A differential approach to the multi-marginal Schrödinger system

Guillaume Carlier, Maxime Laborde

A differential approach to the multi-marginal Schrödinger system

Guillaume Carlier* Maxime Laborde †

November 12, 2018

Abstract

We develop an elementary and self-contained differential approach, in an L^∞ setting, for well-posedness (existence, uniqueness and smooth dependence with respect to the data) for the multi-marginal Schrödinger system which arises in the entropic regularization of optimal transport problems.

Keywords: Multi-marginal Schrödinger system, local and global inverse function theorems, entropy minimization.

MS Classification: 45G15, 49K40.

1 Introduction

Multi-marginal optimal transport problems arise in various applied settings such as economics, quantum chemistry, Wasserstein barycenters... Contrary to the well-developed two-marginals theory (see the textbooks of Villani [13, 12] and Santambrogio [10]), the structure of solutions of such problems is far from being well-understood in general. This explains the need for good numerical/approximation methods among which the entropic approximation (which has its roots in the seminal paper of Schrödinger [11]) method plays a distinguished role both for its simplicity and its efficiency, see Cuturi [6],

*CEREMADE, UMR CNRS 7534, Université Paris IX Dauphine, Pl. de Lattre de Tassigny, 75775 Paris Cedex 16, FRANCE and INRIA-Paris, MOKAPLAN, carlier@ceremade.dauphine.fr

†Department of Mathematics and Statistics, McGill University, Montreal, CANADA, maxime.laborde@mcgill.ca

Benamou et al. [1]. Roughly speaking, as its name indicates, the entropic approximation strategy consists in approximating the initial optimal transport problem by the minimization of a relative entropy with respect to the Gibbs kernel associated to the transport cost. Rigorous Γ-convergence results as well as dynamic formulations for the quadratic transport cost were studied in particular by Léonard, see [8], [9] and the references therein.

At least formally, joint measures that minimize a relative entropy subject to marginal constraints have a very simple structure, their density is the reference kernel multiplied by the tensor product of potentials (which we will call Schrödinger potentials) which are constrained by the prescribed marginal conditions. However, the existence and regularity of Schrödinger potentials cannot be taken for granted as a direct consequence of Lagrange duality because of constraints qualification issues. The problem at stake is a system of nonlinear integral equations where the data are the kernel and the marginals and the unknowns are the Schrödinger potentials. In the two-marginals case, there is a very elegant contraction argument for the Hilbert projective metric which shows the well-posedness of this system, see in particular Borwein, Lewis and Nussbaum [3]. This contraction argument is constructive and gives linear-convergence of the Sinkhorn algorithm which consists in solving alternatively the two integral equations of the system. It is not obvious to us though whether this approach can be extended to the multi-marginal case (for which existence results exist but, apart from the case of finitely supported measures, rely on rather involved and abstract arguments, see for instance Borwein and Lewis [2]). Our goal is to give an elementary differential proof of the well-posedness of the Schrödinger system in an L^∞ setting.

This short paper is organized as follows. Section 2 is devoted to the presentation of the multi-marginal Schrödinger system and its variational interpretation. Section 3 deals with local invertibility whereas section 4 is devoted to global invertibility and well-posedness. Section 5 gives some further properties of the Schrödinger map.

2 Preliminaries

2.1 Data and assumptions

We are given an integer $N \geq 2$, N probability spaces $(X_i, \mathcal{F}_i, m_i)$, $i = 1, \dots, N$ and set

$$X := \prod_{i=1}^N X_i, \mathcal{F} := \bigotimes_{i=1}^N \mathcal{F}_i,\ m := \bigotimes_{i=1}^N m_i. \tag{2.1}$$

Given $i \in \{1, \ldots, N\}$, we will denote by $X_{-i} := \prod_{j \neq i}^N X_j$, $m_{-i} := \bigotimes_{j \neq i}^N m_j$ and will always identify X to $X_i \times X_{-i}$ i.e. will denote $x = (x_1, \ldots, x_N) \in X$ as $x = (x_i, x_{-i})$.

We shall denote by $L^\infty_{++}(X_i, \mathcal{F}_i, m_i)$ (respectively $L^\infty_{++}(X, \mathcal{F}, m)$) the interior of the positive cone of $L^\infty(X_i, \mathcal{F}_i, m_i)$ (respectively $L^\infty(X, \mathcal{F}, m)$) and consider a kernel $K \in L^\infty_{++}(X, \mathcal{F}, m)$ as well as densities $\mu_i \in L^\infty_{++}(X_i, \mathcal{F}_i, m_i)$ with the same total mass:

$$\int_{X_i} \mu_i \mathrm{d}m_i = \int_{X_j} \mu_j \mathrm{d}m_j, \ i, j \in \{1, \ldots, N\}. \tag{2.2}$$

Our aim is to show the well-posedness of the multi-marginal Schrödinger system: find potentials φ_i in $L^\infty(X_i, \mathcal{F}_i, m_i)$ (called Schrödinger potentials) such that for every i and m_i-almost every $x_i \in X_i$ one has:

$$\mu_i(x_i) = e^{\varphi_i(x_i)} \int_{X_{-i}} K(x_i, x_{-i}) e^{\sum_{j \neq i} \varphi_j(x_j)} \mathrm{d}m_{-j}(x_{-j}). \tag{2.3}$$

Clearly if $\varphi = (\varphi_1, \ldots, \varphi_N)$ solves (2.3) so does every family of potentials of the form $(\varphi_1 + \lambda_1, \ldots, \varphi_N + \lambda_N)$ where the λ_i's are constants with zero-sum, it is therefore natural to add as a normalization conditions to (2.3) the additional $N-1$ linear equations:

$$\int_{X_i} \varphi_i \mathrm{d}m_i = 0, \ i = 1, \ldots, N-1. \tag{2.4}$$

2.2 Variational interpretation

It is worth here recalling the origin of the Schrödinger system in terms of minimization problems with multi-marginal constraints. Given $\mu = (\mu_1, \cdots, \mu_N) \in \prod_{i=1}^N L^\infty_{++}(X_i, \mathcal{F}_i, m_i)$ satisfying (2.2), consider the entropy minimization problem

$$\inf_{q \in \Pi(\mu)} H(q|Km) \tag{2.5}$$

where $\Pi(\mu)$ is the set of measures on X having marginals $(\mu_1 m_1, \ldots, \mu_N m_N)$ (the nonemptyness of this set being guaranteed by (2.2)), Km denotes the measure (equivalent to m) having density K with respect to m and H denotes the relative entropy:

$$H(q|Km) := \begin{cases} \int_X \left(\log\left(\frac{1}{K} \frac{dq}{dm} \right) - 1 \right) \mathrm{d}q \text{ if } q \ll m \\ +\infty \text{ otherwise.} \end{cases}$$

A motivation for (2.5) is the following, when $K = e^{-\frac{c}{\varepsilon}}$ is the Gibbs kernel associated to some cost function c and $\varepsilon > 0$ is a small (temperature) parameter, then (2.5) is an approximation of the multi-marginal optimal transport problem which consists in finding a measure in $\Pi(\mu)$ making the average of the cost c minimal (see [8], [9], [5]).

At least formally, (2.5) is dual to the concave unconstrained maximization problem

$$\sup_{\varphi=(\varphi_1,\ldots,\varphi_N)} \sum_{i=1}^N \int_{X_i} \varphi_i \mu_i \mathrm{d}m_i - \int_X K(x) e^{\sum_{j=1}^N \varphi_j(x_j)} \mathrm{d}m(x) \tag{2.6}$$

and if $\varphi \in \prod_{i=1}^N L^\infty(X_i, \mathcal{F}_i, m_i)$ solves (2.6) (the point is that the existence of such a minimizer cannot be taken for granted) it is a critical point of the (differentiable) functional in (2.6) which exactly leads to the Schrödinger system (2.3). Moreover interpreting such a φ as a family of Lagrange multipliers associated to the marginal constraints in (2.5) leads to the guess that the solution q of (2.5) should be of the form $q = \gamma m$ with a density kernel γ of the form

$$\gamma(x_1, \ldots, x_N) = K(x_1, \ldots, x_N) e^{\sum_{j=1}^N \varphi_j(x_j)} \tag{2.7}$$

and the requirement that $q \in \Pi(\mu)$ also leads to (2.3). Of course, by concavity, if φ is a bounded solution of (2.3) it is a maximizer of (2.6) and $q = \gamma m$ given by (2.7) solves (2.5).

3 Local invertibility

Let us define

$$E := \left\{ \varphi := (\varphi_1, \ldots, \varphi_N) \in \prod_{i=1}^N L^\infty(X_i, \mathcal{F}_i, m_i) : \int_{X_i} \varphi_i \mathrm{d}m_i = 0, \ i = 1, \ldots, N-1 \right\}$$

which, equipped with the L^∞ norm, is a Banach space. For $\varphi = (\varphi_1, \ldots, \varphi_N) \in \prod_{i=1}^N L^\infty(X_i, \mathcal{F}_i, m_i)$ define $T(\varphi) = (T_1(\varphi), \ldots, T_N(\varphi)) \in \prod_{i=1}^N L^\infty(X_i, \mathcal{F}_i, m_i)$ by

$$T_i(\varphi)(x_i) := \int_{X^{-i}} K(x_i, x_{-i}) e^{\sum_{j=1}^N \varphi_j(x_j)} \mathrm{d}m_{-i}(x_{-i}). \tag{3.1}$$

Note that $T(E) = T(\prod_{i=1}^N L^\infty(X_i, \mathcal{F}_i, m_i)) \subset F_{++}$ where

$$F_{++} := F \cap \prod_{i=1}^N L^\infty_{++}(X_i, \mathcal{F}_i, m_i), \tag{3.2}$$

and

$$F := \left\{ \mu \in \prod_{i=1}^{N} L^{\infty}(X_i, \mathcal{F}_i, m_i) \ : \int_{X_1} \mu_1 dm_1 = \ldots = \int_{X_N} \mu_N dm_N \right\}. \quad (3.3)$$

It will also be convenient to define the map $\widetilde{T} = (\widetilde{T}_1, \ldots, \widetilde{T}_N)$ by $\widetilde{T}_i(\varphi) := \log(T_i(\varphi))$ for $\varphi = (\varphi_1, \ldots, \varphi_N) \in \prod_{i=1}^{N} L^{\infty}(X_i, \mathcal{F}_i, m_i)$ i.e.

$$\widetilde{T}_i(\varphi)(x_i) := \varphi_i(x_i) + \log\Big(\int_{X^{-i}} K(x_i, x_{-i}) e^{\sum_{j \neq i} \varphi_j(x_j)} \mathrm{d}m_{-i}(x_{-i}) \Big). \quad (3.4)$$

Let us then observe that both $\widetilde{T}$ and T are of class C^{∞}, more precisely for φ and h in $\prod_{i=1}^{N} L^{\infty}(X_i, \mathcal{F}_i, m_i)$, we have

$$\widetilde{T}_i'(\varphi)(h)(x_i) = h_i(x_i) + \frac{\int_{X_{-i}} K(x_i, x_{-i}) e^{\sum_{k \neq i} \varphi_k(x_k)} \sum_{j \neq i} h_j(x_j) \mathrm{d}m_{-i}(x_{-i})}{\int_{X_{-i}} K(x_i, x_{-i}) e^{\sum_{j \neq i} \varphi_j(x_j)} \mathrm{d}m_{-i}(x_{-i})}$$

and

$$T_i'(\varphi)(h)(x_i) = e^{\widetilde{T}_i(\varphi)(x_i)} \widetilde{T}_i'(\varphi)(h)(x_i). \quad (3.5)$$

Let us fix $\varphi := (\varphi_1, \ldots, \varphi_N) \in E$, observe that $\widetilde{T}'(\varphi)$ extends (and we still denote by $\widetilde{T}'(\varphi)$ this extension) to a bounded linear self map of $\prod_{i=1}^{N} L^2(X_i, \mathcal{F}_i, m_i)$ which is of the form

$$\widetilde{T}'(\varphi) := \mathrm{id} + L \quad (3.6)$$

with L a *compact* linear self map of $\prod_{i=1}^{N} L^2(X_i, \mathcal{F}_i, m_i)$. We then have the following:

Proposition 3.1. *Let $\varphi \in E$ then $T'(\varphi)$ is an isomorphism between E and F. In particular, T is a local C^{∞} diffeomorphism between E and F, and $T(E)$ is open in F_{++}.*

Proof. In view of (3.5), the desired invertibility claim amounts to show that $\widetilde{T}'(\varphi)$ is an isomorphism between E and F_{φ} the linear subspace of codimension $N-1$ consisting of $\theta = (\theta_1, \ldots, \theta_N) \in \prod_{i=1}^{N} L^{\infty}(X_i, \mathcal{F}_i, m_i)$ which satisfy

$$\int_{X_1} e^{\widetilde{T}_1(\varphi)} \theta_1 \mathrm{d}m_1 = \ldots = \int_{X_N} e^{\widetilde{T}_N(\varphi)} \theta_N \mathrm{d}m_N. \quad (3.7)$$

Let us also denote by $F_{\varphi,2}$ the set of all $\theta = (\theta_1, \ldots, \theta_N) \in \prod_{i=1}^{N} L^2(X_i, \mathcal{F}_i, m_i)$ which satisfy (3.7).

As noted above, one can write $\widetilde{T}'(\varphi) = \mathrm{id} + L$ on $\prod_{i=1}^{N} L^2(X_i, \mathcal{F}_i, m_i)$ with L compact. Let us define the probability measure Q_φ on X given by

$$Q_\varphi(\mathrm{d}x) = \frac{K(x) e^{\sum_{j=1}^{N} \varphi_j(x_j)} m(\mathrm{d}x)}{\int_X K(x) e^{\sum_{j=1}^{N} \varphi_j(x_j)} \mathrm{d}m(x)}. \tag{3.8}$$

For $i = 1, \dots, N$, let us now disintegrate Q_φ with respect to its i-th marginal Q_φ^i:

$$Q_\varphi(\mathrm{d}x_i, \mathrm{d}x_{-i}) = Q_\varphi^{-i}(\mathrm{d}x_{-i} | x_i) \otimes Q_\varphi^i(\mathrm{d}x_i) \tag{3.9}$$

where $Q_\varphi^{-i}(\mathrm{d}x_{-i}|x_i)$ is the conditional probability of x_{-i} given x_i according to Q_φ. The compact operator L can then conveniently be expressed in terms of the corresponding conditional expectations operators. Indeed, setting $L(h) = (L_1(h), \dots, L_N(h))$, we obviously have

$$L_i(h)(x_i) = \int_{X_{-i}} \Big(\sum_{j \neq i} h_j(x_j) \Big) Q_\varphi^{-i}(\mathrm{d}x_{-i}|x_i) \text{ for } m_i\text{-a.e. } x_i \in X_i.$$

Let $h \in \prod_{i=1}^{N} L^2(X_i, \mathcal{F}_i, m_i)$ be such that $\widetilde{T}'(\varphi)(h) = 0$ (equivalently $T'(\varphi)(h) = 0$) i.e. for every i and m_i-a.e. $x_i \in X_i$, there holds

$$h_i(x_i) = - \int_{X_{-i}} \Big(\sum_{j \neq i} h_j(x_j) \Big) Q_\varphi^{-i}(\mathrm{d}x_{-i}|x_i)$$

multiplying by $h_i(x_i)$ and then integrating with respect to Q_φ^i gives

$$\int_{X_i} h_i^2(x_i) \mathrm{d}Q_\varphi^i(x_i) = - \sum_{j, j \neq i} \int_X h_i(x_i) h_j(x_j) \mathrm{d}Q_\varphi(x)$$

summing over i thus yields

$$\begin{aligned} \int_X \Big(\sum_{i=1}^{N} h_i(x_i) \Big)^2 \mathrm{d}Q_\varphi(x) &= \sum_{i=1}^{N} \int_{X_i} h_i^2(x_i) \mathrm{d}Q_\varphi^i(x_i) + \sum_{i,j, j \neq i} \int_X h_i(x_i) h_j(x_j) \mathrm{d}Q_\varphi(x) \\ &= 0. \end{aligned}$$

Since Q_φ is equivalent to m, we deduce that $\sum_{i=1}^{N} h_i(x_i) = 0$ m-a.e. that is h is constant and its components sum to 0. Hence $\ker(\widetilde{T}'(\varphi))$ has dimension $N-1$ and $\ker(\widetilde{T}'(\varphi)) \cap E = \{0\}$ i.e. $\widetilde{T}'(\varphi)$ is one to one on E.

Since L is a compact operator of L^2 and $\ker(\mathrm{id} + L)$ has dimension $N-1$, it follows from the Fredholm alternative Theorem (see chapter VI of [4]) that $R(\mathrm{id} + L)$ has codimension $N-1$. Differentiating the relation

$$\int_{X_i} e^{\widetilde{T}_i(\varphi)} \mathrm{d}m_i = \int_{X_j} e^{\widetilde{T}_j(\varphi)} \mathrm{d}m_j, \; i, j \in \{1, \dots, N-1\}$$

gives

$$\int_{X_i} e^{\widetilde{T}_i(\varphi)} \widetilde{T}_i'(\varphi)(h) \mathrm{d}m_i = \int_{X_j} e^{\widetilde{T}_j(\varphi)} \widetilde{T}_j'(\varphi)(h) \mathrm{d}m_j, \ i, j \in \{1, \dots, N-1\}$$

i.e. $\widetilde{T}'(\varphi)(h) \in F_\varphi$ for every $h \in \prod_{i=1}^N L^\infty(X_i, \mathcal{F}_i, m_i)$. Likewise, we also have $\widetilde{T}'(\varphi)(h) \in F_{\varphi,2}$, for every $h \in \prod_{i=1}^N L^2(X_i, \mathcal{F}_i, m_i)$. Since $F_{\varphi,2}$ has codimension $N-1$, we get

$$R(\mathrm{id} + L) = \widetilde{T}'(\varphi)\Big(\prod_{i=1}^N L^2(X_i, \mathcal{F}_i, m_i)\Big) = F_{\varphi,2}. \tag{3.10}$$

In particular, for every $\theta \in F_\varphi$ there exists $h \in \prod_{i=1}^N L^2(X_i, \mathcal{F}_i, m_i)$ such that $\theta = h + L(h)$ since obviously L maps $\prod_{i=1}^N L^2(X_i, \mathcal{F}_i, m_i)$ into $\prod_{i=1}^N L^\infty(X_i, \mathcal{F}_i, m_i)$ we have $h \in \prod_{i=1}^N L^\infty(X_i, \mathcal{F}_i, m_i)$. Finally, since $\widetilde{T}'(\varphi)(h) = \widetilde{T}'(\varphi)(\widetilde{h})$ whenever $h - \widetilde{h}$ is a vector of constants summing to zero, we may also assume that $h \in E$. This shows that $\widetilde{T}'(\varphi)(E) = F_\varphi$ or equivalenty $T'(\varphi)(E) = F$.

We have shown that $T'(\varphi)$ is an isomorphism between the Banach spaces E and F, the local invertibility claim thus directly follows from the local inversion Theorem.

□

4 Global invertibility and well-posedness

To pass from local to global invertibiliy of T, we invoke classical arguments à la Caccioppoli-Hadamard (see for instance [7]). First of all, it is easy to see that T is one to one on E:

Proposition 4.1. *The map T is injective on E.*

Proof. If φ and ψ are in E and $T(\varphi) = T(\psi) := \mu$, then both φ and ψ are solutions of the maximization problem (2.6), since the functional in (2.6) is the sum of a linear term and a term that is strictly concave in the direct sum of the potentials we should have $\sum_{i=1}^N \varphi_i(x_i) = \sum_{i=1}^N \psi_i(x_i)$ which by the normalization condition in the definition of E implies that $\varphi = \psi$.

□

Next we observe that:

Lemma 4.2. *$T(E)$ is closed in F_{++}.*

Proof. Let $(\varphi^n)_n \in E^{\mathbb{N}}$ be such that $\mu^n := T(\varphi^n)$ converges in L^∞ to some $\mu \in F_{++}$. Let $\psi^n = (\varphi_1^n + \lambda_1^n, \dots, \varphi_N^n + \lambda_N^n)$ where the λ_i^n's are constant which sum to zero and chosen in such a way that

$$\int_{X_i} e^{\psi_i^n} \mathrm{d}m_i = 1, \ i = 1, \dots, N-1, \tag{4.1}$$

this ensures that $\mu^n := T(\psi^n)$ i.e. for every i and m_i- a.e. $x_i \in X_i$

$$\log(\mu_i^n(x_i)) = \psi_i^n(x_i) + \log\Big(\int_{X_{-i}} K(x_i, x_{-i}) q_{-i}^n(x_{-i}) \mathrm{d}m_{-i}(x_{-i})\Big) \tag{4.2}$$

where

$$q_{-i}^n(x_{-i}) := e^{\sum_{j\neq i} \psi_j^n(x_j)}.$$

Since $(\mu_N^n)_n$ is uniformly bounded and bounded away from 0 and so is K, we deduce that $(e^{\psi_N^n})_n$ is bounded and bounded away from 0 in L^∞ i.e. $(\psi_N^n)_n$ is bounded in $L^\infty(X_N, \mathcal{F}_N, m_N)$. From this L^∞ bound on $(\psi_N^n)_n$, the fact that $K \in L^\infty_{++}(X, \mathcal{F}, m)$ and the uniform bounds from above and from below on μ_i^n, we deduce that ψ_i^n is bounded in L^∞ for $i = 1, \dots, N-1$. In particular, taking subsequences if necessary, we may assume that for every i, $(q_{-i}^n)_n$ converges weakly $*$ in $L^\infty(X_{-i}, \mathcal{F}_{-i}, m_{-i})$ to some q_{-i}, in particular $\int_{X_{-i}} K(x_i, x_{-i}) q_{-i}^n(x_{-i}) \mathrm{d}m_{-i}(x_{-i})$ converges for m_i-a.e. x_i to $\int_{X_{-i}} K(x_i, x_{-i}) q_{-i}(x_{-i}) \mathrm{d}m_{-i}(x_{-i})$. But since $\log(\mu_i^n)$ converges in $L^\infty(X_i, \mathcal{F}_i, m_i)$ to $\log(\mu_i)$, we deduce from (4.2) that ψ_i^n converges m_i-a.e. (and also in L^p for every $p \in [1, +\infty)$ by Lebesgue's dominated convergence Theorem) to some $\psi_i \in L^\infty$. Passing to the limit in (4.2), we then have $\mu = T(\psi)$ or equivalently $\mu = T(\varphi)$ for $\varphi \in E$ such that $\varphi - \psi$ is constant. This shows that $T(E)$ is closed in F_{++}. ∎

We are now in position to state our main result:

Theorem 4.3. *For every $\mu \in F_{++}$, the multi-marginal Schrödinger system (2.3) admits a unique solution $\varphi = S(\mu) \in E$, moreover $S \in C^\infty(F_{++}, E)$.*

Proof. It follows from Proposition 3.1 that $T(E)$ is open in F_{++} and Lemma 4.2 ensures it is closed in F_{++}, since F_{++} is connected (it is actually convex) we deduce that $T(E) = F_{++}$. Together with Proposition 4.1 this implies that T is a bijection between E and F_{++}, the smoothness claim then follows from Proposition 3.1. ∎

5 Further properties of the Schrödinger map

From now on, we refer to the smooth map $S = T^{-1} : F_{++} \to E$ from Theorem 4.3 as the Schrödinger map. Our aim now is to study the (local) Lipschitz behavior of S. Given $M \geq 1$ we define

$$F_{++,M} := \{\mu \in F_{++} \ : \ \frac{1}{M} \leq \mu_i \leq M \ m_i\text{-a.e.}\}. \tag{5.1}$$

Let us start with an elementary a priori bound:

Lemma 5.1. *For every $M \geq 1$ there is a constant R_M such that $S(F_{++,M})$ is included in the ball of radius R_M of $\prod_{i=1}^N L^\infty(X_i, \mathcal{F}_i, m_i)$.*

Proof. Let $\mu \in F_{++,M}$ and $\varphi = S(\mu)$, as in the proof of Lemma 4.2 we introduce constants λ_i with zero sum such that $\mu = T(\psi)$ with $\psi_i = \varphi_i + \lambda_i$ is normalized by (4.1) (instead of (2.4)). Using the fact that K is bounded and bounded away from 0, that $M^{-1} \leq \mu_N \leq M$, (4.1) and $\mu_N = T_N(\psi)$ gives upper and lower bounds on e^{ψ_N} i.e. an L^∞ bound (depending on M and K only) on ψ_N. This bound and $\mu_i = T_i(\psi)$ in turn provide L^∞ bounds on ψ_i for $i = 1, \dots, N-1$. Finally, we get bounds on the constants λ_i since $\lambda_i = \int_{X_i} \psi_i \mathrm{d}m_i$ for $i = 1, \dots, N-1$ and $\lambda_N = -\sum_{i=1}^{N-1} \lambda_i$. This gives the desired bounds on $\varphi = S(\mu)$.

□

More interesting in possible applications, is the Lipschitz behavior of S given by the folllowing

Theorem 5.2. *For every $M \geq 1$ there is a constant C_M, such that*[1] *for every μ and ν in $F_{++,M}$, there holds*

$$\|S(\mu) - S(\nu)\|_{L^2} \leq C_M \|\mu - \nu\|_{L^2}, \tag{5.2}$$

and

$$\|S(\mu) - S(\nu)\|_{L^\infty} \leq C_M \|\mu - \nu\|_{L^\infty}. \tag{5.3}$$

Proof. Let $\mu \in F_{++,M}$ and $\varphi = S(\mu) \in E$, our aim is to estimate the operator norm of $S'(\mu) = [T'(\varphi)]^{-1}$ (first in L^2 and then in L^∞). Let $\theta \in F$ and $h = S'(\mu)\theta$ i.e. $T'(\varphi)h = \theta$ which can be rewritten as

$$\widetilde{T}_i'(\varphi)h = \widetilde{\theta}_i \text{ with } \widetilde{\theta}_i := \frac{\theta_i}{\mu_i}. \tag{5.4}$$

[1] In formulas (5.2) (respectively (5.3)) L^2 (resp. L^∞) is a abbreviated notation for $\prod_{i=1}^N L^2(X_i, \mathcal{F}_i, m_i)$ (resp. $\prod_{i=1}^N L^\infty(X_i, \mathcal{F}_i, m_i)$).

Defining the measure Q_φ by (3.8) and disintegrating it with respect to its i-th marginal as in (3.9) in the proof of proposition 3.1 gives that for every i and m_i-a.e. x_i one has

$$\widetilde{\theta}_i(x_i) = h_i(x_i) + \int_{X_{-i}} \Big(\sum_{j\neq i} h_j(x_j)\Big) Q_\varphi^{-i}(\mathrm{d}x_{-i}|x_i). \tag{5.5}$$

We then argue in a similar way as we did in the proof of proposition 3.1, multiplying (5.5) by h_i and integrating with respect to Q_φ^i and summing over i, we obtain

$$\sum_{i=1}^N \int_{X_i} \widetilde{\theta}_i(x_i) h_i(x_i) \mathrm{d}Q_\varphi^i(x_i) = \int_X \Big(\sum_{j=1}^N h_j(x_j)\Big)^2 \mathrm{d}Q_\varphi(x). \tag{5.6}$$

Next we observe that thanks to the fact that $\mu \in F_{++,M}$, the upper and lower bounds on K and Lemma 5.1 there is a constant $\nu_M \geq 1$ such that

$$\frac{m}{\nu_M} \leq Q_\varphi \leq \nu_M m, \ \frac{m_i}{\nu_M} \leq Q_\varphi^i \leq \nu_M m_i. \tag{5.7}$$

Using the fact that $\|\widetilde{\theta}_i\|_{L^2(X_i,\mathcal{F}_i,m_i)} \leq M\|\theta_i\|_{L^2(X_i,\mathcal{F}_i,m_i)}$, (5.7) and Cauchy-Schwarz inequality, we deduce from (5.6) that there is a constant C_M such that

$$\int_X \Big(\sum_{j=1}^N h_j(x_j)\Big)^2 \mathrm{d}m(x) \leq C_M \sum_{i=1}^N \|\theta_i\|_{L^2(X_i,\mathcal{F}_i,m_i)} \|h_i\|_{L^2(X_i,\mathcal{F}_i,m_i)}. \tag{5.8}$$

Finally recall that since $h \in E$ we have

$$\int_X \Big(\sum_{j=1}^N h_j(x_j)\Big)^2 \mathrm{d}m(x) = \sum_{j=1}^N \int_{X_j} h_j^2(x_j) \mathrm{d}m_j(x_j) =: \|h\|_{L^2}^2$$

hence

$$\|h\|_{L^2} = \|S'(\mu)\theta\|_{L^2} \leq C_M \|\theta\|_{L^2} \text{ i.e. } \sup_{\mu\in F_{++,M}} \|S'(\mu)\|_{\mathcal{L}(L^2)} \leq C_M. \tag{5.9}$$

By the mean-value inequality (5.9) immediately gives the Lipschitz in L^2 estimate (5.2).

As for a bound on the operator norm of $S'(\mu)$ in L^∞ , we first observe that for some positive constant λ_M we have $Q_\varphi^{-i} \leq \lambda_M m_{-i}$, so that (5.5) gives

$$\begin{aligned}\|h_i\|_{L^\infty} &\leq \|\widetilde{\theta}_i\|_{L^\infty(m_i)} + \lambda_M \sum_{j\neq i} \int_{X_j} |h_j(x_j)| \mathrm{d}m_j(x_j) \\ &\leq M\|\theta_i\|_{L^\infty(m_i)} + \lambda_M \sqrt{N} \|h\|_{L^2} \\ &\leq M\|\theta_i\|_{L^\infty(m_i)} + \lambda_M \sqrt{N} C_M \|\theta\|_{L^2} \\ &\leq C'_M \|\theta\|_{L^\infty}\end{aligned}$$

where we have used Cauchy-Schwarz inequality in the second line and (5.9) in the third one. This clearly implies (5.3).

□

Acknowledgments: G.C. is grateful to the Agence Nationale de la Recherche for its support through the project MAGA (ANR-16-CE40-0014).

References

[1] Jean-David Benamou, Guillaume Carlier, Marco Cuturi, Luca Nenna, and Gabriel Peyré. Iterative Bregman projections for regularized transportation problems. *SIAM Journal on Scientific Computing*, 37(2):A1111–A1138, 2015.

[2] J. M. Borwein and A. S. Lewis. Decomposition of multivariate functions. *Canad. J. Math.*, 44(3):463–482, 1992.

[3] J. M. Borwein, A. S. Lewis, and R. D. Nussbaum. Entropy minimization, *DAD* problems, and doubly stochastic kernels. *J. Funct. Anal.*, 123(2):264–307, 1994.

[4] Haïm Brezis. *Analyse fonctionnelle.* Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1983. Théorie et applications. [Theory and applications].

[5] Guillaume Carlier, Vincent Duval, Gabriel Peyré, and Bernhard Schmitzer. Convergence of entropic schemes for optimal transport and gradient flows. *SIAM J. Math. Anal.*, 49(2):1385–1418, 2017.

[6] Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems*, pages 2292–2300, 2013.

[7] Steven G. Krantz and Harold R. Parks. *The implicit function theorem.* Birkhäuser Boston, Inc., Boston, MA, 2002. History, theory, and applications.

[8] Christian Léonard. From the Schrödinger problem to the Monge–Kantorovich problem. *Journal of Functional Analysis*, 262(4):1879–1920, 2012.

[9] Christian Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete Contin. Dyn. Systems, A*, 34(4):1533–1574, 2014.

[10] Filippo Santambrogio. *Optimal transport for applied mathematicians.* Progress in Nonlinear Differential Equations and their Applications, 87. Birkhäuser/Springer, Cham, 2015. Calculus of variations, PDEs, and modeling.

[11] Erwin Schrödinger. Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique. In *Annales de l'institut Henri Poincaré*, volume 2, pages 269–310, 1932.

[12] Cédric Villani. *Topics in Optimal Transportation*, volume 58 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, 2003.

[13] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.