following fact to the reader: there exists $\delta=O(h)$ such that for every $r'<r''$ in $[t,t+h]$ we have:

$$e^{\left(-(1+\delta)B'(t)(r''-r')/(B-A)(t)\right)} \leqslant \frac{B(r'')-A(r')}{B(r')-A(r')} \leqslant e^{\left(-(1-\delta)B'(t)(r''-r')/(B-A)(t)\right)}.$$

We obtain this estimate for each interval $[r_n,r_{n+1}]\subset[t,t+h]$. Multiplying all together, we see that the probability to stay on the same state after m steps is in

$$\left[\exp\left(-(1+\delta)\frac{B'(t)}{(B-A)(t)}h\right),\exp\left(-(1-\delta)\frac{B'(t)}{(B-A)(t)}h\right)\right],$$

where $\delta=O(h)$. A simple Taylor expansion gives (30).

Example 5.2 (Poisson distributions). — Elaborating on the last example we consider, for $t\in\mathbb{R}^+$, $\mu_t=\mathrm{P}(t)$, where $\mathrm{P}(t)$ is the Poisson law of parameter t. In this case $A_k(t)=\sum_{i=0}^k\exp(-t)t^i/i!$ so that the jump rate $q_{k,k+1}(t)$ is constantly 1 for every k and t, and the other rates are zero. We recover the Poisson process. Note that the Poisson laws are in stochastic order, which matches with the increasing trajectories of the Poisson counting process.

Example 5.3 (Binomial distributions). — In this example $\mu_t=\mathrm{B}(n,t)$ for $t\in[0,1]$. Let us define a Markov process $X=(X_t)_{t\in[0,1]}\in\mathrm{Marg}((\mu_t)_t)$ and compute its jump rates; we will then see that $\mathrm{Law}(X)=\mathrm{MQ}$. We define X on the probability space $[0,1]^n$ by $X:(\alpha_1,\dots,\alpha_n)\mapsto\sum_{k=0}^n\mathbb{1}_{\alpha_k\in[0,t]}$, so its law is μ_t. The fact that $(X_t)_{t\in[0,1]}$ is Markov comes from the following coarse argument: provided k coordinates of $\alpha=(\alpha_1,\dots,\alpha_n)$ are smaller than t, the distribution is uniform on $[0,t]^k$ for the k coordinates of the past of t and on $[t,1]^{n-k}$ for the $n-k$ of its future. Between t and $t+h$ the probability to have (at least, as well as exactly) one jump is $[(n-k)h/(1-t)]+O(h^2)$. As $A_k(t)=\sum_{i=0}^k\binom{n}{i}t^i(1-t)^{n-i}$ with the notation of Example 5.1, it can easily be checked that $(n-k)/(1-t)=A_k'(t)/\mu_t(k)$, which proves that $(X_t)_{t\in[0,1]}$ is the Markov-quantile process attached to $(\mu_t)_{t\in[0,1]}$. This example could be of interest with respect to previous works on the entropic interpolation on graphs as, e.g., [17, 32].

5.2. Example of Markov-quantile transport processes. — The following examples are also related to the relations of the Markov-quantile process with the Continuity Equation extensively explained in [8] and that we evoked in the introduction in Section 1.4. In particular, we will consider processes tangent to a non-autonomous vector field on $\mathbb{R}$. Basically, in the examples, μ_t is made of two parts that are translated in opposite directions and cross. We examine three crossing situations for atomic or diffuse measures.

Example 5.4 (One atom crossing a diffuse measure). — Consider $\mu = (\mu_t)_{t\in[0,1]}$ with $\mu_t = \frac{1}{2}\lambda|_{[t-3/4,t-1/4]} + \frac{1}{2}\delta_0$. This is the family of marginals of a simple process P with affine trajectories, defined by $P(t \mapsto 0) = 1/2$ and $P(\{t \mapsto x_0 + t : x_0 \in A\}) = \lambda|_{[-3/4,-1/4]}(A)$. This is not the Markov-quantile process attached to μ but it is a Markov process and it is tangent to the vector field defined by:

$$v_t(x) = 0 \text{ if } x = 0 \quad\text{and}\quad v_t(x) = 1 \text{ otherwise.}$$

Now the theory of the Continuity Equation developed in [2] and that we reexplain in [8] permits to identify v as a minimal vector field (for the Benamou–Brenier functional) since the quantile process is also almost surely an integral curve of it. Therefore, as the quantile process does, the process $Y := (Y_t)_{[0,1]} \sim P$ is minimizing E(E(X)) the expected energy that we briefly presented in Section 1.4, among the processes with marginals μ.

The Markov-quantile process $(X_t)_{t\in[0,1]}$ attached to $(\mu_t)_{t\in[0,1]}$ can be described as follows: the trajectories start according to μ_0 and are piecewise affine, with pieces taken from the affine curves above. Provided $X_0 \in [-3/4,-1/4]$, the first piece is $X_t = X_0 + t$ on $[0,\tau]$, where $\tau = -X_0$. The second affine piece is constant equal to zero on $[\tau, \min(\tau+\eta, 1)]$, where η is an exponential random variable of parameter 2, independent from X_0. The third piece, if it exists, is affine of slope 1, namely $X_t = t - (\tau+\eta)$ on $[\tau+\eta, 1]$.

Finally, in the present example, as P, the Markov-quantile process is minimizing E(E(X)), but unlike it and in addition, it has increasing kernels and is a *strongly* Markovian process.

Example 5.5 (Crossing of two purely atomic measures). — Consider two measures α and β of mass $1/2$, concentrated on the rational numbers of $[0,1]$, with finite or infinite support. Let τ_t be the translation of vector t in R. Set

$$\mu = (\mu_t)_{t\in\mathbb{R}} = ((\tau_t)_\#\alpha + (\tau_{-t})_\#\beta)_{t\in\mathbb{R}}.$$

As in Example 5.4 the measure $P \in \mathrm{Marg}(\mu)$ is concentrated on the space of piecewise affine paths (of slopes 1 and -1) is a minimizer of the action. The two measures $(\tau_t)_\#\alpha$ and $(\tau_{-t})_\#\beta$ are both concentrated on Q when $t \in \mathbb{Q}$ and they are mutually singular if $t \notin \mathbb{Q}$. Hence according to [8], the optimal vector field $(v_t)_{t\in[0,1]}$ satisfies $\lambda|_{[0,1]}$-almost surely $v_t = \pm 1$. It can be checked that the Markov-quantile process is again piecewise affine with a random finite number of changes of slope. Interesting exercises on the Markov-quantile process can be considered, as for instance finding the probability for a trajectory coming from $-\infty$ in $-\infty$ to tend to $+\infty$ in $+\infty$. Note that the situation seems to be well approached by truncating the measure to finitely many 'big' atoms. This corresponds to the case of α and β with finite support. In this particular case the above mentioned exercise reduces to the so-called 'gladiator game' [27] that is a stochastic version of Borel's Blotto game [42].

Example 5.6 (Crossing of two diffuse measures). — Consider

$$\mu_t = \lambda|_{[t-2,t-1]} + \lambda|_{[1-t,2-t]}$$

and again such that

$$(\{t \mapsto t + x_0 : x_0 \in A\}) = \lambda|_{[-2,-1]}(A) \quad \text{and} \quad (\{t \mapsto x_0 - t : x_0 \in A\}) = \lambda|_{[1,2]}(A).$$

Unlike in the previous examples, does not minimize A on $\mathrm{Marg}_{\mathrm{C}}(\mu)$. All the measures μ_t are continuous so that the Markov-quantile process $(X_t)_{t\in\mathbb{R}}$ is the quantile process. It is affine by part and continuous. With probability 1/2, in fact if $X_0 \leqslant 0$, first it has slope 1, then slope 0 on $[(1-X_0)/2,(5+X_0)/2]$ and finally slope -1. If $X_0 > 0$, the process $(X_t)_t$ starts with slope -1, is flat on $[(1+X_0)/2,(5-X_0)/2]$ and continues with slope 1 after $(5-X_0)/2$.

5.3. Theoretic Markov-quantile processes

Example 5.7 (One atom with regular level functions). — Take $(\mu_t)_{t\in[0,1]}$ such that for every t, μ_t has exactly one atom $x_t \in \mathbb{R}$ and the interval of quantile levels of this atom at time t is $]A(t),B(t)[$. Assume moreover that A and B are of class C^1 and piecewise monotone. Then the Markov-quantile process $(X_t)_{t\in[0,1]}$ can be described using two Poisson point processes of jump rates $(A')_+/(B-A)$ and $(B')_-/(B-A)$. Conditionally on $F_{\mu_t}(X_t) \in\ ]A(t),B(t)[$, we have $X_t = x_t$ until the next time $t_0 > t$ in the point process. Then the process $(X_t)_t$ leaves x_t and starts a piece of quantile trajectory constant in the space $[0,1]$ of quantile levels with value $A(t_0)$ or $B(t_0)$. The process may hit again x_t if there exists some $t_1 > t_0$ with $A(t_1) = x_{t_0}$, or $B(t_1) = x_{t_0}$ respectively.

The next remark is of general interest and particularly significant with respect to Remark 5.8. It presents the Markov-quantile process as one end of the spectrum of processes of law in $\mathrm{Marg}(\mu)$ that satisfies (ii) of Theorem A, i.e., have increasing kernels, the other end being the independent process.

Remark 5.8. — The minimality condition (iii) of Theorem A satisfied by the Markov-quantile process $(X_t)_t$ attached to some $(\mu_t)_t$ can also be stated as follows. For every process $(Y_t)_{t\in\mathbb{R}}$ satisfying (i) and (ii) of Theorem A, for every $s < t$ and every $x \in \mathbb{R}$ it holds:

$$\mathrm{Law}(X_t | X_s \leqslant x) \preceq_{\mathrm{sto}} \mathrm{Law}(Y_t | Y_s \leqslant x).$$

A similar relation that concerns maxima of $\preceq_{\mathrm{sto}}$ in place of minima is satisfied by the independent process $(Z_t)_{t\in\mathbb{R}}$. If a process $(Y_t)_t$ has increasing kernels, we obtain:

$$\mathrm{Law}(Y_t | Y_s \leqslant x) \preceq_{\mathrm{sto}} \mathrm{Law}(Y_t | Y_s < +\infty) = \mathrm{Law}(Y_t) = \mathrm{Law}(Z_t | Z_s \leqslant x).$$

We conclude with the following result: Assume that for some $s < t$ and $(X_t)_t$ the Markov-quantile process, X_s is independent of X_t. Then for any process $(Y_t)_t$ satisfying (i) and (ii) of Theorem A, we have for every $x \in \mathbb{R}$:

$$\mu_t = \mathrm{Law}(X_t | X_s \leqslant x) \preceq_{\mathrm{sto}} \mathrm{Law}(Y_t | Y_s \leqslant x) \preceq_{\mathrm{sto}} \mathrm{Law}(Z_t | Z_s \leqslant x) = \mu_t,$$

so that, due to the Markov property, for every $s' \leq s$ and $t' \geq t$, $Y_{s'}$ and $Y_{t'}$ are independent.

Example 5.9 (Two atoms). — We set $\mu_t = a(t)\delta_0 + b(t)\delta_1$ with $a + b = 1$ but do not assume any regularity on the functions a and b. Let $(X_t)_t$ be the Markov-quantile process. We shall show that X_s and X_t are independent if and only if the total variation of a (or b) on $[s,t]$ is infinite or $m_{s,t} := \min(\inf_{[s,t]} a, \inf_{[s,t]} b) = 0$. If $m_{s,t} = 0$ the independence is true for any Markov process. Indeed, by assumption, for any $\varepsilon > 0$, we may take r such that $\mathrm{P}(X_r = 0)$ is small enough so that $\mathrm{P}(X_r = 0|X_s = 0) \leq \varepsilon$, $\mathrm{P}(X_r = 1|X_s = 0) \geq 1-\varepsilon$, and $|\mathrm{P}(X_t = 0|X_r = 1) - \mathrm{P}(X_t = 0)| \leq \varepsilon$. Then:

$$\begin{aligned}
|\mathrm{P}(X_t = 0|X_s = 0) - \mathrm{P}(X_t = 0)|
&= \big|\mathrm{P}(X_t = 0|X_r = 0)\mathrm{P}(X_r = 0|X_s = 0) \\
&\qquad + \mathrm{P}(X_t = 0|X_r = 1)\mathrm{P}(X_r = 1|X_s = 0) - \mathrm{P}(X_t = 0)\big| \\
&\qquad\qquad\qquad \text{since } X \text{ is Markov} \\
&\leq |\mathrm{P}(X_t = 0|X_r = 1) - \mathrm{P}(X_t = 0)| + 2\varepsilon \\
&\leq 3\varepsilon,
\end{aligned}$$

which is the wanted independence.

Hence, we assume that $m_{s,t} > 0$ and that a takes values in $[m_{s,t}, 1 - m_{s,t}]$ in $\mu_t = a(t)\delta_0 + b(t)\delta_1$. We are left with the task to prove that independence is equivalent to a infinite total variation of a on $[s,t]$. Let θ_0 be the uniform measure on $[0, a(s)]$. Our goal reduces to establishing $\lambda = \theta_0\ell_{[s,t]} := \mathrm{sto}\sup_R \theta_0\ell_{r_1}\ell_{r_2}\cdots\ell_{r_m}$, where R ranges among the partitions $\{r_0, \ldots, r_m\}$ with $(r_0, r_m) = (s,t)$. For the measures under consideration, if $a(r_{k-1}) \leq a(r_k) \leq a(r_{k+1})$ or $a(r_{k-1}) \geq a(r_k) \geq a(r_{k+1})$ it holds $\ell_{r_{k-1}}\ell_k\ell_{k+1} = \ell_{r_{k-1}}\ell_{r_{k+1}}$. Therefore we can assume without loss of generality that the sequence $(a(r_k))_{k=0,\ldots,m}$ has increments with alternating sign, for instance $a(r_{2k+1}) \geq a(r_{2k})$ for every k. We define $\theta_n = \theta_0\ell_{r_0}\cdots\ell_{r_n}$.

The measure θ_n can be written in the form:

$$\theta_n = d_n\lambda_{[0,a(r_n)]} + d'_n\lambda_{[a(r_n),1]} \in \mathcal{M}^{\nearrow},$$

where $d_n = a(r_n)^{-1}\theta_n([0, a(r_n)])$ in fact parametrizes the complete measure. Note, after Remark 5.8 that $\theta_n \preceq_{\mathrm{sto}} \lambda$, which means $d_n \geq 1 \geq d'_n$. As $a(r_n) \in [m_{s,t}, 1 - m_{s,t}]$, the sequence converges to λ if and only if $d_n \to 1$.

Recalling the effect of the kernel $\ell_{r_{n+1}}$, described on Figure 3 and defined in Notation 4.4(b), we find:

$$d_{n+1} - 1 = \begin{cases} (d_n - 1)\dfrac{a(r_{n+1})}{a(r_n)} & \text{if } a(r_{n+1}) < a(r_n), \\[2ex] (d_n - 1)\dfrac{b(r_{n+1})}{b(r_n)} & \text{otherwise.} \end{cases}$$

The product $\prod_{n=1}^{m} \min\big(\frac{a(r_{n+1})}{a(r_n)}, \frac{1-a(r_{n+1})}{1-a(r_n)}\big)$ can be arbitrarily close to zero (over all partitions of $[s,t]$) if and only if $a \in [m_{s,t}, 1 - m_{s,t}]$ has infinite total variation. This proves the claimed equivalence.

Example 5.10 (One atom on the lower levels). — Consider $(\mu_t)_{t\in[0,1]}$ such that for every t, μ_t has exactly one atom and this atom is between the quantile levels $A(t) = 0$ and $B(t)$. An example is $\mu_t = B(t)\delta_0 + (1 - B(t))\mathrm{E}(1)$, where δ_0 is the Dirac mass in zero and $\mathrm{E}(1)$ the exponential law of parameter 1. No regularity assumption is made on B. Similar observations as in Example 5.9 permit us to specify the kernel between time s and $t > s$. Let $\alpha_{s,t}$ be $\sup_{r\in[s,t]} B(r)$. Then $L_{[s,t]}$ is simply the uniform measure of mass $\alpha_{s,t}$ on $[0,\alpha_{s,t}]^2$ plus the one-dimensional uniform measure of mass $1 - \alpha_{s,t}$ on the diagonal between $(\alpha_{s,t},\alpha_{s,t})$ and $(1,1)$. The same for the kernel $\ell_{s,t}$ reads:

$$\ell_{s,t}(x,\cdot) = \begin{cases} \alpha_{s,t}^{-1}\lambda|_{[0,\alpha_{s,t}]} & \text{if } x \leqslant \alpha_{s,t},\\ \delta_x & \text{if } x > \alpha_{s,t}.\end{cases}$$

A particle of quantile value $\leqslant \alpha_{s,t}$ at time s is uniformly mapped at time t on the particles of quantile levels $[0,\alpha_{s,t}]$. If the quantile value at time s is greater that $\alpha_{s,t}$, the particle keeps on with the same level until time t as if it were the quantile process.

5.4. Transformations of Markov-quantile processes

Example 5.11 (Markov-quantile processes). — According to $(\mu_t)_t$, a quantile process may be Markov or not. Recalling Remark 1.8 (b), if Q is Markov it coincides with the Markov-quantile process. As proved in [23, Prop. 3], the criterion is the following: the process is *not* Markov if and only if there exists $\alpha \neq \alpha' \in [0,1]$ and $t_1 < t_2 < t_3$ such that the α-quantile and the α'-quantile of μ_{t_2} are equal but that those of μ_{t_1} and μ_{t_3} differ. This can be summarized saying that "X's are forbidden", where X refers to the shape of the letter, the four ends being $G_{\mu_1}(\alpha)$, $G_{\mu_1}(\alpha')$, $G_{\mu_3}(\alpha)$ and $G_{\mu_3}(\alpha')$, the intersection being $G_{\mu_2}(\alpha) = G_{\mu_2}(\alpha')$. Other letters like O, Y and Z are allowed.

Remark 5.12 (Reversal of time and twist of space and time). — If $(X_t)_{t\in\mathbb{R}}$ is the Markov-quantile process attached to $(\mu_t)_{t\in\mathbb{R}}$ then X_{-t} is the Markov-quantile process in $\mathrm{Marg}((\mu_{-t})_{t\in\mathbb{R}})$. This comes from Theorem A (iv) on the limit of products of quantile couplings and the fact that ${}^t\mathrm{Q}(\mu,\nu) = \mathrm{Q}(\nu,\mu)$. More generally, for every homeomorphism φ from $\mathbb{R}$ into $\mathbb{R}$, $X_{\varphi(t)}$ has law the Markov-quantile measure of $\mathrm{Marg}((\mu_{\varphi(t)})_{t\in\mathbb{R}})$. Of course non-injective monotone continuous map φ may be used too.

Moreover, if for all t, $f_t : \mathbb{R} \to \mathbb{R}$ are strictly monotone functions with the same orientation for all t, the process $f_t(X_t)$ is a Markov-quantile process.

Remark 5.13. — The Markov-quantile process being time-reversible, and as a coupling $P \in \mathrm{Marg}(\mu,\nu)$ has increasing kernel if and only if ${}^tP \in \mathrm{M}^{\nearrow}(\nu,\mu)$ and $P \preceq_{\mathrm{lo}} Q$ is equivalent to ${}^tP \preceq_{\mathrm{lo}} {}^tQ$ (recall Remark 3.16 and Definition 3.4), points (ii) and (iii) of in Theorem A can be replaced by

(ii') For every $s < t$, $\mathrm{MQ}^{s,t} \in \mathrm{M}^{\nearrow}(\mu_s,\mu_t)$, and

(iii') for every $s < t$, $\mathrm{MQ}^{s,t}$ is a minimal coupling for $\preceq_{\mathrm{lo}}$ among the processes satisfying (i) and (ii').

5.5. Open questions

5.5.1. *Markovinification*

(a) We may interpret the resulting process in Theorem 2.26 as the Markov process that has infinitesimally the same transitions as P. However, as it depends on the choice of the partitions, this Markov process is not a priori uniquely determined. At which conditions is this Markov process *uniquely determined* and how can it be characterized? If the initial process is a quantile process, we proved in Theorem A that the answer is yes, without condition, and characterized it using orderings.

(b) In Theorem B we proved that Markovinification of the quantile process occurs for processes $(\mathsf{Q}_{[R_n]})_n$ in place of consistent (see Definition 2.14) transport plans. Does an *existence* statement analogous to Theorem 2.26 happen when we consider sequences $(P_{[R_n]})_n$?

5.5.2. *The Kellerer theorem in dimension* $d \geqslant 2$. — Kellerer proved Theorem 1.17 for martingales in $\mathbb{R}$. For measures $(\mu_t)_t$ on $\mathbb{R}^d$, increasing in convex order, it is known that there exists an associated martingale [19, 3] but not whether one of them is Markov. This is a major question. Note that another interpretation of the question in higher dimension is possible when considering martingales indexed by a multidimensional set (see [18, Prob. 7b]). This problem was solved in [23].

5.5.3. *Markov Kamae–Krengel theorem*. — Kamae and Krengel proved in [26] that if $(\mu_t)_{t\in\mathbb{R}}$ are measures on a partially ordered Polish space E such that $t \mapsto \mu_t$ is increasing for the stochastic order, in the sense that $t \mapsto \int f \mathrm{d}\mu_t$ is increasing for any increasing bounded $f : E \to \mathbb{R}$, there exists an increasing process $(X_t)_t$ with law in $\mathrm{Marg}(\mu)$. We proved in Theorem A and C that if E is $\mathbb{R}$, the process can moreover be Markov. A natural problem is whether this is also true for any E.

5.5.4. *Strong Markov property*. — The Markov-quantile process is not always the unique Markov process that minimizes the energy E, as is shown in Example 5.4. However in this example MQ is strongly Markovian. Is the Markov-quantile process strongly Markovian for every $(\mu_t)_t$? Does the strong Markov property characterize it for curves of finite energy?

References

[1] J. M. P. Albin – "A continuous non-Brownian motion martingale with Brownian motion marginal distributions", *Statist. Probab. Lett.* **78** (2008), no. 6, p. 682–686.

[2] L. Ambrosio, N. Gigli & G. Savaré – *Gradient flows in metric spaces and in the space of probability measures*, second ed., Lectures in Math. ETH Zürich, Birkhäuser Verlag, Basel, 2008.

[3] M. Beiglböck, M. Huesmann & F. Stebegg – "Root to Kellerer", in *Séminaire de Probabilités XLVIII*, Lect. Notes in Math., vol. 2168, Springer, Cham, 2016, p. 1–12.

[4] M. Beiglböck & N. Juillet – "On a problem of optimal transport under marginal martingale constraints", *Ann. Probab.* **44** (2016), no. 1, p. 42–106.

[5] ______, "Shadow couplings", *Trans. Amer. Math. Soc.* **374** (2021), no. 7, p. 4973–5002.

[6] J.-D. Benamou & Y. Brenier – "A numerical method for the optimal time-continuous mass transport problem and related problems", in *Monge Ampère equation: applications to geometry and optimization (Deerfield Beach, FL, 1997)*, Contemp. Math., vol. 226, American Mathematical Society, Providence, RI, 1999, p. 1–11.

[7] P. Billingsley – *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics, John Wiley & Sons Inc., New York, 1999.
[8] C. Boubel & N. Juillet – "On absolutely continuous curves in the Wasserstein space over R and their representation by an optimal Markov process", 2021, arXiv:2105.02495.
[9] G. Brunick & S. Shreve – "Mimicking an Itô process by a solution of a stochastic differential equation", *Ann. Appl. Probab.* **23** (2013), no. 4, p. 1584–1628.
[10] W. Doeblin – "Sur certains mouvements aléatoires discontinus", *Skand. Aktuarietidskr.* **22** (1939), p. 211–222.
[11] B. Dupire et al. – "Pricing with a smile", *Risk* **7** (1994), no. 1, p. 18–20.
[12] H. Federer – *Geometric measure theory*, Grundlehren Math. Wiss., vol. 153, Springer-Verlag New York Inc., New York, 1969.
[13] W. Feller – *An introduction to probability theory and its applications. Vol. I*, third ed., John Wiley & Sons, Inc., New York-London-Sydney, 1968.
[14] I. Gyöngy – "Mimicking the one-dimensional marginal distributions of processes having an Itô differential", *Probab. Theory Relat. Fields* **71** (1986), no. 4, p. 501–516.
[15] K. Hamza & F. C. Klebaner – "A family of non-Gaussian martingales with Gaussian marginals", *J. Appl. Math. Stochastic Anal.* (2007), article no. 92723 (19 pages).
[16] P. Henry-Labordère, X. Tan & N. Touzi – "An explicit martingale version of the one-dimensional Brenier's theorem with full marginals constraint", *Stochastic Process. Appl.* **126** (2016), no. 9, p. 2800–2834.
[17] E. Hillion – "$W_{1,+}$-interpolation of probability measures on graphs", *Electron. J. Probab.* **19** (2014), article no. 92 (29 pages).
[18] F. Hirsch, C. Profeta, B. Roynette & M. Yor – *Peacocks and associated martingales, with explicit constructions*, Bocconi & Springer Series, vol. 3, Springer, Milan, 2011.
[19] F. Hirsch & B. Roynette – "On R^d-valued peacocks", *ESAIM Probab. Statist.* **17** (2013), p. 444–454.
[20] F. Hirsch, B. Roynette & M. Yor – "Kellerer's theorem revisited", in *Asymptotic Laws and Methods in Stochastics. Volume in Honour of Miklos Csorgo*, Fields Inst. Commun. Series, vol. 76, Fields Inst. Res. Math. Sci., Toronto, ON, 2015, p. 347–363.
[21] D. G. Hobson – "Fake exponential Brownian motion", *Statist. Probab. Lett.* **83** (2013), no. 10, p. 2386–2390.
[22] ______, "Mimicking martingales", *Ann. Appl. Probab.* **26** (2016), no. 4, p. 2273–2303.
[23] N. Juillet – "Peacocks parametrised by a partially ordered set", in *Séminaire de Probabilités XLVIII*, Lect. Notes in Math., vol. 2168, Springer, Cham, 2016, p. 13–32.
[24] ______, "Martingales associated to peacocks using the curtain coupling", *Electron. J. Probab.* **23** (2018), article no. 8 (29 pages).
[25] O. Kallenberg – *Foundations of modern probability*, second ed., Probability and its Appl. (New York), Springer-Verlag, New York, 2002.
[26] T. Kamae & U. Krengel – "Stochastic partial ordering", *Ann. Probab.* **6** (1978), no. 6, p. 1044–1049.
[27] K. S. Kaminsky, E. M. Luks & P. I. Nelson – "Strategy, nontransitive dominance and the exponential distribution", *Austral. J. Statist.* **26** (1984), no. 2, p. 111–118.
[28] H. G. Kellerer – "Markov-Komposition und eine Anwendung auf Martingale", *Math. Ann.* **198** (1972), p. 99–122.
[29] ______, "Integraldarstellung von Dilationen", in *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, p. 341–374.
[30] ______, "Order conditioned independence of real random variables", *Math. Ann.* **273** (1986), p. 507–528.
[31] ______, "Markov property of point processes", *Probab. Theory Relat. Fields* **76** (1987), p. 71–80.
[32] C. Léonard – "Lazy random walks and optimal transport on graphs", *Ann. Probab.* **44** (2016), no. 3, p. 1864–1915.
[33] S. Lisini – "Characterization of absolutely continuous curves in Wasserstein spaces", *Calc. Var. Partial Differential Equations* **28** (2007), no. 1, p. 85–120.
[34] G. Lowther – "Fitting martingales to given marginals", 2008, arXiv:0808.2319.

[35] ______, "Limits of one-dimensional diffusions", *Ann. Probab.* **37** (2009), no. 1, p. 78–106.
[36] D. B. Madan & M. Yor – "Making Markov martingales meet marginals: with explicit constructions", *Bernoulli* **8** (2002), no. 4, p. 509–536.
[37] M. Nagasawa – *Schrödinger equations and diffusion theory*, Monographs in Math., vol. 86, Birkhäuser Verlag, Basel, 1993.
[38] Y. Ollivier – "Ricci curvature of Markov chains on metric spaces", *J. Funct. Anal.* **256** (2009), no. 3, p. 810–864.
[39] G. Pagès – "Functional co-monotony of processes with applications to peacocks and barrier options", in *Séminaire de Probabilités XLV*, Lect. Notes in Math., vol. 2078, Springer, Cham, 2013, p. 365–400.
[40] B. Pass – "On a class of optimal transportation problems with infinitely many marginals", *SIAM J. Appl. Math.* **45** (2013), no. 4, p. 2557–2575.
[41] S. T. Rachev & L. Rüschendorf – *Mass transportation problems. Vol. I & II*, Probability and its Appl. (New York), Springer-Verlag, New York, 1998.
[42] Y. Rinott, M. Scarsini & Y. Yu – "A Colonel Blotto gladiator game", *Math. Oper. Res.* **37** (2012), no. 4, p. 574–590.
[43] M. Shaked & J. G. Shanthikumar – *Stochastic orders*, Springer Series in Statistics, Springer, New York, 2007.
[44] V. Strassen – "The existence of probability measures with given marginals", *Ann. Math. Statist.* **36** (1965), p. 423–439.
[45] C. Villani – *Topics in optimal transportation*, Graduate Studies in Math., vol. 58, American Mathematical Society, Providence, RI, 2003.
[46] ______, *Optimal transport*, Grundlehren Math. Wiss., vol. 338, Springer-Verlag, 2009.

Manuscript received 3rd February 2020
accepted 20th October 2021

Charles Boubel, Institut de Recherche Mathématique Avancée, UMR 7501, Université de Strasbourg et CNRS
7 rue René Descartes, 67000 Strasbourg, France
E-mail : charles.boubel@unistra.fr
Url : https://irma.math.unistra.fr/~boubel/

Nicolas Juillet, Institut de Recherche Mathématique Avancée, UMR 7501, Université de Strasbourg et CNRS
7 rue René Descartes, 67000 Strasbourg, France
E-mail : nicolas.juillet@uha.fr
Url : https://juillet.perso.math.cnrs.fr/