ON ABSOLUTELY CONTINUOUS CURVES IN THE WASSERSTEIN SPACE OVER $\mathbb{R}$ AND THEIR REPRESENTATION BY AN OPTIMAL MARKOV PROCESS

CHARLES BOUBEL[†] AND NICOLAS JUILLET[†]

ABSTRACT. Let $\mu = (\mu_t)_{t\in\mathbb{R}}$ be a 1-parameter family of probability measures on $\mathbb{R}$. In [13] we introduced its "Markov-quantile" process: a process $X = (X_t)_{t\in\mathbb{R}}$ that resembles at most the quantile process attached to μ, among the Markov processes attached to μ, *i.e.* whose family of marginal laws is μ.

In this article we look at the case where μ is absolutely continuous in the Wasserstein space $\mathcal{P}_2(\mathbb{R})$. Then, X is solution of a Benamou–Brenier transport problem with intermediate marginals μ_t. It provides a *Markov* Lagrangian probabilistic representation of the continuity equation, moreover the *unique* Markov process:

– obtained as a limit for the finite dimensional topology of quantile processes where the past is made independent of the future at finitely many times.

– or, alternatively, obtained as a limit of processes linearly interpolating μ. This raises new questions about ways to obtain *Markov* Lagrangian representations of the continuity equation, and to seek uniqueness properties in this framework.

1. Introduction

In [13] we introduced the "Markov-quantile" process attached to a 1-parameter family $\mu = (\mu_t)_{t\in\mathbb{R}}$ of probability measures on $\mathbb{R}$. It is a process in the broad sense, *i.e.* a 1-parameter family $(X_t)_{t\in\mathbb{R}}$ of random variables defined on the same probability space. For the distribution of $(X_t)_{t\in\mathbb{R}}$ we adopted the notation $\mathfrak{MQ}((\mu_t)_{t\in\mathbb{R}})$, or generally simply $\mathfrak{MQ}$, that is a measure on $\mathbb{R}^{\mathbb{R}}$ equipped with the product σ-field. It can be called Markov-quantile measure but, by abuse of notation, we occasionally identified it with the Markov-quantile process. As usual X_t may namely be chosen to be the projection on the coordinate of label t for the canonical probability space $\Omega = \mathbb{R}^{\mathbb{R}}$ equipped with $\mathfrak{MQ}$ itself. The Markov-quantile measure $\mathfrak{MQ}$ is characterized by the following properties:

- **(a)** μ is the family of its marginal laws, *i.e.* for each t, μ_t is the law of X_t,
- **(b)** it is Markov,
- **(c)** it resembles "as much as possible" the quantile process $\mathfrak{Q}$ attached to μ.

The meaning of (b) is recalled in Definition 1.2 and we give in Remark 3.9 a practical criterion for Markov measures. The meaning of (c) is made precise in §3.

2010 *Mathematics Subject Classification.* 60A10, 28A33, 60J25, 35Q35, 49J55, 49Q22.

The definition of quantile process $\mathfrak{Q}((\mu_t)_{t\in\mathbb{R}})$ appears in Reminder 3.6. Note that it is Markov if and only if $\mathfrak{MQ} = \mathfrak{Q}$, which for instance happens when μ_t is diffuse for every $t \in \mathbb{R}$. For all the details, we refer the reader to our initial article [13], in particular its introduction where we give an intuition of what a Markov process that is as similar as possible to the quantile process looks like.

In this article, we consider $\mathfrak{MQ}$ in a more analytical context than in [13], that of the continuity equation in connection with the dynamical optimal transport theory, notably in continuation with Lisini's work [16]. We prove that $\mathfrak{MQ}$ is a Lagrangian representation of the continuity equation associated with $\mu = (\mu_t)_{t\in\mathbb{R}}$ together with its minimal vector fields $(v_t)_{t\in\mathbb{R}}$ (in the sense of the tangential structure over Wasserstein space [18, 2], loosely recalled in Theorem 7.1 (a)). The novel aspect of this result is of course the Markov property. It comes with several promising questions for which we give an account later.

We give now a few elements that are necessary to understand our Main Theorem, and state it. Then we give the outline of the article.

In [13], we dealt with *any* 1-parameter family of probability measures on $\mathbb{R}$. In this article we consider only the —nevertheless still rich— set of *continuous* curves $(\mu_t)_{t\in[0,1]} : [0,1] \to \mathcal{P}_2(\mathbb{R})$ to the Wasserstein space over $\mathbb{R}$. It provides the advantage that $\mathfrak{Q}$ and $\mathfrak{MQ}$ will be identified with measures on $\mathcal{C}([0,1],\mathbb{R})$ (see Notation 1.1 just below). The reader may already have noticed another (secondary) difference: in this article the time set is $[0,1]$.

Notation 1.1. **(a)** For every Polish (i.e. complete and separable) metric space $(\mathcal{X}, d)$ we denote by $\mathcal{C}([0,1],\mathcal{X})$ the space of continuous curves from $[0,1]$ to $\mathcal{X}$ —or simply by $\mathcal{C}$ specially when $\mathcal{X} = \mathbb{R}$—, with the σ-algebra induced by the topology of $\|\cdot\|_\infty$. We are interested on $\mathcal{P}(\mathcal{C})$ that is the space of probability measures on it and we denote by $\mathrm{Marg}_{\mathcal{C}}(\mu)$ the subset $\{\Gamma \in \mathcal{P}(\mathcal{C}) : \Gamma^t = \mu_t \text{ for every } t \in [0,1]\}$ where Γ^t is Γ pushed forward by the map $\gamma \in \mathcal{C} \to \gamma(t)$. The convergence we consider on $\mathcal{P}(\mathcal{C})$ is the convergence in distribution of Probability Theory, i.e. $\Gamma_n \to \Gamma \in \mathcal{P}(\mathcal{C})$ if and only if $\int f d\Gamma_n \to \int f d\Gamma$ for any bounded and continuous function f defined on $\mathcal{C}$. Be cautious that the same definition, applied to the case where the Γ_n and Γ are considered in $\mathcal{X}^{[0,1]}$ endowed with the product topology, leads to the "finite dimensional convergence" (that of all the finite marginals) that was considered in [13, Reminder 1.1], and also called "weak convergence" in that article. This last one is weaker. In the present article we will specifiy "finite dimensional" for the convergence in $\mathcal{X}^{[0,1]}$; otherwise we mean the weak convergence in $\mathcal{P}(\mathcal{C})$.

(b) For every Polish $\mathcal{X}$ we denote by $\mathcal{P}_2(\mathcal{X})$ the 2-Wasserstein space $\{\mu \in \mathcal{P}(\mathcal{X}) : \int d(x,x_0)^2 \,\mathrm{d}\mu(x) < \infty\}$ over $\mathcal{X}$ (here x_0 is some and in fact any point of $\mathcal{X}$.) The so-called 2-Wasserstein distance W_2 defined on $\mathcal{P}_2(\mathcal{X})$ is recalled in Reminder 4.5.

The Markov property is a classical notion; though as it plays a central role in this article we recall its definition.

Definition 1.2 (Markov measure and Markov process)**.** Let I be an interval and $(X_t)_{t\in I}$ be a process of law Γ. The measure Γ is *Markov* if X is a Markov process in the usual sense, for which one of the formulation is:

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

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 $u \leqslant s$. (In this case (1) is satisfied by any process X' of law Γ).

In our Main Theorem we use also the following notion, precisely built in Definition 3.8. We associate, with any process measure Γ, the distribution "Γ made Markov at a finite set $R \subset \mathbb{R}$ of times", denoted by $\Gamma_{[R]}$. For any interval I disjoint of R, the restrictions to I of (the canonical processes associated with) Γ and $\Gamma_{[R]}$ coincide. But for any two times $s < t$ separated by an element of R, the marginals at times s and t are independent. More generally the future of any $r \in R$ is made independent of its past. With this operation Remark 3.9 also provides a tractable characterization of the Markov measures that is fundamental in this paper.

The (kinetic) energy $\mathcal{E}(\gamma)$ of a curve $\gamma : [0,1] \mapsto \mathcal{X}$ in a metric space $(\mathcal{X}, d)$ may be introduced as follows:

$$\mathcal{E} : \gamma \in \mathcal{C}([0,1], \mathcal{X}) \mapsto \sup_R \sum_{k=0}^{m} \frac{d(\gamma(r_k), \gamma(r_{k+1}))^2}{r_{k+1} - r_k} \in [0, +\infty], \tag{2}$$

where $R = \{r_1, \ldots, r_m\} \subset \,]0,1[$ and $(r_0, r_{m+1}) = (0,1)$, as well as the 2-Wasserstein distance W_2 on $\mathcal{P}_2(\mathcal{X})$, based on the distance on $\mathcal{X}$ itself. Then an inequality involving energies for curves in $\mathcal{X}$ and $\mathcal{P}(\mathcal{X})$ is proved in Remark 5.3 (for $\mathcal{X} = \mathbb{R}^d$): for all $\Gamma \in \mathcal{P}(\mathcal{C}([0,1], \mathbb{R}^d))$, if $(\mathrm{proj}^t)_\#\Gamma \in \mathcal{P}_2(\mathbb{R}^d)$ for all $t \in [0,1]$:

$$\int \mathcal{E}(\gamma)\, \mathrm{d}\Gamma(\gamma) \geqslant \mathcal{E}((\Gamma^t)_{t\in[0,1]}), \quad \text{where } \Gamma^t := (\mathrm{proj}^t)_\#\Gamma, \tag{3}$$

and where, on the left, d is the Euclidean distance on $\mathbb{R}^d$ and, on the right, d is the induced W_2 on $\mathcal{P}_2(\mathbb{R}^d)$. We defined $\mathcal{E}$ for continuous curves; actually, finite energy implies continuity: if Expression (2) is finite for some γ, then γ is continuous, see Proposition 4.4(a).

We prove two convergence results, of close types, Theorem 7.1 and 7.4. The second one involves a geodesic interpolation of a curve, that we introduce in Definition 7.2, and is therefore too technical to be stated in this introduction. The first one, our Main Theorem, may be stated immediately —in fact, in a slighlty simplified version. In its statement, the existence of Γ such that (3) is an equality was established by Lisini (even for metric spaces $\mathcal{X} \neq \mathbb{R}$). However, the existence and uniqueness of a *Markov* measure is completely new.

Main Theorem. *Let $\mu = (\mu_t)_{t\in[0,1]}$ be a curve of finite energy $\mathcal{E}(\mu)$ in $\mathcal{P}_2(\mathbb{R})$.*

(a) (Existence of a Markov representation) *There exists a* Markov *probability measure Γ in $\mathrm{Marg}_{\mathcal{C}}(\mu)$ such that* (3) *is an equality, i.e.:*

$$\int \mathcal{E}(\gamma)\, \mathrm{d}\Gamma(\gamma) = \mathcal{E}(\mu),$$

and such that there exists a nested sequence $(R_n)_{n\in\mathbb{N}}$ of finite subsets of $]0,1[$ such that $\mathfrak{Q}_{[R_n]}$ (see Definition 3.8 for this measure) converges to Γ in $\mathcal{P}(\mathcal{C})$.

(b) (Uniqueness) *If* Γ *is as in (a) then it is* $\mathfrak{MQ}$.

Note that this theorem relies on Theorems A and B and Lemma 2.19 of [13].

We stress that the Markov property was up to now not involved in the a priori rather analytic context of the dynamical Optimal Transport (OT). As explained in §1.3 of [13], we came to involve it while we were considering Kellerer's Theorem, that is nowadays mostly represented in Martingale Optimal Transport (and Peacocks) a young subfield of Optimal Transport that takes advantage of the older tradition of "classical" OT. We found it particularly interesting to bring the other way around with the Markov property a new ingredient back to the parent theory.

Outline of the article. In §2 we give a brief historical overview of the set of problems in which our results take place; this introduces the main concepts at stake and motivates our work. In §3 we gather the few elements of [13] on which the present work relies, and that are necessary to its understanding. In §4–6 we introduce the precise definitions of the notions we need, which are mostly classical, and prove the propositions leading to our two results, Theorem 7.1, a stlightly more precise version of our Main Theorem above, and Theorem 7.4. In §7 we state and prove them. Finally §8 presents some open questions raised by our 1-dimensional result: may the Markov-quantile process be generalized in any dimension and moreover provide a *Markov* minimizer in the Lagrangian form of the continuity equation?

Convention 1.3. 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 $r_1 < \ldots < r_m$, if not otherwise indicated.

2. Framework; motivation of our work in this context

As we briefly mentioned in the introduction and explain below in Reminder 4.5, quantile couplings are optimal transport plans for the quadratic cost function. This suggests that the quantile process $\mathfrak{Q}$ or even the Markov-quantile process $\mathfrak{MQ}$ could be minimizers of dynamical optimal transport problems. This is true and rather well-known for $\mathfrak{Q}$; one approach is in [19] (see also [8]). In this section we show that this also makes sense for $\mathfrak{MQ}$, and in which terms it can be formulated.

Here is the minimization problem at stake. We consider a now classical action introduced by Benamou and Brenier in the context of the incompressible Euler equations, see Definition 5.2. If $X = (X_t)_{t\in\mathbb{R}}$ is a process, its action is:

$$\mathcal{A}(X) = \mathbb{E}\int_0^1 |\dot{X}_t|^2 \,\mathrm{d}t = \int_0^1 \mathbb{E}|\dot{X}_t|^2 \,\mathrm{d}t.$$

Note however that the original definition by Benamou and Brenier involves the velocity vector fields (one usually calls it "Eulerian") while we present its "Lagrangian" dual action involving the trajectories $t \mapsto X_t$. As it will become clear in this section, this action for infinitely many marginals is simply related to the quadratic transport problem with two marginals.

The origin of this research goes back to the interpretation by Arnold in [6] of the solutions of the incompressible Euler equations on a compact Riemannian manifold as geodesic curves in the space of diffeomorphisms preserving the volume. In [9], Benamou and Brenier relaxed the minimisation problem attached to those geodesics and introduced *generalized geodesics* that are, in probabilistic terms, continuous processes $X = (X_t)_{t\in[0,1]}$ with $\mathrm{Law}(X_t) = \mathrm{Vol}$ at every time, where Vol denotes the Riemannian volume. Their minimisation property is encoded in the fact that they minimize $\mathcal{A}$ under the constraint that the marginals $\mathrm{Law}(X_t)$ and $\mathrm{Law}(X_0, X_1)$ are prescribed.

Later, see [15, 18], Otto and his coauthors discovered that the solutions of some PDEs, in particular the Fokker–Planck and porous medium equations can be thought of as curves of maximal (negative) slope for some functionals F in the space of probability measures $\mathcal{P}_2(\mathbb{R}^d)$ endowed with the 2-transport distance (alias Wasserstein distance). It catches a comprehensive picture of the infinite dimensional manifold of measures used in optimal transport, building a differential calculus on it, called "Otto calculus". In this context, the derivative of the curve $(\mu_t)_t$ at time t shall be seen as a vector field v_t of gradient type, square integrable with respect to μ_t, such that the transport (or continuity) equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu_t + \mathrm{div}(\mu_t v_t) = 0 \tag{4}$$

is satisfied. The speed of the curves of maximal slope of F is $\sqrt{\int |v_t|^2 \,\mathrm{d}\mu_t}$, which corresponds to $\sqrt{\mathbb{E}(|\dot{X}_t|^2)}$ in Benamou–Brenier's action; it has to coincide with the opposite of the slope of F at μ_t, hence the derivative of $t \mapsto F(\mu_t)$ is $-\int |v_t|^2 \,\mathrm{d}\mu_t$.

A thorough study of those questions has been conducted in the monograph [2] by Ambrosio, Gigli and Savaré (see also [11, 17, 3]) under very loose assumptions on the curve $(\mu_t)_t$ or the vector field $(v_t)_t$. They proved, in particular, that the vector field $(v_t)_t$ is uniquely determined if $(\mu_t)_t$ is absolutely continuous of order 2 (see "$\mathcal{AC}_2$" in §4). They showed also that a process minimizing the action, for prescribed marginals μ_t, exists, by using limits of solutions of mollified versions of (4). Almost every trajectory of the process is in fact solution of the Cauchy problem $\dot{X}_t = v_t(X_t)$. In a further work [16], Lisini studied, in fact in a broader framework, the $\mathcal{AC}_2$ curves of probability measures on a metric space. In this context where the continuity equation is not defined, he also proved that the action can be minimized in a Lagrangian approach, i.e. in terms of trajectories in place of vector fields.

Now here is the link with our work: In both the results by Ambrosio–Gigli–Savaré and Lisini, no statement is given on the *uniqueness* of the minimizing process $(X_t)_t$. But on $\mathbb{R}$, the Markov-quantile process turns out to be a minimizing process, which yields a canonical minimizer. That notion depends of course on the chosen criterion that makes it canonical: for instance, the quantile process is a minimizer and can also be considered canonical (see the end of (b) below for a discussion). Our criterion is as follows. In this context where $(\mu_t)_t \in \mathcal{AC}_2([0,1], \mathbb{R}^d)$, i.e. has finite energy, using Theorems A and B of [13] gives rise to the two following results when $d = 1$. The first one is a slightly enhanced version of the Main Theorem given in the introduction.

(a) Theorem 7.1 makes explicit under which assumptions and in which sense $\mathfrak{MQ}$ is a canonical minimizer of the action. The existence of such a minimizer, in any dimension d, is classical, and our work adds a uniqueness result when $d = 1$, under the assumption that it is *Markov*, and obtained as a limit of products of couplings.

(b) Theorem 7.4 obtains the process $\mathfrak{MQ}$, which is a minimizer of $\mathcal{A}$, as a limit of geodesically —meaning in this case linearly— interpolating processes belonging to $\mathfrak{Disp}_{R_n}$ (see Definition 7.2) instead of the limit of $(\mathfrak{Q}_{[R_n]})_n$ as in Theorem B of [13]. Using limits of interpolating processes is the classical way to obtain minimizers (see [22, Chapter 7], [16]) in any dimension d, so this places our work within this context. The interest of doing it is that then, our process (that exists for $d = 1$) satisfies a uniqueness property that makes sense for any d. This is different from Theorem 7.1, or from any uniqueness statement on $\mathfrak{Q}$, that both only makes sense for $d = 1$. In Section 8 we formalize open problems related to extensions of Theorem 7.4 in dimension d or in metric spaces.

We can hardly conclude this paragraph without mentioning the later developments around the so-called Brenier–Schrödinger problem (see for instance the works by Arnaudon *et al.* [5], Benamou, Carlier and Nenna [10], Baradat and Léonard [7], and the references therein) that is basically an entropic minimization problem. There the trajectories get a Brownian perturbations leading to another way to compute the action. Since the Brownian motion is Markov it is tempting to imagine that a connection with the Markov-quantile process could exist. However, until now we failed to create this connection, one major obstruction being that the measures μ_t in the family $\mu = (\mu_t)_{t\in[0,1]}$ basically have to be diffuse.

3. Quick reminder on the Markov-quantile process and a few related notions

We gather below the few main notions of [13] the present article relies on.

Notation 3.1. For all measurable space E, $\mathcal{M}(E)$ and $\mathcal{P}(E)$ are the spaces of measures and probability measures on E. If $\mathcal{T}' \subset \mathcal{T}$, $\mathrm{proj}^{\mathcal{T}'}$ is the projection $\prod_{\tau\in\mathcal{T}} E_\tau \to \prod_{\tau\in\mathcal{T}'} E_\tau$; in case $\mathcal{T} = \{\tau_1, \dots, \tau_m\}$ is finite, $\mathrm{proj}^{\tau_1,\dots,\tau_m}$ means $\mathrm{proj}^{\{\tau_1,\dots,\tau_m\}}$. When $P \in \mathcal{P}\big(\prod_{\tau\in\mathcal{T}} E_\tau\big)$ and $s < t$, P^s stands for $(\mathrm{proj}^s)_\# P$ and $P^{s,t}$ for $(\mathrm{proj}^{s,t})_\# P$, and $\mathrm{Marg}((\mu_\tau)_{\tau\in\mathcal{T}})$ denotes $\{P \in \mathcal{P}\big(\prod_{\tau\in\mathcal{T}} E_\tau\big) : \forall\tau \in \mathcal{T}, (\mathrm{proj}^\tau)_\# P = \mu_\tau\}$. When not otherwise specified, what we call the marginals of P are its marginals P^s on a single factor.

Remark 3.2. To any $\Gamma \in \mathrm{Marg}_{\mathcal{C}}((\mu_t)_t)$ corresponds $\Gamma' \in \mathrm{Marg}((\mu_t)_t)$ defined by $\Gamma'(B) = \Gamma(B \cap \mathcal{C}([0,1], \mathbb{R}^d))$ for any B in the cylindrical σ-algebra of $(\mathbb{R}^d)^{[0,1]}$. Notice that for any dense countable set D of $[0,1]$, $\Gamma'((\mathrm{proj}^D)^{-1}(\mathcal{C}(D, \mathbb{R}^d))) = 1$. Conversely, suppose that some $Q \in \mathrm{Marg}((\mu_t)_t)$ satisfies this property. Then we say that Q is "concentrated on $\mathcal{C}((\mu_t)_t)$" and there is a unique $\Gamma_Q \in \mathrm{Marg}_{\mathcal{C}}((\mu_t)_t)$ such that $\Gamma'_Q = Q$. So by a slight abuse, we will not distinguish Γ and Γ' or Q and Γ_Q. For $\Gamma \in \mathrm{Marg}_{\mathcal{C}}((\mu_t)_t)$ and R a finite subset of $\mathbb{R}$, this gives sense, e.g., to $\Gamma_{[R]}$ after Definition 3.8.

Definition 3.3. If $\sharp\mathcal{T} = 2$, i.e. if $\mu \in \mathcal{P}(E)$ and $\nu \in \mathcal{P}(E')$, a measure $P \in \mathrm{Marg}(\mu, \nu)$ is called a *transport (plan)* from μ to ν, or a *coupling* between μ and ν.

Now E stands for some Polish space and $\mathcal{B}(E)$ for the set of its Borel subsets.

Definition/Notation 3.4. A probability kernel, or kernel k from E to E' is a map $k : E \times \mathcal{B}(E') \to [0,1]$ 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 $B \in \mathcal{B}(E')$.

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.

Definition 3.5. The *stochastic order* $\preceq_{\mathrm{sto}}$ on $\mathcal{P}(\mathbb{R})$ is defined by: $\mu \preceq_{\mathrm{sto}} \nu$ if for any $x \in \mathbb{R}$, $\mu(]-\infty, x]) \geqslant \nu(]-\infty, x])$. In Notation 3.1 we say that $P^{s,t} \in \mathcal{P}(\mathbb{R}^{\{s\}} \times \mathbb{R}^{\{t\}})$ has *increasing kernel* if $x \leqslant y$ implies $k_{P^{s,t}}(x, \cdot) \preceq_{\mathrm{sto}} k_{P^{s,t}}(y, \cdot)$ and say that $P \in \mathcal{P}(\prod_{t\in\mathbb{R}} \mathbb{R}^{\{t\}})$ has increasing kernels if $P^{s,t}$ has increasing kernel for every $s < t$.

Reminder 3.6. The *quantile of level* α of a measure $\mu \in \mathcal{M}(\mathbb{R})$ 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 $(Q_\tau)_{\tau\in\mathcal{T}}$, defined on $\Omega = [0,1]$ with the Lebesgue measure, is given by $Q_t(\alpha) = x_{\mu_t}(\alpha)$, and we denote $\mathrm{Law}(Q)$ by $\mathfrak{Q} \in \mathrm{Marg}((\mu_t)_{t\in\mathcal{T}})$. In particular $\mathrm{Law}(Q_t) = \mu_t$ for every $t \in \mathcal{T}$. See Definition 3.23 of [13] for full details. We call also $\mathfrak{Q}$ the quantile process or the *quantile coupling* when $\mathcal{T}$ has cardinal 2. It is again a slight abuse since $\mathfrak{Q}$ is a measure.

Here are the parts of Theorems A and B of [13] that are used in this article; we give just below the definitions necessary to their understandig. In particular Definition 3.8 plays a key role in this work.

Theorems A+B of [13]**.** *Let $(\mu_t)_{t\in\mathbb{R}}$ be a family of probability measures on $\mathbb{R}$.*

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

(i) *$\mathfrak{MQ}$ is Markov,*

(ii) *$\mathfrak{MQ}$ has increasing kernels,*

(iii) *$\mathfrak{MQ}$ has minimal couplings (alias transports) among the measures satisfying* (i) *and* (ii)*, in the sense that it satisfies:*

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

where the minimum is taken among processes $(Y_t)_t$ satisfying (i) and (ii).

(b) *There is an increasing sequence* $(R_n)_{n\in\mathbb{N}}$ *of finite subsets of* $\mathbb{R}$ *such that* $\mathfrak{Q}_{[R_n]} \in \mathrm{Marg}((\mu_t)_{t\in\mathbb{R}})$ *converges, in the finite dimensional sense, to* $\mathfrak{MQ}$*.*

In point (a)(iii), the minimum is for the stochastic order, defined as follows.

The two following important concepts may appear unusual. We invite the interested reader to consult the corresponding parts of [13] for more details.

Definition 3.7 (See Definition 2.8 of [13])**.** If $\mu_i \in \mathcal{P}(E_i)$ 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 $R \in \mathcal{P}(\mathbb{R}^3)$ such that for every $(B_1,B_2,B_3) \in \mathcal{B}(E_1)\times\mathcal{B}(E_2)\times\mathcal{B}(E_3)$:

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

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

Definition 3.8 (See Definition 4.18 of [13])**.** If $M \in \mathrm{Marg}((\mu_t)_t)$ and if $R = \{r_1,\ldots,r_m\} \subset \mathbb{R}$ 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.

$$(\mathrm{proj}^S)_\# M_{[R]} = \underbrace{M^{s_1^0,\ldots,s_{n_0}^0,r_1} \circ M^{r_1,s_1^1,\ldots,s_{n_1}^1,r_2} \circ \ldots \circ M^{r_m,s_1^m,\ldots,s_{n_m}^m}}_{\text{(denoted immediately below by } M_S)},$$

where $S = \{s_1^0,\ldots,s_{n_0}^0,r_1,s_1^1,\ldots,s_{n_1}^1,r_2,\ldots,r_m,s_1^m,\ldots,s_{n_m}^m\}$ and where the first or last term disappears if n_0 or n_m is null, respectively. These marginals are consistent in the sense that for all finite subsets S and S' of $\mathbb{R}$, containing R, $S' \subset S \Rightarrow (\mathrm{proj}^{S'})_\# M_S = M_{S'}$. So by the Kolomogorov-Daniell theorem (see Proposition 2.12 of [13]), 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".

Remark 3.9. Let I be some interval. A process $X = (X_t)_{t\in I}$ and $\Gamma \in \mathcal{P}(\mathbb{R}^I)$ its measure; X is therefore Markov (see Definition 1.2) if and only if, for any finite subset R of I, $\Gamma_{[R]} = \Gamma$.

We finally define the composition of kernels and of couplings.

Definition 3.10 (See also §2.1 of [13])**.** Kernels k from E to E' and k' from E' to E'' can be *composed* as follows:

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

If $P \in \mathrm{Marg}(\mu,\mu')$ and $Q \in \mathrm{Marg}(\mu',\mu'')$ are transport plans, we can compose them in a similar way, getting what we call their *product*:

$$P.Q \coloneqq (\mathrm{proj}^{E\times E''})_\#(P\circ Q), \text{ so that: } k_{P.Q} = k_P.k_Q.$$

4. Absolutely continuous curves of order 2 and the Wasserstein distance

Recall Convention 1.3: "$R = \{r_1, \dots, r_m\}$" implies $r_1 < \dots < r_m$.

Definition 4.1. A *partition* of an interval $[a, b]$ is a finite subset $R = \{r_0, \dots, r_{m+1}\}$ of $[a, b]$ with $(r_0, r_{m+1}) = (a, b)$. We denote the set of partitions of $[a, b]$ by $\mathrm{Part}([a, b])$. The *mesh* $|R|$ of R is $\max_{k=0}^{m} |r_{k+1} - r_k|$.

Reminder/Notation 4.2. Let γ be a curve in $\mathcal{C}([0, 1], \mathcal{X})$. **(a)** For $0 \leqslant a < b \leqslant 1$ the (possibly infinite) length of γ on $[a, b]$ is defined as $\mathcal{L}_a^b(\gamma) = \sup_{R \in \mathrm{Part}([a,b])} \sum_{k=0}^{m} d(\gamma(r_k), \gamma(r_{k+1}))$, where $R = \{r_0, r_1, \dots, r_m, r_{m+1}\}$.

(b) The curve γ is said to be absolutely continuous if for every $\delta > 0$ there exists ε such that for any family of intervals $]a_k, b_k[$ satisfying $\sum (b_k - a_k) \leqslant \varepsilon$ it holds $\sum d(\gamma(a_k), \gamma(b_k)) \leqslant \delta$. We denote by $\mathcal{AC}([0, 1], \mathcal{X})$ the space of such curves. As explained for instance in [2], where the definition is slightly different but equivalent, these curves admit for almost every t a metric derivative which we denote by $|\dot\gamma|(t)$:

$$|\dot\gamma|(t) = \lim_{h \to 0} \frac{d(\gamma(t + h) - \gamma(t))}{h}$$

(if $\mathcal{X} = \mathbb{R}^n$ and γ is differentiable at t, this is $|\dot\gamma(t)|$, so the notation is consistent). Then $\mathcal{L}_a^b(\gamma)$ equals $\int_a^b |\dot\gamma|(t)$ and $\mathcal{L}_a^b(\gamma)$ coincides with the total variation of γ on $[a, b]$. Equivalent definitions of absolutely continuous curves are that $t \mapsto \mathcal{L}_0^t(\gamma)$ is absolutely continuous, or that there exists an integrable function $m : [0, 1] \to \mathbb{R}^+$ such that $d(\gamma(a), \gamma(b)) \leqslant \int_a^b m \, \mathrm{d}\lambda$ for every $a < b$.

(c) We also introduce the space $\mathcal{AC}_2([0, 1], \mathcal{X}) \subset \mathcal{AC}([0, 1], \mathcal{X})$ of *absolutely continuous curves* γ *of order two*, i.e. such that $\int_0^1 |\dot\gamma|^2 < +\infty$. Notice that Lipschitzian curves are absolutely continuous of order two.

Now we introduce the notion of energy and the subsequent Proposition 4.4, which seems classical but for which we could not find any reference in the literature. Similar results concerning the length, in particular for geodesic curves, can be found in [4, 2]. We will consider them as known.

Definition 4.3. Let γ be a mapping from $[0, 1]$ to a metric space $(\mathcal{X}, d)$. For $0 \leqslant a < b \leqslant 1$ the energy $\mathcal{E}_a^b(\gamma)$ of γ on $[a, b]$ is defined as:

$$\mathcal{E}_a^b(\gamma) = \sup_{R \in \mathrm{Part}([a,b])} \mathcal{E}_a^b(\gamma, R), \text{ where:} \tag{6}$$

$$\mathcal{E}_a^b(\gamma, \{r_0, \dots, r_{m+1}\}) = \sum_{k=0}^{m} d(\gamma(r_k), \gamma(r_{k+1}))^2 / (r_{k+1} - r_k).$$

Note. For $[a, b] = [0, 1]$ we may denote $\mathcal{E}_0^1$ and $\mathcal{L}_0^1$ by $\mathcal{E}$ and $\mathcal{L}$.

Proposition 4.4. *Let γ be a mapping from $[0, 1]$ to $\mathcal{X}$. Then:*

(a) *If $\mathcal{E}(\gamma) < \infty$ then γ is continuous.*

(b) **(i)** *If a partition $R' \in \mathrm{Part}([a, b])$ is finer than R, $\mathcal{E}_a^b(\gamma, R) \leqslant \mathcal{E}_a^b(\gamma, R')$.* **(ii)** *If γ is continuous, the limit $\lim_{|R| \to 0} \mathcal{E}(\gamma, R)$ is well-defined and equals $\mathcal{E}(\gamma)$.* **(iii)**

$\mathcal{E}(\gamma)$ is finite if and only if $\gamma \in \mathcal{AC}_2([0,1],\mathcal{X})$; in this case $\mathcal{E}_a^b(\gamma) = \int_a^b |\dot\gamma|^2(t)\mathrm{d}t$ for all $a < b$.

(c) *$\mathcal{E}(\gamma)$ is lower semi-continuous for the uniform convergence.*

Proof. **(a)** If $\mathcal{E}(\gamma) < \infty$, there is a bound M such that for any s and $t > s$, $\frac{d(\gamma(s),\gamma(t))^2}{|s-t|} \leqslant M$, i.e. $d(\gamma(s),\gamma(t)) \leqslant M\sqrt{|s-t|}$, which gives the result.

(b)(i) This follows from the fact that, for $\alpha, \beta > 0$ and $a + b \geqslant c$:

$$\frac{1}{\alpha}a^2 + \frac{1}{\beta}b^2 \geqslant \frac{1}{\alpha+\beta}c^2,$$

itself given by the inequality $\left(a\sqrt{\beta/\alpha} - b\sqrt{\alpha/\beta}\right)^2 \geqslant 0$.

(ii) We treat (ii) in the case $\mathcal{E}(\gamma) < \infty$, letting the reader adapt the details in the case $\mathcal{E}(\gamma) = \infty$. Take $\varepsilon > 0$ and $R = \{r_0, \ldots, r_{m+1}\} \in \mathrm{Part}([0,1])$ such that $\mathcal{E}(\gamma, R) \geqslant \mathcal{E}(\gamma) - \varepsilon$. Set $\alpha = \frac{|R|}{3}$, so that if $R' = \{r'_0, \ldots, r'_{m'+1}\} \in \mathrm{Part}([0,1])$ and $|R'| \leqslant \alpha$ then for all $k \in \{0, \ldots, m\}$, $\sharp\{i \in \mathbb{N} : r'_i \in [r_k, r_{k+1}]\} \geqslant 2$. For any $R' \in \mathrm{Part}([0,1])$ such that $|R'| \leqslant \alpha$, we denote $(\min(R' \cap [r_k, r_{k+1}]), \max(R' \cap [r_k, r_{k+1}]))$ by (r_k^+, r_{k+1}^-). Since $\lim_{(s,t)\to(r_k,r_{k+1})} \frac{d(\gamma(s),\gamma(t))^2}{t-s} = \frac{d(\gamma(r_k),\gamma(r_{k+1}))^2}{r_{k+1}-r_k}$ for all k, there is an $\alpha_1 > 0$ such that $|R'| \leqslant \min(\alpha, \alpha_1)$ ensures the second inequality below, hence (ii):

$$\mathcal{E}(\gamma, R') \geqslant \sum_{k=0}^{m} d(\gamma(r_k^+) - \gamma(r_{k+1}^-))^2/(r_k^+ - r_{k+1}^-) \geqslant \mathcal{E}(\gamma, R) - \varepsilon \geqslant \mathcal{E}(\gamma) - 2\varepsilon.$$

(iii) Notice that a similar argument as above gives the Chasles relation $\mathcal{E}_a^c(\gamma) = \mathcal{E}_a^b(\gamma) + \mathcal{E}_b^c(\gamma)$ for $a < b < c$. Then we proceed in three steps.

– First, $\mathcal{E}_0^1(\gamma) < \infty$ implies that $t \mapsto \mathcal{L}_0^t(\gamma)$ is absolutely continuous, i.e. γ is. By contradiction, assume that $\mathcal{E}_0^1(\gamma) < \infty$ and that for some $\varepsilon > 0$ and every $\delta > 0$, there exists disjoint intervals $[a_k, b_k]$ with $\sum(b_k - a_k) \leqslant \delta$ and $\sum \mathcal{L}_{a_k}^{b_k}(\gamma) > \varepsilon$. Take now $\delta < \varepsilon^2/2\mathcal{E}(\gamma)$. The convexity of the scalar square gives that $\mathcal{L}_a^b(\gamma)^2 \leqslant (b-a)\mathcal{E}_a^b(\gamma)$, hence, together with the Chasles relation, the last inequality in (7) below. The second inequality of (7) is the Cauchy-Schwarz inequality. Since (7) is a contradiction, we are done.

$$\varepsilon < \sum \mathcal{L}_{a_n}^{b_n}(\gamma) \leqslant \sqrt{\sum \mathcal{L}_{a_n}^{b_n}(\gamma)^2/(b_n - a_n) \sum (b_n - a_n)} < \varepsilon/\sqrt{2}. \tag{7}$$

– Now $\gamma \in \mathcal{AC}_2 \Rightarrow \mathcal{E}(\gamma) < \infty$. Indeed, $|\gamma(b) - \gamma(a)|^2/(b-a) \leqslant (\int_a^b |\dot\gamma|)^2/(b-a) \leqslant \int_a^b |\dot\gamma|^2$, so if $\gamma \in \mathcal{AC}_2$, $\int_0^1 |\dot\gamma|^2 < \infty$ so that $\mathcal{E}(\gamma) \leqslant \int_0^1 |\dot\gamma|^2 < \infty$.

– Finally suppose that $\mathcal{E}(\gamma) < \infty$. Then $\gamma \in \mathcal{AC}_2([0,1],\mathcal{X})$. Indeed, we showed above that $\gamma \in \mathcal{AC}$. Now take $\varepsilon > 0$ and $h \in]0, \varepsilon]$. For all h let n be an integer such that $(n+1)h \geqslant 1-\varepsilon$. Then:

$$\int_0^{1-\varepsilon} \Big(\underbrace{\frac{d(\gamma(t+h)-\gamma(t))}{h}}_{a_h}\Big)^2 \mathrm{d}t \leqslant \frac{1}{h}\int_0^h \sum_{i=0}^n \frac{d(\gamma(ih+t),\gamma((i+1)h+t)^2}{h}\mathrm{d}t$$
$$\leqslant \mathcal{E}(\gamma).$$

Now, since $\gamma \in \mathcal{AC}$, $|\dot\gamma|$ is almost surely defined, hence $\liminf_{h\to 0} a_h = |\dot\gamma|$. By the Fatou lemma we get:

$$\int_0^{1-\varepsilon} |\dot\gamma|^2(t) \leqslant \mathcal{E}(\gamma).$$

This holds for every ε, so that $\gamma \in \mathcal{AC}_2([0,1],\mathcal{X})$ and that the announced formula $\int_0^1 |\dot\gamma|^2(t) = \mathcal{E}(\gamma)$ is satisfied.

(c) This holds since $\mathcal{E}$ is a supremum of functions continuous on $\mathcal{C}$ for the uniform topology on $\mathcal{C}([0,1],\mathcal{X})$. □

Reminder 4.5. On $\mathcal{P}(\mathbb{R}^d)^2$ the following infimum (minimum by the Prokhorov Theorem) has all the properties of a distance except that it may be infinite; it is called the 2-Wasserstein distance:

$$W_2(\mu,\nu) = \min_{P\in\mathrm{Marg}(\mu,\nu)} \sqrt{\int \|y-x\|^2 \mathrm{d}P(x,y)}. \tag{8}$$

On the Wasserstein space $\mathcal{P}_2(\mathbb{R}^d) = \{\mu \in \mathcal{P}(\mathbb{R}^d) |\ \int \|x\|^2\, \mathrm{d}\mu(x) < \infty\}$, W_2 is finite, thus is a true distance. Consequently, if $(\mu_t)_t \in \mathcal{C}([0,1],\mathcal{P}_2(\mathbb{R}^d))$, then $\mathcal{L}_a^b((\mu_t)_t, R)$ and $\mathcal{E}_a^b((\mu_t)_t, R)$ are finite for every $R \in \mathrm{Part}([a,b])$.

A minimizer P of (8) is called an *optimal transport plan* between μ and ν. If $d=1$ and $W_2(\mu,\nu) < \infty$ the quantile coupling $\mathfrak{Q}(\mu,\nu)$ introduced in Reminder 3.6 is the unique optimal transport plan, see for instance [20]. Therefore, for the quantile process $\mathfrak{Q}((\mu_t)_t) \in \mathrm{Marg}((\mu_t)_t)$:

$$W_2(\mu_s,\mu_t) = \sqrt{\int |y-x|^2\, \mathrm{d}\mathfrak{Q}^{s,t}(x,y)}.$$

5. Action – expected energy of a random curve

Notation 5.1. Now μ denotes a family $(\mu_t)_{t\in[0,1]}$ of probability measures on $\mathbb{R}^d$.

Definition 5.2. If $\Gamma \in \mathcal{P}((\mathbb{R}^d)^{[0,1]})$ is concentrated (see Remark 3.2) on $\mathcal{C}([0,1],\mathbb{R}^d)$ its action $\mathcal{A}(\Gamma)$ is defined as:

$$\mathcal{A}(\Gamma) = \int_{\mathcal{C}} \mathcal{E}(\gamma)\mathrm{d}\Gamma(\gamma).$$

Before we state Proposition 5.4 about the action of $\mathfrak{MQ}$ a long remark is in order concerning the action of the elements of the approximating sequence $(\mathfrak{Q}_{[R_n]})_n$ (given by Theorem A of [13]).

Remark 5.3. **(a)** If $\mathcal{A}(\Gamma) < +\infty$, Γ is in fact concentrated on $\mathcal{AC}_2$.

(b) If Γ is a measure on $\mathcal{C}$, e.g., an element of $\mathrm{Marg}_{\mathcal{C}}(\mu)$, then:

$$\mathcal{A}(\Gamma) := \int_{\mathcal{C}} \lim_{|R|\to 0} \mathcal{E}(\gamma, R)\,\mathrm{d}\Gamma(\gamma) = \lim_{|R|\to 0} \int_{\mathcal{C}} \mathcal{E}(\gamma, R)\,\mathrm{d}\Gamma(\gamma) \tag{9}$$

because of the monotone convergence theorem: use a monotone sequence of partitions and Proposition 4.4(b).

(c) If $\Gamma \in \mathrm{Marg}_{\mathcal{C}}(\mu)$, then:

$$\mathcal{A}(\Gamma) \geqslant \mathcal{E}(\mu). \tag{10}$$

In case $d = 1$, this is an equality if (but in general not only if) $\Gamma = \mathfrak{Q}(\mu)$. Indeed:

$$\begin{aligned}\int_{\mathcal{C}} \mathcal{E}(\gamma, R)\,\mathrm{d}\Gamma(\gamma) &= \int_{\mathcal{C}} \sum_{k=1}^{m} \|\gamma(r_k) - \gamma(r_{k+1})\|^2/(r_{k+1} - r_k)\,\mathrm{d}\Gamma(\gamma) \\ &= \sum_{k=1}^{m} \left(\int_{\mathcal{C}} \|\gamma(r_k) - \gamma(r_{k+1})\|^2/(r_{k+1} - r_k)\,\mathrm{d}\Gamma(\gamma) \right) \\ &\geqslant \sum_{k=1}^{m} W_2(\mu_{r_k}, \mu_{r_{k+1}})^2/(r_{k+1} - r_k) = \mathcal{E}(\mu, R).\end{aligned} \tag{11}$$

The inequality comes from the fact that $(\mathrm{proj}^{r_k, r_{k+1}})_\# \Gamma$ is in $\mathrm{Marg}(\mu_{r_k}, \mu_{r_{k+1}})$, so that $\int_{\mathcal{C}} \|\gamma(r_k) - \gamma(r_{k+1})\|^2\,\mathrm{d}\Gamma(\gamma) \geqslant W_2(\mu_{r_k}, \mu_{r_{k+1}})^2$ with equality, when $d = 1$, if $(\mathrm{proj}^{r_k, r_{k+1}})_\# \Gamma = \mathfrak{Q}(\mu_{r_k}, \mu_{r_{k+1}})$. Now, thanks to (9), when $|R|$ tends to 0 this provides $\mathcal{A}(\Gamma) \geqslant \mathcal{E}(\mu)$, with the announced equality case.

(d) If equality occurs in (c) for some measure Γ, it holds also for any $\Gamma_{[R]}$ introduced in Definition 3.8 —hence, if $d = 1$, for the measures $\mathfrak{Q}_{[R]}(\mu)$. Indeed, consider (11) only for partitions finer than R (by Proposition 4.4(b)(i), the minimum in (6) remains the same): you get that $\mathcal{A}(\Gamma) = \mathcal{A}(\Gamma_{[R]})$.

Remark 5.3 (d) "passes to the (finite dimensional) limit" when $(R_n)_n$ is such that $\mathfrak{Q}_{[R_n]}(\mu) \underset{n\to\infty}{\longrightarrow} P$, where $P \in \mathrm{Marg}_{\mathcal{C}}(\mu)$ coincides with the Markov-quantile measure $\mathfrak{MQ}$ (in the sense of Remark 3.2). Recall that, for simplicity, depending on the context we see $\mathfrak{MQ}$ (or $\mathfrak{Q}$) as an element of $\mathrm{Marg}_{\mathcal{C}}(\mu) \subset \mathcal{P}(\mathcal{C})$ or $\mathrm{Marg}(\mu) \subset \mathcal{P}(\mathbb{R}^{[0,1]})$.

Proposition 5.4. *The Markov-quantile process* $\mathfrak{MQ} \in \mathrm{Marg}(\mu)$ *satisfies* $\mathcal{A}(\mathfrak{MQ}) = \mathcal{E}(\mu)$. *Moreover for every* $(R_n)_n$ *as in Theorem B of* [13], $(\mathfrak{Q}_{[R_n]})_n$ *tends to* $\mathfrak{MQ}$ *in* $\mathrm{Marg}_{\mathcal{C}}(\mu) \subset \mathcal{P}(\mathcal{C})$.

Proof. We need the following classical claim.

Claim. If $f : (\mathcal{C}, \|\cdot\|_\infty) \to \mathbb{R}^+$ is lower semi-continuous, then $F : \mathcal{P}(\mathcal{C}) \to \mathbb{R}$ defined by $F(\Gamma) = \int_{\mathcal{C}} f\mathrm{d}\Gamma$ is also lower semi-continuous.

To check it, take $(\Gamma_n)_{n\in\mathbb{N}^*} \in \mathcal{P}(\mathcal{C})^{\mathbb{N}^*}$ tending to some Γ_0. Then $\liminf_n F(\Gamma_n) \geqslant F(\Gamma_0)$ by Lemma 4.3 of [22], the desired claim.

The claim and Proposition 4.4 (c) give that the action $\mathcal{A} : \mathcal{P}(\mathcal{C}) \to \mathbb{R}$ is lower semi-continuous. Now take $(R_n)_{n\in\mathbb{N}^*}$ a sequence of finite subsets of $\mathbb{R}$ as given by

Theorem B of [13] and set $\Gamma_n \coloneqq \mathfrak{Q}_{[R_n]}(\mu)$. If we show that Γ_n converges weakly to $\mathfrak{MQ}$ in $\mathcal{P}(\mathcal{C})$, we will get that $\mathcal{A}(\mathfrak{MQ}) \leqslant \liminf_n \mathcal{A}(\Gamma_n)$, hence the result since $\mathcal{E}(\mu) \leqslant \mathcal{A}(\mathfrak{MQ})$ by Remark 5.3 (c) and $\mathcal{A}(\Gamma_n) = \mathcal{E}(\mu)$ for all n by Remark 5.3 (d). So let us show this. By the Chebyshev inequality, for every ε there exists $\alpha > 0$ such that, for all $n \in \mathbb{N}^*$:

$$\Gamma_n(\{\gamma \in \mathcal{C} : \mathcal{E}(\gamma) > \alpha\}) < \varepsilon \quad \text{and} \quad \Gamma_n(\{\gamma \in \mathcal{C} : |\gamma(0)| > \alpha\}) < \varepsilon.$$

Therefore $\mathcal{N} \coloneqq \{\gamma \in \mathcal{C} : \mathcal{E}(\gamma) \leqslant \alpha\} \cap \{\gamma \in \mathcal{C} : |\gamma(0)| \leqslant \alpha\}$ has Γ_n-mass greater than $1 - 2\varepsilon$ for all n. It follows from its definition that on $\mathcal{N}$, $\int_0^1 |\dot\gamma|^2$ and thus also $\int_0^1 |\gamma|^2$ are bounded, hence $\mathcal{N}$ is included in a ball of the Sobolev space $\mathcal{W}^{1,2}([0,1])$. This Banach space is compactly embedded in $\mathcal{C}$, see [14, Theorem 8.8], so that $\mathcal{N}$ is relatively compact in $\mathcal{C}$, and that thus the family $(\Gamma_n)_n$ is tight. So by the Prokhorov theorem any subsequence of $(\Gamma_n)_n$ has a (weak) limit point. But by Theorem B of [13], each finite marginal of $(\Gamma_n)_n$ tends weakly to the corresponding finite marginal of $\mathfrak{MQ}$, hence all such limit point must be $\mathfrak{MQ}$, hence $(\Gamma_n)_n$ tends weakly to $\mathfrak{MQ}$. □

6. The continuity equation

Notation 6.1. For all $t \in [0,1]$, proj^t is the projection $(\mathbb{R}^d)^{[0,1]} \to (\mathbb{R}^d)^{\{t\}} = \mathbb{R}^d$, i.e. $\mathrm{proj}^t(\gamma) = \gamma(t)$. On $\mathcal{AC}([0,1], \mathbb{R}^d)$ we also define proj^t_2 by $\mathrm{proj}^t_2(\gamma) = \dot\gamma(t)$ on the set where $\dot\gamma$ is defined and $\mathrm{proj}^t_2(\gamma) = 0$ on its (null) complement.

As defined in [2, Definition 5.4.2] we introduce the barycentric projection.

Definition 6.2. Take $\Gamma \in \mathcal{P}(\mathcal{AC}([0,1], \mathbb{R}^d))$ and for all $t \in [0,1]$ denote $(\mathrm{proj}^t)_\#\Gamma$ by μ_t, $(\mathrm{proj}^t \times \mathrm{proj}^t_2)_\#\Gamma$ by M_t and by κ_t a kernel such that $M_t = \mu_t.(\mathrm{id}, \kappa_t)$, which can also be written $\forall B, B' \in \mathcal{B}(E) \times \mathcal{B}(E'), M_t(B \times B') = \int_B \kappa_t(x, B')\mathrm{d}\mu_t(x)$. The barycentric projection of M_t is the μ_t-almost surely defined vector field u^Γ_t on $\mathbb{R}^d$ such that $u^\Gamma_t(x)$ is the barycentre of $\kappa_t(x, .)$. Alternatively, it is defined by the equation:

$$\int \langle v, u^\Gamma_t\rangle(x)\mathrm{d}\mu_t(x) = \int \langle v(x), u\rangle \mathrm{d}M_t(x,u) = \int \langle v(\gamma_t), \dot\gamma_t\rangle \mathrm{d}\Gamma(\gamma), \tag{12}$$

for every continuous bounded vector field v.

Reminder 6.3. If $(\mu_t)_t = (f_t\lambda_{\mathbb{R}^d})_t$ is a family of measures on $\mathbb{R}^d$ with density $(x,t) \mapsto f_t(x)$ smooth with compact support, a smooth vector field v_t transports the measure μ_t, in the sense that its flow Φ^t makes $\mu_t(\Phi^t(B))$ constant for any Borel set B, if and only if v_t satisfies the continuity equation:

$$\partial_t\mu_t + \mathrm{div}_{\mu_t}(v_t) = 0 \quad (\text{or } \partial_t\mu_t + \mathrm{div}(v_t\mu_t) = 0, \text{ see below}), \tag{13}$$

$\mathrm{div}_{\mu_t}(v_t)$ standing for the signed measure $\mathcal{L}_{v_t}\mu_t$, where $\mathcal{L}$ is the Lie derivative. Now (13) keeps a weak meaning in $\mathbb{R}^d \times [0,1]$ in our framework, namely:

$$\int_0^1 \int_{\mathbb{R}^d} (\partial_t\varphi(x,t) + \langle v_t(x), \nabla_x\varphi(x,t)\rangle)\,\mathrm{d}\mu_t(x)\,\mathrm{d}t = 0 \tag{14}$$

for every smooth function $\varphi : \mathbb{R}^d \times [0,1] \to \mathbb{R}$ with compact support in $\mathbb{R}^d \times]0,1[$. In (13), $\operatorname{div}_\nu v$ depends only on the product $v\nu$, so may be written $\operatorname{div}(v\nu)$. Indeed, for $g, h \in \mathcal{C}^\infty(\mathbb{R}^d)$, $\operatorname{div}_{g\nu}(hv) = (\mathrm{d}(gh).v)\nu + gh \operatorname{div}_\nu(v)$.

Reminder 6.4. An important result proved in [2] (see Theorem 8.2.1) is that for every solution (μ_t, v_t) of (13) with $\int_0^1 \int |v_t|^2 \,\mathrm{d}\mu_t \,\mathrm{d}t < \infty$ there exists Γ with:

$$\dot\gamma(t) = v_t$$

$\Gamma \otimes \lambda$-almost surely. In particular $\int_0^1 \int |v_t|^2 \,\mathrm{d}\mu_t \,\mathrm{d}t = \mathcal{A}(\Gamma)$ and therefore:

$$\int_0^1 \int |v_t|^2 \,\mathrm{d}\mu_t \,\mathrm{d}t \geqslant \mathcal{E}(\mu). \tag{15}$$

Notice that, unlike for Lipschitz ODE, Γ is not unique in general.

Proposition 6.5. **(a)** *Let Γ be a probability measure on $\mathcal{AC}([0,1], \mathbb{R}^d)$ such that $\mathcal{A}(\Gamma) < \infty$ and denote $(\mathrm{proj}^t)_\#\Gamma$ by μ_t. Then $(\mu_t, u_t^\Gamma)_{t\in[0,1]}$ (see Definition 6.2) satisfies the continuity equation* (14).

(b) *If moreover Γ minimizes $\mathcal{A}(\Gamma)$ on $\mathrm{Marg}((\mu_t)_{t\in[0,1]})$, then:*

$$\dot\gamma(t) = u_t^\Gamma,$$

$\Gamma \otimes \lambda$-almost surely. In particular Γ is concentrated on integral curves of the time-dependent vector field u_t^Γ.

Proof. **(a)** We have to show that: $\int_0^1 \int \partial_t\varphi(x,t) + \nabla_x\varphi(x,t).u_t^\Gamma \mathrm{d}\mu_t \mathrm{d}t = 0$, with φ as in (14). Let M be an upper bound for $\|\nabla_x \varphi\|$. We first consider $F : (\gamma, t) \mapsto \partial_t\big(\varphi(\gamma(t), t)\big) = \partial_t\varphi(\gamma,t) + \nabla_x\varphi(\gamma,t).\dot\gamma$, then $\|F\|^2$ is bounded by $2M^2(1 + \|\dot\gamma(t)\|^2)$ and has integral on $\mathcal{C} \times [0,1]$ bounded by $2M^2(1 + \mathcal{A}(\Gamma)) < \infty$. Thus $F \in L^2(\Gamma \otimes \lambda)$. Integrating firstly with respect to t and secondly with respect to Γ, we see that $\iint F(\gamma, t) = 0$. If we now use the Fubini theorem, with (12) we obtain the desired equality.

(b) Take κ_t the kernel given in Definition 6.2, then for almost all (x,t), $\int \|v\|^2 \mathrm{d}\kappa_t(x,v) \geqslant \|u_t^\Gamma(x)\|^2$. But by (a), (μ_t, u_t^Γ) satisfies (13) hence by Reminder 6.4, $\int_0^1 \int \|u_t^\Gamma\|^2 \,\mathrm{d}\mu_t \,\mathrm{d}t \geqslant \mathcal{E}(\mu)$. This gives the inequality below:

$$\mathcal{A}(\Gamma) = \int_0^1 \int \|\dot\gamma\|^2(t) \,\mathrm{d}\Gamma(\gamma) \,\mathrm{d}t = \int_0^1 \int \int \|v\|^2 \kappa_t(x, \mathrm{d}v) \,\mathrm{d}\mu_t(x) \,\mathrm{d}t \geqslant \mathcal{E}(\mu).$$

Now if $\mathcal{A}(\Gamma)$ is minimal, i.e. $\mathcal{A}(\Gamma) = \mathcal{E}(\mu)$, all the inequalities above are equalities, which ensures that $\dot\gamma = u_t^\Gamma$ almost surely. □

7. Our resulting theorems on $\mathfrak{MQ}$ as a minimizer in this context

In this section we state and prove our two principal results, Theorems 7.1 and 7.4. The first result, Theorem 7.1 gathers:

– well-known facts, actually true on any $\mathcal{P}_2(\mathbb{R}^d)$ for $d \geqslant 1$, namely (a) and (b)(i), i.e. the existence of measures Γ for which (10) is an equality,

– enhancements of them following from Theorems A and B of [13] and Proposition 5.4, notably the uniqueness in the *Lagrangian* statement for $d = 1$ —the uniqueness of the field v_t in (a) is classical, but recall that it does not imply that of the minimizing process Γ tangent to it.

Notice that it is not known whether the process can be chosen Markov for $d \geqslant 2$ (see Open questions in 8). Moreover $\mathfrak{Q}_{[R_n]}$ is only defined for $d = 1$.

Theorem 7.1 (Existence and uniqueness of representations)**.** *Take a curve $\mu = (\mu_t)_{t\in[0,1]}$ in Wasserstein space $\mathcal{P}_2(\mathbb{R})$ with finite energy $\mathcal{E}(\mu)$. Then:*

(a) (Eulerian statement.) *There exists a vector field v_t satisfying the continuity equation* (13) *and such that Inequality* (15)*:*

$$\int_0^1 \int |v_t|^2 \,\mathrm{d}\mu_t \,\mathrm{d}t \geqslant \mathcal{E}(\mu)$$

is an equality. This vector field is unique.

(b) (Lagrangian statement.) *There exists $\Gamma \in \mathrm{Marg}_{\mathcal{C}}(\mu)$ such that:*

(i) *Inequality* (10)*: $\mathcal{A}(\Gamma) \geqslant \mathcal{E}(\mu)$ is an equality,*

(ii) *the measure Γ is Markov,*

(iii) *it is the limit in $\mathcal{P}(\mathcal{C})$ of a sequence $(\mathfrak{Q}_{[R_n]})_n$.*

Such a Γ is unique in $\mathrm{Marg}_{\mathcal{C}}(\mu)$; it is the Markov-quantile process $\mathfrak{MQ}$.

(c) (Link between them.) *For any Γ minimizing the action, i.e. making* (10) *an equality, the curve $\gamma \in \mathcal{C}$ is Γ-almost surely a solution of the ODE:*

$$\dot{\gamma}(t) = v_t(\gamma_t),$$

for almost every time.

Proof. **(a)** With u_t^Γ given by Definition 6.2, note that $\mathcal{A}(\Gamma) = \int_0^1 \int |u_t^\Gamma|^2 \mathrm{d}\mu_t \,\mathrm{d}t$ for every Γ, so that Proposition 6.5 gives the existence of the field. Its uniqueness comes from a standard argument: if u_t and v_t satisfy (13), so does $w_t := (u_t + v_t)/2$, but if they both make (15) an equality and differ on a non-null subset, $\int_0^1 \int |w_t|^2 \,\mathrm{d}\mu_t \,\mathrm{d}t < \mathcal{E}(\mu)$, which contradicts (15).

(b) Proposition 5.4 shows that $\Gamma = \mathfrak{MQ}$ suits. By Theorem B of [13], the conditions of Theorem 7.1(b) characterize the Markov-quantile process, which ensures the uniqueness.

(c) Use Proposition 6.5(a) and the uniqueness in (a). □

To state our second result, Theorem 7.4, we need to introduce the following definition. In it, remember that an optimal transport is defined in Reminder 4.5.

Definition 7.2. Let $R = \{r_0, r_1, \ldots, r_m, r_{m+1}\}$ be a partition in $\mathrm{Part}([0,1])$. We denote by $\mathfrak{Disp}_R$ the set of measures $M \in \mathcal{P}(\mathcal{C})$ that are *dynamical transports made Markov at the points of R, and linearly* (hence in fact optimally) *interpolating $(\mu_t)_{t\in[0,1]}$ between them*, defined as follows.

(a) For each $i \in \{0, \ldots, m\}$, the coupling $M^{r_i, r_{i+1}} \in \mathrm{Marg}(\mu_{r_i}, \mu_{r_{i+1}})$ is an optimal transport plan between μ_{r_i} and $\mu_{r_{i+1}}$,

(b) for $\{\lambda_1, \ldots, \lambda_n\} \subset [0,1]$ and $i^\lambda : (x,y) \in (\mathbb{R}^d)^2 \mapsto \lambda y + (1-\lambda)x$, we have:

$$(i^{\lambda_1}, \ldots, i^{\lambda_n})_\# M^{r_i, r_{i+1}} = M^{\lambda_1 r_i + (1-\lambda_1) r_{i+1}, \ldots, \lambda_n r_i + (1-\lambda_n) r_{i+1}},$$

(c) for all finite S containing $\{r_1, \ldots, r_m\}$,

$$(\mathrm{proj}^S)_\# M = M^{s_1^0, \ldots, s_{n_0}^0, r_1} \circ M^{r_1, s_1^1, \ldots, s_{n_1}^1, r_2} \circ \ldots \circ M^{r_m, s_1^m, \ldots, s_{n_m}^m},$$

where $S = \{s_1^0, \ldots, s_{n_0}^0, r_1, s_1^1, \ldots, s_{n_1}^1, r_2, \ldots, r_m, s_1^m, \ldots, s_{n_m}^m\}$ and where the first and/or last terms disappear if n_0 and/or n_m is null.

Remark 7.3. Note that $\#\mathfrak{Disp}_R = 1$ if and only if each set $\mathrm{Marg}(\mu_{r_i}, \mu_{r_{i+1}})$, appearing in (a), contains a unique optimal transport. It is the case when $d = 1$, where $\mathrm{Marg}(\mu_{r_i}, \mu_{r_{i+1}}) = \{\mathfrak{Q}(\mu_{r_i}, \mu_{r_{i+1}})\}$, see Reminder 4.5.

Theorem 7.4. *Let d be a positive integer and $\mu = (\mu_t)_{t\in[0,1]}$ a curve of finite energy in $\mathcal{P}_2(\mathbb{R}^d)$. For every nested sequence $(R_n)_n$ of finite subsets R_n of $[0,1]$, with $R_\infty := \cup_n R_n$ dense in $[0,1]$, and $\Gamma_n \in \mathfrak{Disp}_{R_n}$ for all $n \in \mathbb{N}$, there exists $\Gamma \in \mathrm{Marg}_{\mathcal{C}}(\mu)$ that is the limit in $\mathcal{P}(\mathcal{C}([0,1], \mathbb{R}^d))$ of a subsequence of $(\Gamma_n)_n$. Moreover for every Γ obtained in this way the action $\mathcal{A}(\Gamma)$ is minimal, i.e. such that Inequality* (10) *is an equality.*

Moreover, in dimension $d = 1$, a Markov limit Γ exists and if a limit Γ is Markov, it is the Markov-quantile measure in $\mathrm{Marg}_{\mathcal{C}}((\mu_t)_{t\in[0,1]})$.

In the proof we use the following partial order on the measures on $\mathbb{R}^d$ with a given total mass. It is a generalization of $\preceq_{\mathrm{sto}}$ (Definition 3.5) for any dimension d.

Definition 7.5. If $d \in \mathbb{N}^*$ and $m \in]0, +\infty[$, following [21, Section 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, for all point $(x_1, \ldots, x_d)$ of $\mathbb{R}^d$, $\mu(]-\infty, x_1] \times \cdots \times]-\infty, x_d]) \geqslant \nu(]-\infty, x_1] \times \cdots \times]-\infty, x_d])$.

Proof. Adapting [22, Chapter 7] (written in the spirit of [12]), [16] or Proposition 5.4 to our context we obtain the first part of the theorem for every $d \geqslant 1$. This requires slight modifications that we do not detail: Villani's chapter is in fact written for geodesic curves $(\mu_t)_t$ between prescribed μ_0 and μ_1 whereas Lisini's processes are attached to curves $(\mu_t)_{t\in[0,1]}$ of finite energy but the processes of the sequence are constant on each interval between two consecutive points of the partition, whereas ours is linear. Note, as an indication, that our measures Γ_n minimize $\mathcal{A}$ in $\{\Gamma \in \mathcal{P}(\mathcal{C}([0,1], \mathbb{R}^d)) : \forall r \in R_n, \Gamma^r = \mu_r\}$, the minimum being $\mathcal{A}(\Gamma_n) = \mathcal{E}(\mu, R_n)$.

In case $d = 1$, take as before a nested sequence $(R_n)_n$ given by Theorem A of [13], then $\mathfrak{Q}_{[R_n]}$ converges to $\mathfrak{MQ}$ in $\mathcal{P}(\mathcal{C})$ by Proposition 5.4. Up to taking a subsequence, the same sequence of partitions permits $\Gamma_n \in \mathfrak{Disp}_{R_n}$ to converge to some Γ. By Definitions 3.8 and 7.2, for every $S \subset R_n$ the measure $(\mathrm{proj}^S)_\# \Gamma_n$ coincides with $(\mathrm{proj}^S)_\# \mathfrak{Q}_{[R_n]}$ and:

$$(\mathrm{proj}^S)_\# \Gamma = (\mathrm{proj}^S)_\# \mathfrak{MQ}.$$

As R_∞ is dense in $[0,1]$ and the measures are concentrated on $\mathcal{C}$ it follows that $\Gamma = \mathfrak{MQ}$. This proves the existence part in case $d = 1$

For the uniqueness statement, take as before a nested sequence $(R_n)_n$ and let Γ_n be the single element of $\mathfrak{Disp}_{R_n}$ (see Remark 7.3). Assume that $(\Gamma_n)_n$ has a Markov limit Γ. By Definitions 3.8 and 7.2, for every $S \subset R_n$ the measure $(\mathrm{proj}^S)_\#\Gamma_n$ coincides with $(\mathrm{proj}^S)_\#\mathfrak{Q}_{[R_n]}$. Using the same argument as for Proposition 5.4, up to taking a subsequence, $(\mathfrak{Q}_{[R_n]})_n$ converges to an element of $\mathrm{Marg}_{\mathcal{C}}(\mu)$ that we denote by Γ'. Hence for every $S \subset R_\infty$,

$$(\mathrm{proj}^S)_\#\Gamma = (\mathrm{proj}^S)_\#\Gamma'.$$

As R_∞ is dense in $[0,1]$ and the measures are concentrated on $\mathcal{C}$ it follows $\Gamma' = \Gamma$. Note now that for every $n \in \mathbb{N}$, the measure $\mathfrak{Q}_{[R_n]}$ has increasing kernels, which is by Lemma 2.24 of [13] a closed condition for the weak topology, so that it also holds for Γ'. Finally Γ' is a process satisfying (a)(i) and (ii) of Theorem A+B p. 7. For $s < t$, on the one hand we have $\Gamma^{s,t} \succeq_{\mathrm{lo}} \mathfrak{MQ}^{s,t}$ because the Markov-quantile measure is minimal (Theorem A+B, point (a)(iii)). On the other hand, for every $s < t$ in R_∞ we have $\mathfrak{MQ}^{s,t} \succeq_{\mathrm{lo}} \Gamma^{s,t}$ because $(\Gamma')^{s,t}$ is a limit of products of quantile couplings, and $\mathfrak{MQ}^{s,t}$ is defined in Proposition 4.16 of [13] as a supremum in this class for $\preceq_{\mathrm{lo}}$. Hence, the Markov processes $\Gamma = \Gamma'$ and $\mathfrak{MQ}$ have the same law on R_∞, thus coincide as measures on $\mathcal{C}$. □

8. Open questions: a Markov minimizer for the action in metric spaces

(a) Theorem 7.4 states the existence of a process Γ minimizing the Lagrangian action $\mathcal{A}(\Gamma)$ as well as a uniqueness part concerning $\mathfrak{MQ}$. This second statement is proved for $d = 1$. It also makes sense in $\mathbb{R}^d$ for $d \geqslant 2$ and even for geodesic Polish spaces (see the next point) but it is unknown whether it holds true. For $\mathbb{R}^d$, combined with Proposition 6.5 (b), a positive answer would in particular imply that there exists a *Markov* Lagrangian representation of the continuity equation.

(b) The action $\mathcal{A}$ and energy $\mathcal{E}$ are defined on metric spaces. Definition 7.2 can also be extended to geodesic Polish metric spaces $\mathcal{X}$ in a natural way based on processes representing the geodesics of $\mathcal{P}_2(\mathcal{X})$ as in [22, Corollary 7.22] or [1, §2.2]. The first part of Theorem 7.4 is also true in this setting; we did not prove it to avoid technicalities. However, the question of the existence of an analogue of the Markov-quantile process on such a metric space seems us very interesting from an Optimal Transport perspective.

(c) Other questions are listed in §5 of our first paper [13]. It is for instance unknown whether $\mathfrak{MQ}$ is strongly Markov. One may wonder if in $\mathbb{R}^d$ or $\mathcal{X}$ there always exists a strongly Markov process minimizing the Lagrangian action of $\mu \in \mathcal{AC}_2$ and if it would be uniquely determined by this property. A similar statement for the simple Markov property is already false for $d = 1$ as attests [13, Example 5.4].

References

[1] L. Ambrosio and N. Gigli. A user's guide of optimal transport. *preprint*, 2011.

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

[3] L. Ambrosio and G. Savaré. Gradient flows of probability measures. In *Handbook of differential equations: evolutionary equations. Vol. III*, Handb. Differ. Equ., pages 1–136. Elsevier/North-Holland, Amsterdam, 2007.

[4] L. Ambrosio and P. Tilli. *Topics on analysis in metric spaces*, volume 25 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2004.

[5] M. Arnaudon, A. B. Cruzeiro, C. Léonard, and J.-C. Zambrini. An entropic interpolation problem for incompressible viscous fluids. *Ann. Inst. Henri Poincaré Probab. Stat.*, 56(3):2211–2235, 2020.

[6] V. Arnold. Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. *Ann. Inst. Fourier (Grenoble)*, 16(fasc. 1):319–361, 1966.

[7] A. Baradat and C. Léonard. Minimizing relative entropy of path measures under marginal constraints, 2020.

[8] M. Beiglböck and C. Griessler. A land of monotone plenty. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 19(1):109–127, 2019.

[9] J.-D. Benamou and 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)*, volume 226 of *Contemp. Math.*, pages 1–11. Amer. Math. Soc., Providence, RI, 1999.

[10] J.-D. Benamou, G. Carlier, and L. Nenna. Generalized incompressible flows, multi-marginal transport and Sinkhorn algorithm. *Numer. Math.*, 142(1):33–54, 2019.

[11] P. Bernard. Young measures, superposition and transport. *Indiana Univ. Math. J.*, 57(1):247–275, 2008.

[12] P. Bernard and B. Buffoni. Optimal mass transportation and Mather theory. *J. Eur. Math. Soc. (JEMS)*, 9(1):85–121, 2007.

[13] C. Boubel and N. Juillet. The Markov-quantile process attached to a family of marginals. *arXiv:1804.10514*, 2018.

[14] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011.

[15] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.

[16] S. Lisini. Characterization of absolutely continuous curves in Wasserstein spaces. *Calc. Var. Partial Differential Equations*, 28(1):85–120, 2007.

[17] S. Maniglia. Probabilistic representation and uniqueness results for measure-valued solutions of transport equations. *J. Math. Pures Appl. (9)*, 87(6):601–626, 2007.

[18] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.

[19] B. Pass. On a class of optimal transportation problems with infinitely many marginals. *SIAM J. Math. Anal.*, 45(4):2557–2575, 2013.

[20] S. T. Rachev and L. Rüschendorf. *Mass transportation problems. Vol. I*. Probability and its Applications (New York). Springer-Verlag, New York, 1998. Theory.

[21] M. Shaked and J. G. Shanthikumar. *Stochastic orders*. Springer Series in Statistics. Springer, New York, 2007.

[22] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, 2009.

Institut de Recherche Mathématique Avancée, UMR 7501, Université de Strasbourg et CNRS, 7 rue René Descartes, 67 000 Strasbourg, France

Email address: charles.boubel@unistra.fr, nicolas.juillet@unistra.fr