On absolutely continuous curves in the Wasserstein space over $\mathbb{R}$ and their representation by an optimal Markov process

Charles Boubel[1,a], Nicolas Juillet[2,3,b]

[1]*Institut de Recherche Mathématique Avancée, UMR 7501, Université de Strasbourg et CNRS, 7 rue René Descartes, 67 000 Strasbourg, France.*
[a]*charles.boubel@math.unistra.fr*
[2]*Université de Haute-Alsace, IRIMAS UR 7499, 68 100 Mulhouse, France.* [b]*nicolas.juillet@uha.fr*
[3]*Université de Strasbourg, France*

Abstract. Let $\boldsymbol{\mu} = (\mu_t)_{t \in \mathbb{R}}$ be a 1-parameter family of probability measures on $\mathbb{R}$. In [11] we introduced its "Markov-quantile" process: a process $X = (X_t)_{t \in \mathbb{R}}$ that resembles as much as possible the quantile process attached to μ, among the Markov processes attached to μ, *i.e.* whose family of marginal laws is $\boldsymbol{\mu}$.

In this article we look at the case where $\boldsymbol{\mu}$ is absolutely continuous in the Wasserstein space $\mathcal{P}_2(\mathbb{R})$. Then X is solution of a dynamical transport problem with marginals $(\mu_t)_t$. It provides a *Markov* minimal Lagrangian probabilistic representative of $\boldsymbol{\mu}$, which is moreover unique among the processes obtained as certain types of limits: limits for the finite dimensional topology of quantile processes where the past is made independent of the future conditionally on the present at finitely many times, or limits of processes linearly interpolating $\boldsymbol{\mu}$.

This raises new questions about ways to obtain *Markov* Lagrangian representatives, and to seek uniqueness properties in this framework.

Résumé. Soit $\boldsymbol{\mu} = (\mu_t)_{t \in \mathbb{R}}$ une famille à un paramètre de mesures de probabilité sur $\mathbb{R}$. Dans [11] nous introduisons le processus "Markov-quantile" qui lui est attaché: c'est le processus $X = (X_t)_{t \in \mathbb{R}}$ qui ressemble le plus qu'il est possible au processus quantile associé à $\boldsymbol{\mu}$, parmi les processus markoviens associés à $\boldsymbol{\mu}$, c'est-à-dire dont la famille de marges est $\boldsymbol{\mu}$.

Dans cet article nous considérons le cas où $\boldsymbol{\mu}$ est absolument continue dans l'espace de Wasserstein $\mathcal{P}_2(\mathbb{R})$. Alors X est solution d'un problème de transport dynamique, de marges $(\mu_t)_t$. Il fournit un représentant probabiliste lagrangien minimal *markovien* de $\boldsymbol{\mu}$. Il est en outre unique parmi les processus obtenus comme certains types de limites : limites pour la topologie de dimension finie de processus quantiles dont le passé est rendu indépendant du futur, conditionnellement au présent, en un nombre fini d'instants, ou limites de processus interpolant linéairement $\boldsymbol{\mu}$.

Ceci soulève de nouvelles questions sur les manières d'obtenir des représentants lagrangiens *markoviens*, et de demander des propriétés d'unicité dans ce cadre.

MSC2020 subject classifications: Primary 60A10, 49Q22, 60J25; secondary 35Q35, 28A33, 49J55
Keywords: Markov process, Lagrangian action, Optimal transport

1. Introduction

In [11] we introduced the "Markov-quantile" process attached to a 1-parameter family $\boldsymbol{\mu} = (\mu_t)_{t \in \mathbb{R}}$ 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) $\boldsymbol{\mu}$ 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 $\boldsymbol{\mu}$.

The meaning of (a) is $\forall t \in \mathbb{R}, \mu_t = \mathfrak{MQ} \circ X_t^{-1}$, and that of (b) is recalled in Definition 1.2. Remark 3.6 gives a practical criterion for Markov measures. The meaning of (c) is made precise in §3. The quantile process $\mathfrak{Q}((\mu_t)_{t \in \mathbb{R}})$ is the 1-parameter family $(Q_t)_{t \in \mathbb{R}}$ of random variables on $[0,1]$ with the Lebesgue measure, defined by: $Q_t(\alpha) = x_{\mu_t(\alpha)}$, where $x_{\mu_t(\alpha)}$ is the quantile of μ_t of order $\alpha \in [0,1]$; see Reminder 2.4 for details. With this notation it is well-known that the law of (Q_{t_1}, Q_{t_2}) is the optimal transport plan for the quadratic cost between μ_{t_1} and μ_{t_2}, as we recall in Reminder 2.5.

For all the details on $\mathfrak{MQ}$, we refer the reader to our initial article [11], in particular its introduction and its §1.5 where we give an intuition of what a Markov process that is as similar as possible to the quantile process looks like, and why it is difficult to define it. However, to avoid natural confusions some comments are in order:

- When the quantile process $\mathfrak{Q}$ is Markov, we have $\mathfrak{MQ} = \mathfrak{Q}$. In fact both properties are equivalent. This happens in particular for any $\boldsymbol{\mu}$ satisfying $\mu_t(x) = 0$ for every x and t (purely non atomic measures), see [11, Remark 1.8(a)].
- If the time index set $[0,1]$ or $\mathbb{R}$ is replaced by $\mathbb{N}$ (or a finite set $R = \{r_1, \ldots, r_n\} \subset \mathbb{R}$ with $r_1 < \cdots < r_n$) there exists a trivial solution to our problem. The process attached to $(\mu_n)_{n \in \mathbb{N}}$ can be called the *quantile Markov chain* and consists in the inhomogeneous Markov chain with the quantile couplings (see Reminder 2.4 for this notion) from μ_n to μ_{n+1} used as transition kernels, see [11, Remark 1.16(b)].
- Our approach to define $\mathfrak{MQ}$ can be seen *in very far approximation* as a "discrete to continuous" procedure where we use partitions $(R_n)_n$ of the time index set $[0,1]$ and the corresponding quantile Markov chains, defined as suggested in the previous point. The difficulty in [11] is not in extracting converging subsequences but showing that an adequate choice of the sequence of partitions enables to retain the Markov property at the limit. This last property is in fact a problem as soon as $(\mu_t)_{t \in T}$ is defined on a non discrete (but possibly still countable) $T \subset [0,1]$.

In this article, we consider $\mathfrak{MQ}$ in a more analytical context than in [11], that of the dynamical optimal transport theory in duality with the continuity equation, notably in continuation with Lisini's work [17]. We prove that $\mathfrak{MQ}$ satisfies a minimality property: the expected kinetic energy of the random curve of law $\mathfrak{MQ}$ is as small as it can be for a process that exactly interpolates $\boldsymbol{\mu}$. 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, summed up in the last section.

Now 1.1 gives a few necessary elements for understanding and stating our Main Theorem, and 1.2 the article's outline.

1.1. Preliminaries and Main Theorem

In [11], 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, Remark 3.1 and Remark 3.2). 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 $(\mathcal{X}, \mathrm{d})$ 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 in $\mathcal{P}(\mathcal{C})$, that is the space of probability measures on it and we denote by $\mathrm{Marg}_{\mathcal{C}}(\boldsymbol{\mu})$ 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 $\mathrm{proj}^t : \gamma \in \mathcal{C} \to \gamma(t)$. Similarly we denote by $\mathrm{Marg}(\boldsymbol{\mu})$ the subset of $\mathcal{P}\left(\mathcal{X}^{[0,1]}\right)$ defined by $\{\Gamma \in \mathcal{P}\left(\mathcal{X}^{[0,1]}\right) : \mathrm{proj}_\#^t \Gamma = \mu_t \text{ for every } t \in [0,1]\}$. The convergence we consider on $\mathcal{P}(\mathcal{C})$ is the usual weak convergence of measures used in 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 measures Γ_n and Γ are considered in $\mathcal{X}^{[0,1]}$ endowed with the product topology is equivalent to the convergence of the finite marginals, see [11, Reminder 1.11]. Both concepts can accurately be called "weak convergence". In the present paper to avoid confusion we call the less stringent notion of convergence "finite dimensional convergence" and the convergence in $\mathcal{P}(\mathcal{C})$ "weak convergence".

(b) Joint marginals of $\Gamma \in \mathcal{P}(\mathcal{C})$ or $\Gamma \in \mathcal{P}\left(\mathcal{X}^{[0,1]}\right)$ on several indices are denoted by $(\mathrm{proj}^{r_1, \ldots, r_m})_\# \Gamma$ or $(\mathrm{proj}^R)_\# \Gamma$ for $R = (r_1, \ldots, r_m)$ where $\mathrm{proj}^{r_1, \ldots, r_m}$ is the projection map $\mathrm{proj}^{r_1, \ldots, r_m} : \gamma \mapsto (\gamma(r_1), \ldots, \gamma(r_m))$. We also adopt the shorthands Γ^R and $\Gamma^{r_1, \ldots, r_n}$. In case $\pi = \Gamma^{s,t}$ with $\mu_s = \Gamma^s$ and $\mu_t = \Gamma^t$ we take the vocabulary of Optimal transport saying that π is a *transport plan* or a *coupling* of μ_s and μ_t. The corresponding set is denoted by $\mathrm{Marg}(\mu_s, \mu_t)$. Generally we note $\mathrm{Marg}((\mu_t)_{t \in \mathcal{T}})$ the set of measures with marginals $(\mu_t)_{t \in \mathcal{T}}$. Note finally that the finite dimensional convergence $\Gamma_n \to \Gamma$ of the previous paragraph writes $(\mathrm{proj}^{r_1, \ldots, r_m})_\# \Gamma_n \to (\mathrm{proj}^{r_1, \ldots, r_m})_\# \Gamma$, for every $r_1, \ldots, r_m \in [0,1]$.

(c) For every Polish space $\mathcal{X}$ we denote by $\mathcal{P}_2(\mathcal{X})$ the 2-Wasserstein space $\{\mu \in \mathcal{P}(\mathcal{X}) : \int \mathrm{d}(x, x_0)^2 \, \mathrm{d}\mu(x) < \infty\}$ over $\mathcal{X}$ (here x_0 is some and in fact any point of $\mathcal{X}$). The distance W_2 defined on $\mathcal{P}_2(\mathcal{X})$ is recalled in Reminder 2.5.

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 formulations is:

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

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 3.5. 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 at least one element of R, the marginals at times s and t are independent knowing the value of the process at any intermediate time in R. More generally the future of any $r \in R$, conditional upon the present, is made independent of its past. With this operation Remark 3.6 also provides a tractable characterization of the Markov measures that is fundamental in this paper.

Convention 1.3. When we introduce sets $\{r_1, \ldots, r_m\}$, we mean implicitly $r_1 < \ldots < r_m$, if not otherwise indicated.

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

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

where $R = \{r_1, \ldots, r_m\} \subset \,]0,1[$ and $(r_0, r_{m+1}) = (0,1)$. Actually, finite energy implies continuity: if Expression (2) is finite, then γ is continuous (see Proposition 2.3(a)). Furthermore, $\mathcal{E}(\gamma) = \int_0^1 |\dot\gamma|^2(t)dt$ in a sense that is recalled in §2.2. This notion of energy leads to the well-known notion of action, which is central in our article:

Definition 1.4. If $\Gamma \in \mathcal{P}((\mathbb{R}^d)^{[0,1]})$ is concentrated (see Remark 3.1) on $\mathcal{C}([0,1], \mathbb{R}^d)$ its action $\mathcal{A}(\Gamma)$ is defined as:

$$
\mathcal{A}(\Gamma) = \int_{\mathcal{C}([0,1], \mathbb{R}^d)} \mathcal{E}(\gamma) \mathrm{d}\Gamma(\gamma).
$$

The action satisfies the following classical inequality involving energies for curves in $\mathbb{R}^d$ and $\mathcal{P}(\mathbb{R}^d)$; its proof will be recalled in Remark 2.6: 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 $t \in [0,1]$,

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

where the distance d involved in the definition of $\mathcal{A}$, through (2), is the Euclidean distance on $\mathbb{R}^d$, and the distance involved in that of $\mathcal{E}$, on the right side, the Wasserstein distance W_2 induced by d on $\mathcal{P}_2(\mathbb{R}^d)$. Note that a central aspect of the present paper is the analysis of how the equality case in (3) can occur.

Definition 1.5. We call "minimal Lagrangian representative" of $\boldsymbol{\mu}$ a measure $\Gamma \in \mathrm{Marg}_{\mathcal{C}}(\boldsymbol{\mu})$ such that (3) is an equality.

We prove two convergence results, of close types, Theorem 4.3, *i.e.*, our Main Theorem, and Theorem 4.6, both presented in §4. Our Main Theorem comes as a refinement of well-known results on minimal representatives attached to a curve $(\mu_t)_t$ that are gathered in Theorem 2.1 ; it rests on our building of the Markov-quantile process and gives the existence of a *Markov* minimal Lagrangian representative, which is completely new and is the main point of the present article. This gives naturally rise to the question of whether a *Markov* process is unique among the minimal Lagrangian representatives. The answer is no, see Example 1.6. Though we state a weak result of uniqueness: it follows from the uniqueness of $\mathfrak{MQ}$ proved in [11] that such a Markov representative is unique among the limits of measures of quantile processes made Markov "at a finite set of times" in $[0,1]$.

Main Theorem (*i.e.*, Theorem 4.3). *Let* $\boldsymbol{\mu} = (\mu_t)_{t \in [0,1]}$ *be a curve of finite energy* $\mathcal{E}(\boldsymbol{\mu}) < \infty$ *in* $\mathcal{P}_2(\mathbb{R})$.

(a) (Existence of a Markov minimal Lagrangian representative) *There exists a minimal Lagrangian representative* Γ *for* $\boldsymbol{\mu}$, i.e., *such that* (3) *is an equality, namely* $\int \mathcal{E}(\gamma) \, \mathrm{d}\Gamma(\gamma) = \mathcal{E}(\boldsymbol{\mu})$, *that satisfies:*

- Γ *is Markov,*
- *there exists a nested* (i.e., *increasing*) *sequence* $(R_n)_{n \in \mathbb{N}}$ *of finite subsets of* $]0,1[$ *such that* $\mathfrak{Q}_{[R_n]}$ *(see page 3 and Definition 3.5 for this measure) converges to* Γ *in* $\mathcal{P}(\mathcal{C})$.

(b) (Weak uniqueness property) *If a minimal Lagrangian representative* Γ *satisfies* both *points of (a) then it is* $\mathfrak{MQ}$.

The following example shows that in general *Markov* minimal Lagrangian representatives are not uniquely determined.

Example 1.6 (Non uniqueness for Markov minimal Lagrangian representatives, see Example 5.4 in [11]). Let μ_t be $\lambda_{[t-3/4, t-1/4]} + \frac{1}{2}\delta_0$ (where $\lambda_{[a,b]}$ is the Lebesgue measure restricted to $[a,b]$) and $\Gamma \in \mathbb{P}(\mathcal{C}([0,1], \mathbb{R}))$ be a measure concentrated on the affine trajectories defined by $\Gamma(t \mapsto 0) = 1/2$ and $\Gamma(\{t \mapsto x_0 + t : x_0 \in A\}) = \lambda_{[-3/4, 1/4]}(A)$. This measure Γ is a Lagrangian representative of the continuity equation attached to $(\mu_t)_{t \in [0,1]}$. It is a minimizer of the action under marginal constraints. It is also Markov but it is not the Markov-quantile measure.

It is an open question for us to find properties enhancing the first point of (a), *e.g.*, perhaps the *strong* Markov property, or properties alternative to the second point, to make $\mathfrak{MQ}$ the unique Markov minimal Lagrangian representative. Note that [11] provides further characterizations of the Markov-quantile measure that are based on the stochastic orders.

We end §4 with Theorem 4.6. Its statement is too technical to be given in this introduction. It obtains the process $\mathfrak{MQ}$ as the unique Markov limit of a sequence of processes that geodesically –in the sense of Optimal Transport, see Definition 4.4– interpolate $(\mu_t)_{t \in [0,1]}$, instead of the sequence $(\mathfrak{Q}_{[R_n]})_n$ in the Main Theorem. The point is that it provides a type of construction that does not rely on the quantile process and that makes sense in $\mathbb{R}^d$ for any d and still furnishes minimal Lagrangian representative. However, for $d > 1$ it is not known whether some adequate choices in the construction can make it Markov. One of our main sources of inspiration is Lisini's paper [17] whose results are similar though our construction differs from it in several points, see Remark 4.7. Lisini uses a sequence of dyadic partitions to attach a minimal Lagrangian representative to each absolutely continuous of order $p > 1$ curve $(\mu_t)_{t \in [0,1]}$ of measures on metric spaces $\mathcal{X}$ that are more general than $\mathbb{R}^d$.

We stress that the Markov property was up to now not involved in the a priori rather analytic context of the dynamical Optimal Transport. As explained in §1.3 of [11], we came to involve it while we were considering Kellerer's Theorem [16], 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" Optimal Transport. We found it particularly interesting to bring the other way around with the Markov property a new ingredient back to the parent theory.

1.2. 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 [11] on which the present work relies, and that are necessary to its understanding. Theorems 2.1 and 3.8 are good summaries of the these two prelininary sections. In §4 we state and prove Theorem 4.3, that is our Main Theorem above, and Theorem 4.6. Finally §5 presents some open questions raised by our 1-dimensional result.

2. State of the art

As we briefly mentioned at the beginning of the introduction and explain below in Reminder 2.5, quantile couplings are optimal transport plans for the quadratic cost function. This suggests that the quantile process $\mathfrak{Q}$ is a minimizer for some dynamical optimal transport problems. This is true and rather well-known ; one approach is in [20] (see also [7]). In this section we recall another standard approach that formulates optimality in terms of minimal Lagrangian representatives. Subsection §2.1 explains the framework and concludes with Theorem 2.1. In §2.2 we prepare the following with useful definitions and results on $\mathcal{E}$ and $\mathcal{A}$.

2.1. Historical framework and reminders on minimal representatives

The origin of this research goes back to the interpretation by Arnold in [5] 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 [12], Brenier relaxed the minimization problem attached to those geodesics and introduced *generalized geodesics* that are, in probabilistic terms, continuous processes $X = (X_t)_{t \in [0,1]}$ with $\mathrm{Law}(X_t)$ equal to the Riemannian volume at every time. The quantity to minimize is the action $\mathcal{A}(X) = \mathbb{E} \int_0^1 |\dot{X}_t|^2 \, \mathrm{d}t = \int_0^1 \mathbb{E} |\dot{X}_t|^2 \, \mathrm{d}t$, under the constraint that the marginals $\mathrm{Law}(X_t)$ and $\mathrm{Law}(X_0, X_1)$ are prescribed. Later, see [15, 19], 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 entropy functionals 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 continuity equation:

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

is satisfied. This special (non homogeneous) vector field is precisely the minimizer of $\iint |v_t|^2 \, \mathrm{d}\mu_t \mathrm{d}t$ among the vector fields satisfying (4), the corresponding value being $\mathcal{E}(\boldsymbol{\mu})$. A thorough study of those questions has been conducted in the

monograph [2] by Ambrosio, Gigli and Savaré (see also [3, 9, 18]) under very loose assumptions on the curve $(\mu_t)_t$ and the vector field $(v_t)_t$. They proved, in particular, that the vector field $(v_t)_t$ of minimal energy is uniquely determined if $(\mu_t)_t$ is absolutely continuous of order 2 (see "AC^2" in §2.2). 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 this process is in fact solution of the Cauchy problem $\dot{X}_t = v_t(X_t)$. Furthermore, all the minimal Lagrangian representatives $(X_t)_t$ are tangent to the minimizing vector field $v = (v_t)_t$ attached to $\boldsymbol{\mu}$. Note that in the case of smooth enough curves $(\mu_t)_t$, the vector field $(v_t)_t$ is also smooth and the minimal Lagrangian representative is uniquely determined. But in general whereas this field v is unique, no uniqueness statement is satisfied by $(X_t)_t$. In a further work [17], Lisini studied, in fact in a broader framework, the AC^2 curves of probability measures on a metric space. In this context where the continuity equation is not defined, he also proved that there exists a minimal Lagrangian representative. The following is standard and based on the works of Ambrosio–Gigli–Savaré and Lisini.

Theorem 2.1 (Existence and uniqueness for minimal representatives). *Take a curve* $\boldsymbol{\mu} = (\mu_t)_{t \in [0,1]}$ *in Wasserstein space* $\mathcal{P}_2(\mathbb{R}^d)$ *with finite energy* $\mathcal{E}(\boldsymbol{\mu})$. *Then:*

(a) (Eulerian statement) *There exists a family* $(v_t)_{t \in [0,1]}$ *of vector fields satisfying the continuity equation* (4) *and such that the inequality:*

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

becomes an equality. This family is unique.

(b) (Lagrangian statement) *There exists* $\Gamma \in \mathrm{Marg}_{\mathcal{C}}(\boldsymbol{\mu})$ *such that Inequality* (3): $\mathcal{A}(\Gamma) \geqslant \mathcal{E}(\boldsymbol{\mu})$ *is an equality.*

(c) (Link between them) *For any* Γ *minimizing the action,* i.e., *making* (3) *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.

The uniqueness of $(v_t)_{t \in [0,1]}$ in (a) encourages to seek, among the processes $(X_t)_t$ in (b), processes satisfying some additional properties, trying by that to yield uniqueness. This is what we do in Theorems 4.3 and 4.6 with $\mathfrak{M}\mathfrak{Q}$.

2.2. *Some reminders on the energy $\mathcal{E}$ of a curve of probabilities and the action $\mathcal{A}$ of a probability on curves.*

The definitions and results on curves, their energy and the Wasserstein distance recalled here are close to Brenier's paper [12, Section 3]. These reminders are required to prove later that $\mathfrak{Q}$ and $\mathfrak{M}\mathfrak{Q}$ are minimal Lagrangian representatives, what is done respectively in Remark 2.6 and Section 4, the former being rather basic the latter being new.

Reminder/Notation 2.2. Let $(\mathcal{X}, \mathrm{d})$ be a metric space and γ a curve in $\mathcal{C}([0,1], \mathcal{X})$. The curve γ is said to be absolutely continuous of order $p \geqslant 1$ and we note $\gamma \in AC^p([0,1], \mathcal{X})$ (or simply AC^p) if there exists $m \in L^p([0,1], \mathbb{R})$ such that $\mathrm{d}(\gamma(a), \gamma(b)) \leqslant \int_a^b m \, \mathrm{d}\lambda$ for every $a < b$. If $\gamma \in AC^p$, an admissible choice for m is the so-called metric derivative $|\dot{\gamma}|$ defined for almost every t by:

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

(if $(\mathcal{X}, \mathrm{d}) = (\mathbb{R}^n, \|\cdot\|)$ and γ is differentiable at t, this is $\|\dot{\gamma}(t)\|$, so the notation is consistent).

Recall that $\mathcal{E}(\gamma)$ was introduced in (2) for a curve γ parametrized on $[0,1]$. The definition extends trivially for curves on $[a,b]$. A *partition* of an interval $[a,b]$ is a finite subset $R = \{r_0, \ldots, r_{m+1}\}$ of $[a,b]$ with $(r_0, r_{m+1}) = (a,b)$. The *mesh* $|R|$ of R is $\max_{k=0}^m |r_{k+1} - r_k|$. We denote by $\mathcal{E}(\gamma, R)$ the quantity approximating $\mathcal{E}(\gamma)$ on the right-hand side in (2). The next proposition gathers well-known facts on $\mathcal{E}$.

Proposition 2.3. *Let γ be a mapping from $[a,b]$ to $\mathcal{X}$. Then:*

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

(b) (i) *If a partition R' is finer than R, $\mathcal{E}(\gamma, R) \leqslant \mathcal{E}(\gamma, R')$.* **(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 $\gamma \in AC^2([a,b], \mathcal{X})$; in this case $\mathcal{E}(\gamma) = \int_a^b |\dot{\gamma}|^2(t)\mathrm{d}t$.*

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

Remark 2.6 recalls properties of $\mathcal{A}$ introduced in Definition 1.4. Its points **(c, d)** use Reminders 2.4 and 2.5.

Reminder 2.4 (Quantiles). 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 [11] for full details. When $\mathcal{T}$ has cardinal 2, $\mathfrak{Q}$ is called the *quantile transport (plan)* or the *quantile coupling* (it is a slight abuse since couplings are *usually* random variables).

Reminder 2.5 (Optimal transport). 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:

$$(5) \qquad W_2(\mu, \nu) = \min_{P \in \mathrm{Marg}(\mu,\nu)} \sqrt{\int \|y - x\|^2 \mathrm{d}P(x,y)}.$$

On the Wasserstein space (recall Notation 1.1), W_2 is finite, thus is a true distance. A minimizer P of (5) 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 2.4 is the unique optimal transport plan, see for instance [21]. Therefore, for the quantile process $\mathfrak{Q} \in \mathrm{Marg}((\mu_t)_t)$:

$$(6) \qquad W_2(\mu_s, \mu_t) = \sqrt{\int |y - x|^2 \, \mathrm{d}\mathfrak{Q}^{s,t}(x,y)}.$$

Remark 2.6. **(a)** If $\mathcal{A}(\Gamma) < +\infty$, Γ is in fact concentrated on AC^2.
  (b) If Γ is a measure on $\mathcal{C}$, *e.g.*, an element of $\mathrm{Marg}_{\mathcal{C}}(\boldsymbol{\mu})$, then:

$$(7) \qquad \mathcal{A}(\Gamma) := \int_{\mathcal{C}} \lim_{|R| \to 0} \mathcal{E}(\gamma, R) \, \mathrm{d}\Gamma(\gamma) = \lim_{|R| \to 0} \int_{\mathcal{C}} \mathcal{E}(\gamma, R) \, \mathrm{d}\Gamma(\gamma)$$

because of the monotone convergence theorem: use a monotone sequence of partitions and Proposition 2.3(b).
  (c) If $\Gamma \in \mathrm{Marg}_{\mathcal{C}}(\boldsymbol{\mu})$, then:

$$(8) \qquad \mathcal{A}(\Gamma) \geqslant \mathcal{E}(\boldsymbol{\mu}).$$

Indeed:

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

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 $\int_{\mathcal{C}} \|\gamma(r_k) - \gamma(r_{k+1})\|^2 \, \mathrm{d}\Gamma(\gamma) \geqslant W_2(\mu_{r_k}, \mu_{r_{k+1}})^2$. Now, thanks to (7), when $|R|$ tends to 0 this provides $\mathcal{A}(\Gamma) \geqslant \mathcal{E}(\boldsymbol{\mu})$.
  (d) In dimension 1 if $\mathcal{E}(\boldsymbol{\mu}) < +\infty$ then $\mathfrak{Q}$ is a minimal Lagrangian representative: by (6) endowed in (c), equality occurs in (9) for $\Gamma = \mathfrak{Q}$, thus $\mathcal{A}(\mathfrak{Q}) = \mathcal{E}(\boldsymbol{\mu})$. We used that $\mathfrak{Q}$ is concentrated on $\mathcal{C}$. This is discussed in Remark 3.2.

3. The Markov-quantile process $\mathfrak{MQ}$ attached to μ

We gather below the main notions of [11] the present article relies on like concatanation (Definition 3.4) and measure made Markov at the times of a partition (Definition 3.5). Theorem 3.8 that concludes the section is the core of the theorems in [11]. However, let us start with an important measure theoretic remark.

Remark 3.1. As will deal with measures in $\mathrm{Marg}_{\mathcal{C}}((\mu_t)_t)$, but make use of theorems about $\mathrm{Marg}((\mu_t)_t)$, we wish to see $\mathrm{Marg}_{\mathcal{C}}((\mu_t)_t)$ as a subset of $\mathrm{Marg}((\mu_t)_t)$, *i.e.*, to give a meaning to the subset "$\{Q \in \mathrm{Marg}((\mu_t)_t) \,|\, Q(\mathcal{C}([0,1], \mathbb{R}^d)) = 1\}$", which makes no sense as $\mathcal{C}([0,1], \mathbb{R}^d)$ is not in the cylindrical σ-algebra of $(\mathbb{R}^d)^{[0,1]}$. It is classically done as follows. For any $Q \in \mathrm{Marg}((\mu_t)_t)$, we will say that Q is "*concentrated on* $\mathcal{C}$" if, for any dense countable subset D of $[0,1]$, $Q(\{f \in (\mathbb{R}^d)^{[0,1]} \,|\, f_{|D}$ is uniformly continuous$\}) = 1$; the latter subset is in the cylindrical σ-algebra of $(\mathbb{R}^d)^{[0,1]}$, as it is a countable union of countable intersections of open sets of the product topology. Notice that the uniform continuity

condition amounts to the fact that $f_{|D}$ extends as a continuous function on $[0,1]$. Then $\mathrm{Marg}_{\mathcal{C}}((\mu_t)_t)$ and the set of measures of $\mathrm{Marg}((\mu_t)_t)$ concentrated on $\mathcal{C}$ are in 1-1 correspondence as follows.

– If $\Gamma \in \mathrm{Marg}_{\mathcal{C}}((\mu_t)_t)$, you can define $\hat{\Gamma} \in \mathrm{Marg}((\mu_t)_t)$ concentrated on $\mathcal{C}$, by $\hat{\Gamma} : B \mapsto \Gamma(B \cap \mathcal{C}([0,1], \mathbb{R}^d))$.

– If $Q \in \mathrm{Marg}((\mu_t)_t)$, take any (its choice will not matter) countable dense subset D of $[0,1]$ and define $\check{Q} \in \mathrm{Marg}_{\mathcal{C}}((\mu_t)_t)$ by $\check{Q} : B \mapsto Q(\{f \in (\mathbb{R}^d)^{[0,1]} \mid \exists g \in B : f_{|D} = g_{|D}\})$. We let the reader check that, as D is countable, the latter subset is in the cylindrical σ-algebra and that, in restriction to the the set of measures concentrated on $\mathcal{C}$, the definition of $\check{Q}$ is independent of the choice of D, $Q \mapsto \check{Q}$ is injective, and $\Gamma \mapsto \hat{\Gamma}$ is its inverse function.

So by a slight abuse, we will not distinguish Γ and $\hat{\Gamma}$ or Q and $\check{Q}$. 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 3.5.

Remark 3.2. In Remark 2.6(d) we used (7) voluntarily without justification to simplify the purpose. In fact since we don't know whether $t \in [0,1] \mapsto Q_t(\alpha)$ (remind Reminder 2.4) is continuous for almost every $\alpha \in [0,1]$ we need to prove that $\mathfrak{Q}$ is concentrated on $\mathcal{C}$ in the sense of Remark 3.1. In a nutshell this can be shown as follows: if $\mathcal{E}(\boldsymbol{\mu}) < +\infty$, for every increasing sequence of partitions $(R_n)_n$ with $R_\infty = \bigcup_n R_n$ dense in $[0,1]$ we have $\lim_{n \to \infty} \int_{\mathbb{R}^{[0,1]}} \mathcal{E}(\gamma, R_n) \mathrm{d}\mathfrak{Q}(\gamma) \leqslant \mathcal{E}(\boldsymbol{\mu}) < +\infty$. Therefore, by the monotone convergence theorem, $\gamma_{|R_\infty}$ has $\mathfrak{Q}$-almost surely finite energy as a mapping defined on R_∞. Since R_∞ is arbitrary chosen, this suffices to prove that $\mathfrak{Q}$ is concentrated on $\mathcal{C}$ as defined in Remark 3.1.

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

Definition/Notation 3.3. 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, for every bounded continuous function f:

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

The two following concepts may appear unusual. The interested reader is invited to consult [11] for more details.

Definition 3.4 (See [11], Definition 2.8). If $\mu_i \in \mathcal{P}(E_i)$ for $i \in \{1,2,3\}$, $P_{1,2} \in \mathrm{Marg}(\mu_1, \mu_2)$ and $P_{2,3} \in \mathrm{Marg}(\mu_2, \mu_3)$, their *concatenation* $P_{1,2} \circ P_{2,3}$ is the unique $R \in \mathcal{P}(E_1 \times E_2 \times E_3)$ such that for every $(B_1, B_2, B_3) \in \prod_{i=1}^3 \mathcal{B}(E_i)$:

$$(10) \qquad R(B_1 \times B_2 \times B_3) = \int_{x \in B_1} \int_{y \in B_2} \int_{z \in B_3} \mathrm{d}\mu_1(x) k_{1,2}(x,\mathrm{d}y) k_{2,3}(y,\mathrm{d}z).$$

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 3.5 (See Definition 4.18 of [11]). If $\Gamma \in \mathrm{Marg}((\mu_t)_t)$ and if $R = \{r_1, \ldots, r_m\} \subset \mathbb{R}$ we denote by $\Gamma_{[R]} \in \mathrm{Marg}((\mu_t)_t)$ the measure Γ *made Markov at the points of R* defined by the data of its finite marginals $(\mathrm{proj}^S)_\# \Gamma_{[R]}$, for all finite S containing R, as follows.

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

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, $S' \subset S \Rightarrow (\mathrm{proj}^{S'})_\# \Gamma_S = \Gamma_{S'}$. So by the Kolomogorov-Daniell theorem (see Proposition 2.12 of [11]), this defines $\Gamma_{[R]}$. We also commit an abuse of language: $\Gamma_{[R]}$ is rather the "*law of a process X of law Γ, made Markov at the points of R*".

Remark 3.6. 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, $\Gamma_{[R]} = \Gamma$.

Remark 3.7. Note that if Γ is concentrated on $\mathrm{Marg}_{\mathcal{C}}(\boldsymbol{\mu})$ then $\Gamma_{[R]}$ is also concentrated on $\mathrm{Marg}_{\mathcal{C}}(\boldsymbol{\mu})$.

Here are the parts of Theorems A and B of [11] that are used in this article.

Theorem 3.8 (From the main theorems in [11]). *There exists one and only one Markov law $\mathfrak{MQ}$ that is a limit in the finite-dimensional sense of sequences of laws of type $(\mathfrak{Q}_{[R_n]})$, being (R_n) an increasing sequence of finite subsets of $\mathbb{R}$. Moreover, one can assume that $R_\infty = \cup_n R_n$ is dense in $\mathbb{R}$.*

Proof. The existence of such an increasing sequence $(R_n)_{n \in \mathbb{N}}$ such that $\mathfrak{Q}_{[R_n]}$ converges to $\mathfrak{MQ}$ in the finite-dimensional sense comes from [11, Theorem B] (where the finite-dimensional convergence is called *weak* convergence). The uniqueness comes from the uniqueness of $\mathfrak{MQ}$ as a Markov measure satisfying (iv) in [11, Theorem A]. The density statement comes from (c)(i) in [11, Theorem 4.21] that is a more detailed version of Theorem B. $\square$

Remark 3.9. The claim of page 1 that $\mathfrak{M}\mathfrak{Q}$ resembles as much as possible the quantile process $\mathfrak{Q}$ attached to $\boldsymbol{\mu}$ clearly appears in Theorem 3.8. It is also strengthened from the side of the stochastic orders by (a)(iii) of Theorem A in [11].

4. Our resulting theorems on $\mathfrak{M}\mathfrak{Q}$ as a minimizer in this context

In this section we state and prove our theorems. In Lemma 4.1 and Proposition 4.2 we pursue our investigation on $\mathfrak{Q}$ started in Remark 2.6 with new results on $\mathfrak{Q}_{[R]}$ and $\mathfrak{M}\mathfrak{Q}$, respectively. Then we prove Theorem 4.3 and Theorem 4.6.

Lemma 4.1 ($\mathfrak{Q}$ and $\mathfrak{Q}_{[R]}$ are minimal Lagrangian representatives). *Let $\boldsymbol{\mu} = (\mu_t)_{t \in [0,1]}$ be a family of real measures in $\mathcal{P}_2(\mathbb{R})$ and $\mathfrak{Q}$ the attached quantile process. We assume that $\mathfrak{Q}$ is concentrated on $\mathcal{C}$ so that $\mathcal{A}(\mathfrak{Q})$ makes sense (this happens as soon as $\mathcal{E}(\boldsymbol{\mu}) < \infty$, recall Remark 3.2). Let R be a partition of $[0,1]$. Then $\mathcal{A}(\mathfrak{Q}_{[R]}) = \mathcal{A}(\mathfrak{Q})\,(= \mathcal{E}(\boldsymbol{\mu}) \in [0, +\infty])$.*

Proof. The equality $\mathcal{A}(\mathfrak{Q}) = \mathcal{E}(\boldsymbol{\mu})$ is part of Remark 2.6. In fact the other equality $\mathcal{A}(\Gamma) = \mathcal{A}(\Gamma_{[R]})$ is satisfied not only for $\Gamma = \mathfrak{Q}$ but for any Γ concentrated on $\mathcal{C}([0,1], \mathbb{R}^d)$ even for $d > 1$. Please look at Remark 2.6 and consider the equality in (9) to see that for partitions $(R_n)_{n \in \mathbb{N}}$ finer than R one has $\int \mathcal{E}(\gamma, R_n) d\Gamma(\gamma) = \int \mathcal{E}(\gamma, R_n) d\Gamma_{[R]}(\gamma)$. For a sequence of such partitions, by (7) one gets:

$$\mathcal{A}(\Gamma_{[R]}) = \lim_{|R_n| \to 0} \int \mathcal{E}(\gamma, R_n) d\Gamma_{[R]}(\gamma) = \lim_{|R_n| \to 0} \int \mathcal{E}(\gamma, R_n) d\Gamma(\gamma) = \mathcal{A}(\Gamma). \qquad \square$$

Lemma 4.1 "passes to the (finite dimensional) limit" when $(R_n)_n$ is such that $\mathfrak{Q}_{[R_n]}(\boldsymbol{\mu}) \xrightarrow{n \to \infty} P$, where $P \in \mathrm{Marg}_{\mathcal{C}}(\boldsymbol{\mu})$ coincides with the Markov-quantile measure $\mathfrak{M}\mathfrak{Q}$ (in the sense of Remark 3.1). Recall that, for simplicity, depending on the context we see $\mathfrak{M}\mathfrak{Q}$ (or $\mathfrak{Q}$) as an element of $\mathrm{Marg}_{\mathcal{C}}(\boldsymbol{\mu}) \subset \mathcal{P}(\mathcal{C})$ or $\mathrm{Marg}(\boldsymbol{\mu}) \subset \mathcal{P}(\mathbb{R}^{[0,1]})$.

Proposition 4.2. *The Markov-quantile process $\mathfrak{M}\mathfrak{Q} \in \mathrm{Marg}(\boldsymbol{\mu})$ satisfies $\mathcal{A}(\mathfrak{M}\mathfrak{Q}) = \mathcal{E}(\boldsymbol{\mu})$. Moreover for every $(R_n)_n$ as in Theorem 3.8, $(\mathfrak{Q}_{[R_n]})_n$ converges weakly to $\mathfrak{M}\mathfrak{Q}$ in $\mathrm{Marg}_{\mathcal{C}}(\boldsymbol{\mu}) \subset \mathcal{P}(\mathcal{C})$.*

Proof. Let $(R_n)_n$ be a sequence of partitions of $[0,1]$ such that $\mathfrak{Q}_{[R_n]} \to \mathfrak{M}\mathfrak{Q}$ in the finite-dimensional sense as in Theorem 3.8. To get the result, it suffices to recall that $\mathcal{A}$ is lower semi-continuous (Proposition 2.3(c)) and that it is known to have compact sublevels in the weak topology, see [1, Proof of Theorem 3.3]. With Lemma 4.1, it implies that any subsequence s of $(\mathfrak{Q}_{[R_n]})_n$ admits a (weak, and hence finite-dimensional) limit point Γ_s. By uniqueness Γ_s is always $\mathfrak{M}\mathfrak{Q}$ so that we have proved $\mathfrak{Q}_{[R_n]} \to \mathfrak{M}\mathfrak{Q}$ weakly. Hence $\mathcal{A}(\mathfrak{M}\mathfrak{Q}) \leqslant \mathcal{A}(\mathfrak{Q}_{[R_n]}) = \mathcal{E}(\boldsymbol{\mu})$, so by (8), $\mathcal{A}(\mathfrak{M}\mathfrak{Q}) = \mathcal{E}(\boldsymbol{\mu})$. $\qquad \square$

Here is our Main Theorem. Notice that by Theorem 2.1(c), the random curves of the Markov-quantile process are integral curves of the minimizing vector field in Theorem 2.1(a).

Theorem 4.3 ($\mathfrak{M}\mathfrak{Q}$ is a Markov minimal Lagrangian representative). *Take a curve $\boldsymbol{\mu} = (\mu_t)_{t \in [0,1]}$ in Wasserstein space $\mathcal{P}_2(\mathbb{R})$ with finite energy $\mathcal{E}(\boldsymbol{\mu})$. There exists $\Gamma \in \mathrm{Marg}_{\mathcal{C}}(\boldsymbol{\mu})$ such that:*

(a) *Inequality (8): $\mathcal{A}(\Gamma) \geqslant \mathcal{E}(\boldsymbol{\mu})$ is an equality,*
(b) *the measure Γ is Markov,*
(c) *it is the limit in $\mathcal{P}(\mathcal{C})$ of a sequence $(\mathfrak{Q}_{[R_n]})_n$.*

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

Proof. Proposition 4.2 shows that $\Gamma = \mathfrak{M}\mathfrak{Q}$ satisfies (a)–(c). Theorem 3.8 implies uniqueness from (b) and (c). $\qquad \square$

To state our second result, Theorem 4.6, we need to introduce the following definition. In it, remember that an optimal transport plan is defined in Reminder 2.5.

Definition 4.4. Let $R = \{r_0, r_1, \ldots, r_m, r_{m+1}\}$ be a partition of $[0,1]$ and $\boldsymbol{\mu} = (\mu_t)_{t \in [0,1]} \in \mathcal{P}(\mathbb{R}^d)^{[0,1]}$. We denote by $\mathfrak{Disp}_R(\boldsymbol{\mu})$ or more simply $\mathfrak{Disp}_R$ the set of measures $\Gamma \in \mathcal{P}(\mathcal{C}([0,1], \mathbb{R}^d))$ such that: (i) conditionally on any 'present' time $r \in R$, the past is independent from the future; (ii) Γ interpolates linearly (hence in fact optimally) μ_{r_i} and $\mu_{r_{i+1}}$. The conditions for Γ to be in $\mathfrak{Disp}_{[R]}$ are more concretely the following: for each $i \in \{0, \ldots, m\}$,

(a) the coupling $\Gamma^{r_i, r_{i+1}} \in \mathrm{Marg}(\mu_{r_i}, \mu_{r_{i+1}})$ 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 $\mathrm{m}^\lambda : (x,y) \in (\mathbb{R}^d)^2 \mapsto \lambda y + (1-\lambda)x$, we have:

$$(\mathrm{m}^{\lambda_1}, \ldots, \mathrm{m}^{\lambda_n})_{\#} \Gamma^{r_i, r_{i+1}} = \Gamma^{\mathrm{m}^{\lambda_1}(r_i, r_{i+1}), \ldots, \mathrm{m}^{\lambda_n}(r_i, r_{i+1})},$$

(c) for all finite $S = \{s_1^0, \ldots, s_{n_0}^0, r_1, s_1^1, \ldots, s_{n_1}^1, r_2, \ldots, r_m, s_1^m, \ldots, s_{n_m}^m\}$ containing $R \setminus \{r_0, r_{m+1}\} = \{r_1, \ldots, r_m\}$,

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

where the first and/or last terms disappear if n_0 and/or n_m is null.

Remark 4.5. 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 2.5.

Theorem 4.6. *Let d be a positive integer and $\boldsymbol{\mu} = (\mu_t)_{t \in [0,1]}$ a curve of finite energy in $\mathcal{P}_2(\mathbb{R}^d)$. For every nested (i.e., increasing) 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 $\Gamma \in \mathrm{Marg}_\mathcal{C}(\boldsymbol{\mu})$ 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 (8) 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 $\mathrm{Marg}_\mathcal{C}((\mu_t)_{t \in [0,1]})$.

Proof. Adapting [22, Chapter 7] (written in the spirit of [10]), [17] or Proposition 4.2 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 $\mathcal{A}(\Gamma_n) = \mathcal{E}(\boldsymbol{\mu}, R_n)$.

In case $d = 1$, take the nested sequence $(R_n)_n$ given by Theorem 3.8, then $\mathfrak{Q}_{[R_n]}$ converges to $\mathfrak{MQ}$ in $\mathcal{P}(\mathcal{C})$ by Proposition 4.2. 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 3.5 and 4.4, for every $S \subset R_n$ the measure $(\mathrm{proj}^S)_\# \Gamma_n$ coincides with $(\mathrm{proj}^S)_\# \mathfrak{Q}_{[R_n]}$ so that

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

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

To establish uniqueness, take as before a nested sequence $(R_n)_n$ and let Γ_n be the single element of $\mathfrak{Disp}_{R_n}$ (see Remark 4.5). Assume that $(\Gamma_n)_n$ has a Markov limit Γ. By Definitions 3.5 and 4.4, 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 4.2, up to taking a subsequence, $(\mathfrak{Q}_{[R_n]})_n$ converges to an element of $\mathrm{Marg}_\mathcal{C}(\boldsymbol{\mu})$ 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 Γ, Γ' are concentrated on $\mathcal{C}$ we have $\Gamma' = \Gamma$. Therefore Γ' is Markov. Uniqueness in Theorem 3.8 implies $\Gamma' = \mathfrak{MQ}$. Thus $\mathfrak{MQ}$ is the unique possible Markov limit for $(\Gamma_n)_n$. $\square$

Remark 4.7. Our work differs from Lisini's paper [17] in several points. In Theorem 4.6: (i) we restrict the range of application to $\mathcal{X} = \mathbb{R}^d$, (ii) our interpolations are continuous and piecewise linear instead of piecewise constant, (iii) we consider the uniform distance between the curves and the resulting weak convergence, instead of the weak topology on $L^p([0,1], \mathcal{X})$, (iv) our partitions are adapted in order to ensure (in case $d = 1$) the Markov property at the limit while the partitions in [17] are dyadic.

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

Let us finish by mentioning possible connection of our theorems with a stream of research whose latest developments are to be found in the so-called Brenier–Schrödinger problem (see for instance the works by Arnaudon *et al.* [4], Benamou, Carlier and Nenna [8], Baradat and Léonard [6], and the references therein). In this modified problem the trajectories become diffusion trajectories with drift and the new setting comes together with a natural action functional for the quantification of large deviations. It corresponds to an entropic minimization problem over the flows (the name given there for Γ) with marginals prescribed at any times (in the basic problem, the same measure for every $t \in [0,1]$) and prescribed joint law between the terminal measures. As a referee pointed out to us the situation is even closer to the setting studied by Dawson and Gärtner [13] where, as in our situation, the last condition is not prescribed. Since the minimizer of the entropy is Markov (see [14, Section 1.4 of Chapter II]) it is tempting to figure out that some alternative approach could exist for constructing the Markov-quantile process. However, until now we failed to create this connection, one major obstruction being that the measures μ_t in the family $\boldsymbol{\mu} = (\mu_t)_{t \in [0,1]}$ apparently have to be diffuse, another related fundamental obstruction being the non stability of the Markov property for the family of processes attached to a one-parameter family of mollifiers $(\boldsymbol{\mu}^\varepsilon)_{\varepsilon > 0}$.

We gather here the main questions arising in the paper.

(a) We proved that choosing the sequence $(R_n)_n$ properly, the approach introduced by Lisini to build Lagrangian representatives converges in dimension $d = 1$ towards a Markov process, so that there exists a Markov minimal Lagrangian representative. Is it still true in higher dimension? In geodesic spaces? Also, in dimension $d = 1$, we saw that there is only one possible Markov limit for this approach, namely the Markov-quantile process. Can also this be generalized?

(b) Can $\mathfrak{MQ}$ or more general objects in Polish spaces be equivalently introduced through a large deviation approach inspired by the Schrödinger problem? See the paragraph just before.

(c) Other questions are listed in §5 of our first paper [11]. Is for instance $\mathfrak{MQ}$ a strongly Markov process? Example 1.6 shows that the simple Markov property fails to uniquely determine $\mathfrak{MQ}$ among minimal Lagrangian representatives. Can it be characterized by a more stringent stochastic property?

Aknowledgement

References

[1] AMBROSIO, L. and FIGALLI, A. (2009). Geodesics in the space of measure-preserving maps and plans. *Arch. Ration. Mech. Anal.* **194** 421–462. https://doi.org/10.1007/s00205-008-0189-2 MR2563635

[2] AMBROSIO, L., GIGLI, N. and SAVARÉ, G. (2008). *Gradient flows in metric spaces and in the space of probability measures*, second ed. *Lectures in Mathematics ETH Zürich*. Birkhäuser Verlag, Basel. MR2401600

[3] AMBROSIO, L. and SAVARÉ, G. (2007). Gradient flows of probability measures. In *Handbook of differential equations: evolutionary equations. Vol. III. Handb. Differ. Equ.* 1–136. Elsevier/North-Holland, Amsterdam. https://doi.org/10.1016/S1874-5717(07)80004-1 MR2549368 (2011b:49115)

[4] ARNAUDON, M., CRUZEIRO, A. B., LÉONARD, C. and ZAMBRINI, J.-C. (2020). An entropic interpolation problem for incompressible viscous fluids. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 2211–2235. https://doi.org/10.1214/19-AIHP1036 MR4116723

[5] ARNOLD, V. (1966). 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** 319–361. MR0202082 (34 ##1956)

[6] BARADAT, A. and LÉONARD, C. (2020). Minimizing relative entropy of path measures under marginal constraints.

[7] BEIGLBÖCK, M. and GRIESSLER, C. (2019). A land of monotone plenty. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **19** 109–127. MR3923841

[8] BENAMOU, J.-D., CARLIER, G. and NENNA, L. (2019). Generalized incompressible flows, multi-marginal transport and Sinkhorn algorithm. *Numer. Math.* **142** 33–54. https://doi.org/10.1007/s00211-018-0995-x MR3968944

[9] BERNARD, P. (2008). Young measures, superposition and transport. *Indiana Univ. Math. J.* **57** 247–275. https://doi.org/10.1512/iumj.2008.57.3163 MR2400257 (2009b:49101)

[10] BERNARD, P. and BUFFONI, B. (2007). Optimal mass transportation and Mather theory. *J. Eur. Math. Soc. (JEMS)* **9** 85–121. https://doi.org/10.4171/JEMS/74 MR2283105

[11] BOUBEL, C. and JUILLET, N. (2022). The Markov-quantile process attached to a family of marginals. *Journal de l'École polytechnique — Mathématiques* **9** 1–62. https://doi.org/10.5802/jep.177

[12] BRENIER, Y. (1989). The least action principle and the related concept of generalized flows for incompressible perfect fluids. *J. Am. Math. Soc.* **2** 225–255. https://doi.org/10.2307/1990977

[13] DAWSON, D. A. and GÄRTNER, J. (1987). Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics* **20** 247–308. https://doi.org/10.1080/17442508708833446

[14] FÖLLMER, H. (1988). Random fields and diffusion processes. Calcul des probabilités, Éc. d'Été, Saint-Flour/Fr. 1985-87, Lect. Notes Math. 1362, 101-203 (1988).

[15] JORDAN, R., KINDERLEHRER, D. and OTTO, F. (1998). The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.* **29** 1–17. https://doi.org/10.1137/S0036141096303359 MR1617171 (2000b:35258)

[16] KELLERER, H. G. (1972). Markov-Komposition und eine Anwendung auf Martingale. *Math. Ann.* **198** 99–122. MR0356250 (50 ##8721)

[17] LISINI, S. (2007). Characterization of absolutely continuous curves in Wasserstein spaces. *Calc. Var. Partial Differential Equations* **28** 85–120. https://doi.org/10.1007/s00526-006-0032-2 MR2267755 (2007k:49001)

[18] MANIGLIA, S. (2007). Probabilistic representation and uniqueness results for measure-valued solutions of transport equations. *J. Math. Pures Appl. (9)* **87** 601–626. https://doi.org/10.1016/j.matpur.2007.04.001 MR2335089

[19] OTTO, F. (2001). The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations* **26** 101–174. MR1842429 (2002j:35180)

[20] PASS, B. (2013). On a class of optimal transportation problems with infinitely many marginals. *SIAM J. Math. Anal.* **45** 2557–2575. https://doi.org/10.1137/120881427 MR3093868

[21] RACHEV, S. T. and RÜSCHENDORF, L. (1998). *Mass transportation problems. Vol. I. Probability and its Applications (New York)*. Springer-Verlag, New York. Theory. MR1619170 (99k:28006)

[22] VILLANI, C. (2009). *Optimal transport. Grundlehren der Mathematischen Wissenschaften* **338**. Springer-Verlag. MR2459454