

# OPTIMAL TRANSPORTATION THEORY WITH REPULSIVE COSTS

SIMONE DI MARINO, AUGUSTO GEROLIN, AND LUCA NENNA

## CONTENTS

Introduction	1
Organization of this Survey	3
What is new?	3
1. Why Multi-marginal Transport Theory for repulsive costs?	4
1.1. Brief introduction to Quantum Mechanics of $N$ -body systems	4
1.2. Probabilistic Interpretation and Marginals	7
1.3. Density Functional Theory (DFT)	8
1.4. “Semi-classical limit” and Optimal Transport problem	9
2. DFT meets Optimal Transportation Theory	12
2.1. Couplings and Multi-marginal Optimal Transportation problem	12
2.2. Multimarginal Optimal Transportation Problem	13
2.3. Dual formulation	15
2.4. Geometry of the Optimal Transport sets	17
2.5. Symmetric Case	18
3. Multi-marginal OT problem with Coulomb Cost	19
3.1. General theory: duality, equivalent formulations and many particles limit	20
3.2. The Monge problem: deterministic examples and counterexamples	22
4. Multi-marginal OT problem with repulsive Harmonic Cost	24
5. Multi-marginal OT for the Determinant	30
6. Numerics	33
6.1. The regularized problem and the Iterative Proportional Fitting Procedure	33
6.2. Numerical experiments: Coulomb cost	36
6.3. Numerical experiments: repulsive Harmonic cost	37
6.4. Numerical experiments: Determinant cost	38
Conclusion	40
Acknowledgements	40
References	40

ABSTRACT. RICAM survey repulsive cost in multi-marginal optimal transportation theory and predictions for the future!

## INTRODUCTION

The main objective of this survey is to present recent developments in optimal transport theory for repulsive costs with finitely many marginals. More precisely, we are interested in characterize, for a class of repulsive cost functions, the minimizers of the multi-marginal optimal

transportation problem

$$(\mathcal{MK}) \quad \min \left\{ \int_{\mathbb{R}^{dN}} c(x_1, \dots, x_N) d\gamma : \gamma \in \Pi(\mathbb{R}^{dN}, \mu_1, \dots, \mu_N) \right\},$$

where  $c : (\mathbb{R}^d)^N \rightarrow \mathbb{R} \cup \{\infty\}$  is a lower semi-continuous cost function and  $\Pi(\mathbb{R}^{dN}, \mu_1, \dots, \mu_N)$  denotes the set of couplings, i.e. probability measures  $\gamma(x_1, \dots, x_N)$  having  $\mu_i$  as marginals, for all  $i = 1, \dots, N$ . Typically,  $c$  can be given by a sum of convex decreasing functions, as the Coulomb interaction cost [16, 33, 65, 52, 65, 93]

$$c(x_1, \dots, x_N) = \sum_{i=1}^N \sum_{j=i+1}^N \frac{1}{|x_j - x_i|},$$

which is the case of most interest in DFT (see section 1). Also related to Quantum Mechanics, in the so-called weak interaction regime [16, 42, 85, 96], we will consider the repulsive harmonic potential as cost function  $c$ ,

$$c_w(x_1, \dots, x_N) = - \sum_{i=1}^N \sum_{j=i+1}^N |x_j - x_i|^2.$$

In the last section, we also discuss briefly the multimarginal problem for the determinant cost

$$c(x_1, \dots, x_N) = \det(x_1, \dots, x_N), \quad \text{or} \quad c(x_1, \dots, x_N) = |\det(x_1, \dots, x_N)|,$$

studied by Carlier and Nazaret in [21].

The literature on multi marginal problems in optimal transport is growing very fast both from theoretical and applied point of view. Optimal Transport Theory for finitely many marginals showed to be powerful in many applications as, for instance, in economics [19, 23, 24, 44, 78], mathematical finance [6, 8, 9, 38, 39, 46], image processing [99], tomography [1], statistics [20] and astrophysics [41].

Although very natural in such applied problems, up to our knowledge, optimal transport problems for 3 or more marginals was only considered in the literature during the 90's by Knott and Smith, Olkin and Rachev and Rüschendorf and Uckelmann [60, 76, 90], and later by Gangbo and Świech [47]. In [47], the authors proved the existence and uniqueness of Monge-type solution for the problem  $(\mathcal{MK})$  for the quadratic cost  $c(x_1, \dots, x_N) = \sum_{i,j=1}^N |x_j - x_i|^2$ .

A systematic study of the structure of transport plans and general conditions in order to guarantee the existence of Monge-type solutions for the problem  $(\mathcal{MK})$  was made by Pass [79, 80, 81] and Kim and Pass [56, 57]. Several theoretical developments were made in the last years, as the study of partial multi-marginal transport [59], symmetric Monge-Kantorovich problem [28, 29, 49, 50, 51, 62, 71, 72, 100] and duality [10, 36, 55, 100]. Applications in Geometry and Analysis as the Wasserstein Barycenter problem and decoupling a system of PDE's can be found in [2, 58, 51].

For a complete presentation of the multimarginal optimal transport theory with general costs functions, including a discussion of many applications, we refer the survey [77]. In this notes, we are rather analyzing the Multimarginal Optimal Transport Theory starting from the point of view of researchers working on Density Functional Theory.

Paola Gori-Giorgi, Michel Seidl *et al* were interested in compute accurately the ground state of a quantum molecule with a large number of atoms. As we will discuss later, after doing suitable approximations, the solutions of the optimal transport  $(\mathcal{MK})$  for Coulomb cost (more

precisely, the optimal potential), or more generally repulsive costs, can be seen as a good *ansatz* in order to construct solutions for Hohenberg-Kohn functional (1.12).

From a mathematical perspective, in those problems, the interesting case is when all marginals in the coupling  $\gamma$  in  $(\mathcal{MK})$  are absolutely continuous with respect to Lebesgue measure and are all the same. In particular, since the cost functions are “*repulsive*”, if Monge-type solutions exists, they should follow the rule “*the further, the better!*”, which means that we want to move the mass as much as we can. In other words, in the present case, optimal transport plans tend to be as spread as possible.

**Organization of this Survey.** In the section 1, we present the physical motivation in considering repulsive costs and give an introduction of Density Functional Theory through a mathematical viewpoint. The main goal of this section is explain how optimal couplings in Monge-Kantorovich problem can be useful to compute the ground state energy of a quantum system of  $N$  electrons.

The section 2 introduce briefly the general theory of multimarginal optimal transport, including a discussion about the Kantorovich duality and the notion of  $c$ -cyclical monotonicity applied to this framework. In addition, we give some highlights in the interesting case of symmetric transport plans.

In the sections 3, 4 and 5, respectively, we discuss the recent development for the specific cases of Coulomb, repulsive harmonic costs and the determinant.

Finally, in section 6 we introduce a recent iterative method to solve (multi marginal) optimal transportation problems and we present some numerical experiments for Coulomb, repulsive harmonic and the determinant costs.

**What is new?** We present in the section 4 some new results we obtained for cost functions

$$c(x_1, \dots, x_N) = f(|x_1 + \dots + x_N|), \quad f : \mathbb{R}^d \rightarrow \mathbb{R} \text{ is convex,}$$

which, in particular include the case of repulsive harmonic costs (see corollary 4.7). In addition, we show:

- (i) *Existence of Fractal-like cyclical Monge solutions:* we show the existence of couplings  $\gamma = (Id, T, T^2, \dots, T^{N-1})_{\#} \mu$  concentrated on the graphs of functions  $T, T^2, \dots, T^{N-1}$ , where  $\mu$  a uniform probability measure on the  $d$ -dimensional cube. Moreover,  $T$  is not differentiable in any point, which differs substantially to the classical regularity theory in the classical 2 marginals case [17, 18].
- (ii) *Optimal Transport plans supported on sets of Hausdorff dimension  $> 1$ :* In the classical optimal transport theory with 2 marginals the the question if Kantorovich type solutions of  $(\mathcal{MK})$  are also Monge-type solutions is fundamental and optimal transport plans are concentrated in the graph of a function  $T$ . In the multimarginal setting, it was known that, for the repulsive harmonic cost, there exist optimal transport plans which are not concentrated in one dimensional sets [42]. In Section 4, we give a concrete general example where optimal transport plans are supported on 2-dimensional Hausdorff sets.
- (iii) *Numerical results:* we present some interesting results for the repulsive harmonic and for the determinant cost.

## 1. WHY MULTI-MARGINAL TRANSPORT THEORY FOR REPULSIVE COSTS?

**1.1. Brief introduction to Quantum Mechanics of  $N$ -body systems.** Energies and geometries of a molecule depend sensibly on the kind of atom on a chemical environment. Semi-empirical models as *Lewis structure* explain some aspects, but they are far to be satisfactory, mostly because they are not quantitative.

Classical behavior of atoms and molecules is described accurately, at least in a theoretical point of view, by quantum mechanics. However, in order to predict the chemical behavior of a molecule with a large number of electrons we need to deal with computational aspects and approximations are needed. For example, if we use a direct approximation of 10 grid points in  $\mathbb{R}$  of the time-dependent Schrödinger equation in order to simulate the chemical behavior of the water molecule ( $H_2O$ ), which has 10 electrons and its position is represented in the space  $\mathbb{R}^{30}$ , then we need to compute  $10^{30}$  grid points.

Quantum Chemistry studies the ground state of individual atoms and molecules, the excited states and the transition states that occur during chemical reactions. Many systems use the so-called *Born-Oppenheimer approximation* and many computations involve iterative and other approximation methods.

The main goal of Quantum Chemistry is to describe accurately results for small molecular systems and to increase the size of large molecules that can be processed, which is very limited by scaling considerations.

It means that from a Quantum Model, for which the only input is the atomic number of a molecule, we want to predict the evolution of a molecular system. Classical models consider a molecule with  $N$  electrons and  $M$  nuclei. We denote by a point  $x_i \in \mathbb{R}^3$  the position coordinates of the  $i$ -th electron of mass  $m_e$  and by  $s_i$  its  $i$ -th electron spin. The charges, the mass and the positions of the  $\alpha$ -th nucleus are represented, respectively, by  $Z_\alpha \in \mathbb{R}^M$ ,  $m_\alpha \in \mathbb{R}_+^M$  and  $R_\alpha \in (\mathbb{R}^3)^M$ .

The state of the system is described by a time-dependent wave function  $\psi \in L^2([0, T]; \otimes_{i=1}^N L^2(\mathbb{R}^3 \times \mathbb{Z}^2) \times \otimes_{\alpha=1}^M L^2(\mathbb{R}^3))$ ,  $\psi(x_1, s_1, \dots, x_N, s_N; R_1, \dots, R_M, t)$ . The group  $\mathbb{Z}^2 = \{\uparrow, \downarrow\}$  represents the spin of a particle.

We say that a wave functions is antisymmetric or *fermionic* if their change sign under a simultaneous exchange of two electrons  $i$  and  $j$  with the space coordinates  $x_i$  and  $x_j$  and spins  $s_i$  and  $s_j$ , that is,

$$\psi(x_{\sigma(1)}, s_{\sigma(1)}, \dots, x_{\sigma(N)}, s_{\sigma(N)}) = \text{sign}(\sigma) \psi(x_1, s_1, \dots, x_N, s_N) \quad \sigma \in \mathcal{S}_N$$

where  $\mathcal{S}_N$  denotes the permutation set of  $N$  elements.

In the physics literature, *Fermionic* wave function obeys the *Fermi-Dirac statistics* and, in particular, the *Pauli exclusion principle*. On the other hand, a wave function is called symmetric or *bosonic* when obeys *Bose-Einstein statistics* which also implies that when one swaps two bosons, the wave function of the system is unchanged.

If it is not mentioned explicitly, we will suppose that the wave function is spinless, i.e.  $\psi$  is a function depending only on time, electrons and nuclei position  $\psi(t, x_1, x_2, \dots, x_N, R_1, \dots, R_M)$ .

The evolution of a wave function  $\psi$  boils down to a (time-dependent) Schrödinger Equation

$$i\partial_t \psi = H\psi \quad , \quad H = T_n + T_e + V_{ne} + V_{ee} + V_{nn} \tag{1.1} \quad \boxed{\text{GSE}}$$

where the operators  $T_n, T_e, V_{ne}, V_{ee}$  and  $V_{nn}$  are defined by

$$\begin{aligned}
T_n &= -\hbar \sum_{\alpha=1}^M \frac{1}{2m_\alpha} \Delta_{R_\alpha} \quad (\text{Nuclei Kinetic Energy}) \\
T_e &= -\hbar \sum_{i=1}^N \frac{1}{2m_e} \Delta_{x_i} \quad (\text{Electron Kinetic Energy}) \\
V_{ne} &= -\sum_{i=1}^N v_R(x_i) = -\sum_{i=1}^N \left( \sum_{\alpha=1}^M \frac{Z_\alpha}{|x_i - R_\alpha|} \right) \quad \left( \begin{array}{l} \text{Potential Energy of Inte-} \\ \text{raction Nuclei-Electron} \end{array} \right) \\
V_{ee} &= \sum_{i=1}^N \sum_{j=1, j \neq i}^N \frac{1}{|x_i - x_j|} \quad (\text{Interaction Electron-Electron}) \\
V_{nn} &= \sum_{\alpha=1}^M \sum_{\beta=1, \beta \neq \alpha}^M \frac{Z_\alpha Z_\beta}{|R_\alpha - R_\beta|} \quad (\text{Interaction Nuclei-Nuclei})
\end{aligned}$$

**Born-Oppenheimer Approximation:** The Born-Oppenheimer approximation (named for its original inventors, Max Born and Robert Oppenheimer) is based on the fact that nuclei are several thousand times heavier than electrons. The proton, itself, is approximately 2000 times more massive than an electron. Roughly speaking, in Born-Oppenheimer approximation we suppose that nuclei behave as point classical particles. This reasonable, since typically  $2000m_e \leq m_\alpha \leq 100000m_e$ .

In other words, we suppose  $m_\alpha \gg m_e$ , we fix  $m_e = 1$  and we consider an *ansatz* of type  $\psi(x, R, t) = \psi_e(x, R)\chi(R, t)$ , where physically means that the dynamical of the electron  $\psi_e$  are decoupled of the dynamics of the nuclei  $\chi$ . Substituting that *ansatz* in (1.1), we can show that the electronic part  $\psi_e(x, R)$  solves the following eigenvalue problem, so-called *Electronic Schrödinger Equation*

$$H_e \psi_e = \lambda_e \psi_e, \quad H_e = T_e + V_{ne} + V_{ee} + V_{nn} \quad (1.2) \quad \boxed{\text{ESE}}$$

where the eigenvalue  $\lambda_e = \lambda_e(R)$  depends on the position vector  $R = (R_1, \dots, R_M)$  of atomic nuclei. The function  $H_e = H_e(R)$  is called *Electronic Hamiltonian*. From the other side, through a formal argument, the nuclei wave function  $\chi(R, t)$  is a solution of a Schrödinger Equation restrict to the nuclei with potential energy  $\lambda_e$ ,

$$i\partial_t \chi(R) = (T_n + \lambda_e(R))\chi(R).$$

We refer to [14] for all computations and formal deduction of those formulas. Another consequence of that approximation is that the nuclear components of the wave function are spatially more localized than the electronic component of the wave function. In the classical limit, the nuclear are fully localized about single points representing classical point particles.

In Quantum Chemistry is interesting to study the of minimizing the nuclei configuration, called *Geometric Optimization Problem*.

*The Geometric Optimization Problem:* Compute the following minimizer

$$\inf \left\{ E(R_1, \dots, R_M) + \sum_{\alpha=1}^M \sum_{\beta=1, \beta \neq \alpha}^M \frac{Z_\alpha Z_\beta}{|R_\alpha - R_\beta|} : (R_1, \dots, R_M) \in \mathbb{R}^{3M} \right\} \quad (1.3) \quad \boxed{\text{pb:GOP}}$$

where

- The second term is the nuclei-nuclei iteration  $V_{nn}$ , as defined in the (1.1).

- The function  $E(R) = E(R_1, \dots, R_M)$  corresponds to the effective potential created by the electrons and is, itself, given by a minimization problem (see *Electronic minimization problem* below).

The value of the potential  $E_0 = E_0(R_1, \dots, R_M)$  for given fixed positions of the nuclei is obtained by solving the electronic problem:

**The Electronic minimization problem:** Compute the lowest eigenvalue ( “ground state energy”)  $E_0$  of the following linear operator, called *Electronic Hamiltonian*

**Definition 1.1** (Electronic Hamiltonian). *The Electronic Hamiltonian is a linear operator*

$$H_{el} : L_{anti}^2(\mathbb{R}^{3N}) \rightarrow L_{anti}^2(\mathbb{R}^{3N}), \quad H_{el} = -\frac{\hbar^2}{2} \sum_{i=1}^N \Delta_{x_i} + \sum_{i=1}^N v_R(x_i) + \sum_{i=1}^N \sum_{j=i+1}^N f(x_j - x_i)$$

where  $L_{anti}^2(\mathbb{R}^{3N})$  denote the set of square-integrable antisymmetric functions  $\psi : \mathbb{R}^{3N} \rightarrow \mathbb{R}$ ,  $v_R : \mathbb{R}^3 \rightarrow \mathbb{R}$  is a  $L^2(\mathbb{R}^{3N})$  function and  $f : \mathbb{R} \rightarrow \mathbb{R}$  is a continuous function.

The first term of the Electronic Hamiltonian is the Kinetic Energy and the second term, the function  $v_R(x_i)$ , depends only on the single electron. Typically,  $v_R(x_i) = -\sum_{\alpha=1}^M \frac{Z_\alpha}{|x_i - R_\alpha|}$  is the interaction electron-nuclei energy of the single electron. The function  $f$  which depends only on the distance of two electrons and measures the electron-electron potential interaction; typically  $f$  is the Coulomb Interaction  $f(x_j - x_i) = 1/|x_j - x_i|$  or the repulsive harmonic interaction  $f(x_j - x_i) = -|x_j - x_i|^2$ . From both mathematical and physical viewpoint, may be also useful to study more general convex and concave functions.

The *ground state* of the functional (1.1) is given by the *Rayleigh-Ritz variational principle*

$$E_0 = E(R_1, \dots, R_M) = \min\{\langle \psi, H_{el} \psi \rangle : \psi \in \mathcal{H}\}$$

where  $\mathcal{H} = \{\psi \in H^1(\mathbb{R}^{3N}) : \psi \text{ is antisymmetric and } \|\psi\| = 1\}$ . Equivalently,

$$E_0 = \min \left\{ T_e[\psi] + V_{ne}[\psi] + V_{ee}[\psi] : \psi \in \mathcal{H} \right\} \quad (1.4) \quad \boxed{\text{eq. groundstateenergy}}$$

where  $T[\psi]$  is the *Kinetic energy*,

$$T_e[\psi] = \frac{\hbar^2}{2} \sum_{s_1 \in \mathbb{Z}^2} \int_{\mathbb{R}^3} \dots \sum_{s_N \in \mathbb{Z}^2} \int_{\mathbb{R}^3} \sum_{i=1}^N |\nabla_{x_i} \psi(x_1, s_1, \dots, x_N, s_N)|^2 dx_1 \dots dx_N$$

$V_{ne}[\psi]$  is the electron-nuclei interaction energy,

$$V_{ne}[\psi] = \sum_{s_1 \in \mathbb{Z}^2} \int_{\mathbb{R}^3} \dots \sum_{s_N \in \mathbb{Z}^2} \int_{\mathbb{R}^3} \sum_{i=1}^N v_R(x_i) |\psi(x_1, s_1, \dots, x_N, s_N)|^2 dx_1 \dots dx_N$$

and,  $V_{ee}[\psi]$  is the electron-electron interaction energy

$$V_{ee}[\psi] = \sum_{s_1 \in \mathbb{Z}^2} \int_{\mathbb{R}^3} \dots \sum_{s_N \in \mathbb{Z}^2} \int_{\mathbb{R}^3} \sum_{i=1}^N \sum_{j=i+1}^N \frac{1}{|x_j - x_i|} |\psi(x_1, s_1, \dots, x_N, s_N)|^2 dx_1 \dots dx_N$$

We refer here as the ground state quantum  $N$ -body problem the problem to find equilibrium states for such a kind of system in (1.1). A theorem proved first by Zhislin [101], guarantee the existence of the minimum in (1.4). Some variants can be found in the literature due, for instance, to Lieb & Simon [68], Lions [70] and Friesecke [43].

In computational chemistry, the Electronic minimization problem methods are used to compute the equilibrium configuration of molecules and solids. However, the most computationally practicable methods in this context are not numerical methods (in the sense in which this terminology is used in mathematics), but power series of analytical solutions of reduced models.

**1.2. Probabilistic Interpretation and Marginals.** The absolute value square of the wave function  $\psi(x_1, s_1, \dots, x_N, s_N)$  can be interpreted as an  $N$ -point probability density distribution that electrons are in the points  $x_i$  with spins  $s_i$ .

$$\int_{\mathbb{R}^{3N}} |\psi(x_1, s_1, \dots, x_N, s_N)|^2 dx = 1.$$

If  $\psi$  is a  $L^2(\mathbb{R}^{3N})$  function we can define the single particle density by

$$\rho^\psi(x_1) = N \int_{\mathbb{R}^{3(N-1)}} \rho_N^\psi(x_1, s_1, \dots, x_N, s_N) dx_2, \dots, dx_N, \quad (1.5)$$

eq.singlede

where  $\rho_N^\psi$  represents the  $N$ -point position density by

$$\rho_N^\psi(x_1, \dots, x_N) = \sum_{s_1, \dots, s_N \in \mathbb{Z}_2} |\psi(x_1, s_1, \dots, x_N, s_N)|^2. \quad (1.6)$$

eq.Npointde

Analogously, we can define the  $k$ -density

$$\rho_k^\psi(x_1, \dots, x_k) = \binom{N}{k} \int_{\mathbb{R}^{3(N-k)}} \rho_N^\psi(x_1, s_1, \dots, x_N, s_N) dx_{k+1} \dots dx_N \quad (1.7)$$

eq.pairdens

We remark that when the particle is spinless the single particle density (1.5) can be simply be written as

$$\rho^\psi(x_1) = N \int_{\mathbb{R}^{3(N-1)}} |\psi(x_1, \dots, x_N)|^2 dx_2, \dots, dx_N.$$

The relevance of  $\rho^\psi$  and  $\rho_2^\psi$  to compute the ground state energy (1.4) is due to the fact that the integration nuclei-nuclei energy  $V_{ne}[\psi]$  and electron-electron energy  $V_{ee}[\psi]$  depends only on these, which is summated in the following lemma

lemma:rhoesem

**Lemma 1.2** ([33]). *If  $\psi, \nabla\psi \in L^2((\mathbb{R}^3 \times \mathbb{Z}^2)^N, \mathbb{R})$ ,  $\psi$  antisymmetric and  $\|\psi\|_2 = 1$ , then*

$$V_{ne}[\psi] = \int_{\mathbb{R}^3} v(x) \rho^\psi(x) dx, \quad V_{ee}[\psi] = \int_{\mathbb{R}^6} \frac{1}{|x-y|} \rho_2^\psi(x, y) dx dy.$$

In the next subsection we will denote the densities  $\rho^\psi$  and  $\rho_2^\psi$  simply by  $\rho$  and  $\rho_2$ , omitting the superscript symbol.

It is natural to wonder about the space of those densities arising from a function  $\psi$  is defined, i.e, the space  $\mathcal{A}$  of densities  $\rho : \mathbb{R}^3 \rightarrow \mathbb{R}$  verifying (1.5) for an antisymmetric  $\psi$  such that  $\psi, \nabla\psi \in L^2((\mathbb{R}^3 \times \mathbb{Z}^2)^N, \mathbb{R})$  and  $\|\psi\|_2 = 1$ . An explicit characterization exists:

$$\mathcal{A} = \left\{ \rho : \mathbb{R}^3 \rightarrow \mathbb{R} : \rho \geq 0, \sqrt{\rho} \in H^1(\mathbb{R}^3) \text{ and } \int_{\mathbb{R}^3} \rho(x) dx = N \right\}. \quad (1.8)$$

eq.spacerho

**1.3. Density Functional Theory (DFT).** The Density Functional Theory is the standard approximation to quantum mechanics in simulations of a system with more than a dozen electrons and showed to be successful in many instances but has rare drastic failures as, for instance, in predicting the behavior of  $Cr_2$  molecules [27]. DFT theory approximates quantum mechanics via variational principles for the marginal density

$$\rho_1(x_1) := \int_{\mathbb{R}^{d(N-1)}} |\psi(x_1, \dots, x_N)|^2 dx_2 \dots dx_N, \quad (1.9) \quad \text{eq:rho1}$$

where  $\psi$  is a wave function associated to the  $N$ -body quantum problem as in equation (1.5). Roughly speaking, DFT models are semi-empirical models of the pair density

$$\rho_2(x_1, x_2) := \int_{\mathbb{R}^{3(N-2)}} |\psi(x_1, \dots, x_N)|^2 dx_3 \dots dx_N. \quad (1.10) \quad \text{eq:rho2}$$

in terms of its marginal  $\rho$ . We simply write  $\psi \rightarrow \rho_1$  and  $\psi \rightarrow \rho_2$  to denote the relation between  $\psi$  and  $\rho_1$ , and  $\psi$  and  $\rho_2$ . This means, respectively, that  $\psi$  has single particle density  $\rho$  and pair density  $\rho_2$ .

Concerning the ground state problem (1.4), we notice that, after the lemma 1.2, we are able to write the *Electronic minimization problem* (1.4) as

$$E_0 = \inf_{\rho \in \mathcal{A}} \left\{ F_{HK}[\rho] + \int_{\mathbb{R}^3} v(x) \rho(x) dx \right\}, \quad (1.11) \quad \text{eq:groundstatedensity}$$

with

$$F^{HK}[\rho] = \inf \left\{ T[\psi] + V_{ee}[\rho_2] : \rho \in \mathcal{A}, \psi \rightarrow \rho \right\}. \quad (1.12) \quad \text{eq:FHK}$$

where  $F_{HK}$  is the so called *Hohenberg-Kohn* or *Levy-Lieb functional*. For a mathematical seek of completeness, in this notes, instead of the Coulombian interaction  $V_{ee}[\rho]$  we are also consider a weak repulsive interaction given by

$$\hat{V}_{ee}[\rho] = \int_{\mathbb{R}^{3N}} \sum_{i=1}^N \sum_{j=1}^N -|x_j - x_i|^2 \rho_N^\psi(x_1, \dots, x_N) dx_1 \dots dx_N \quad \text{in a compact set } K \subset \mathbb{R}^{3N} \quad (1.13) \quad \text{eq:confinedrepulsiveha}$$

and  $+\infty$  otherwise.

The next theorem is the heart of Density Functional theory which was created by Walter Kohn.

:HK

**Theorem 1.3** (Hohenberg-Kohn, Levy-Lieb). *Let  $\rho \in \mathcal{P}(\mathbb{R}^3)$  be a probability density such that  $\rho \in H^1(\mathbb{R}^3)$  and  $f : \mathbb{R} \rightarrow \mathbb{R}$  be a continuous function. There exists a functional  $F^{HK} : \mathcal{P}(\mathbb{R}^3) \rightarrow \mathbb{R}$  depending only on the single-particle density  $\rho$  such that for any potential  $V_{nn}$ , the exact Quantum Mechanics ground state energy (1.4) satisfies*

$$E_0 = \min \left\{ \left( F^{HK}[\rho] + N \int_{\mathbb{R}^3} v_R(x) \rho(x) dx \right) : \rho \in \mathcal{P}(\mathbb{R}^3) \right\} \quad (1.14) \quad \text{eq:HK}$$

Moreover,  $F^{HK}[\rho]$  is given itself as a minimum problem

$$F^{HK}[\rho] = \min \left\{ \left\langle \psi, \left( -\frac{\hbar^2}{2} \Delta + \sum_{i=1}^N \sum_{j=i+1}^N f(x_j - x_i) \right) \psi \right\rangle : \psi \in H^1(\mathbb{R}^{3N}), \psi \rightarrow \rho \right\}$$

where,  $\psi \rightarrow \rho$  means that  $\psi$  has single-particle density  $\rho$  and  $\mathcal{A}$  is the set of  $\rho : \mathbb{R}^3 \rightarrow \mathbb{R}$  such that  $\rho \geq 0$ ,  $\sqrt{\rho} \in H^1(\mathbb{R}^3)$  and  $\int_{\mathbb{R}^3} \rho(x) dx = N$ .



The Hohenberg-Kohn theorem states that the functional  $F^{HK}$  is said to be *universal* in the sense that it does not depend on the molecular system under consideration. From a Physics point of view, it guarantees that in a molecular system of  $N$  electrons, the first electron position density  $\rho_1$  determines the pair density of the system (see corollary (1.4) below). From a mathematical perspective, the proof of this theorem is a functional analysis exercise that, for the sake of completeness, we are going to present a version we learnt from Gero Friesecke.

*Proof.* First remark that the non-universal part of the energy only depends on  $\rho^\psi$ :

$$\langle \psi, \sum_{i=1}^N v(x_i) \psi \rangle = \int_{\mathbb{R}^{3N}} \sum_{i=1}^N v(x_i) |\psi(x_1, \dots, x_N)|^2 d\mathbf{x} = N \int_{\mathbb{R}^3} v(x) \rho^\psi(x) dx$$

Now, we do a standard argument of *Electronic minimization problem* in (1.4), transforming it in a double inf problem

$$\begin{aligned} E_0 &= \inf_{\psi} \left( \langle \psi, H_{el} \psi \rangle + N \int v(r) \rho^\psi(r) dr \right) \\ &= \inf_{\psi} \left( \inf_{\psi \mapsto \rho} \left( \langle \psi, H_{el} \psi \rangle \right) + N \int v(r) \rho^\psi(r) dr \right) \\ &= \inf_{\psi} \left( F^{HK}[\rho] + N \int v(r) \rho^\psi(r) dr \right) \end{aligned}$$

where in the last line we denoted  $F^{HK}[\rho] = \inf_{\psi \mapsto \rho} \left( \langle \psi, H_{el} \psi \rangle \right)$ .  $\square$

The present version stated on theorem 1.3 is due by Levy and Lieb. The next corollary contains the main physical and mathematical consequence of the Hohenberg-Kohn theorem (1.3).

**cor:HK**

**Corollary 1.4** (HK Theorem and Coupling problem). *Let  $\rho \in \mathcal{P}(\mathbb{R}^3)$  be a measure and  $\psi \in H^1(\mathbb{R}^{3N})$  be an antisymmetric function with  $\|\psi\|_2 = 1$ . There exists a universal map from single-particle density  $\rho(x_1)$  to a pair densities  $\rho_2(x_1, x_2)$  - or even  $k$ -density  $\rho_k(x_1, \dots, x_k)$  - which gives the exact pair density of any  $N$ -electron molecular ground state  $\psi(x_1, \dots, x_N)$  in terms of its single-particle density.*

*Proof.* Consider

$$\bar{\psi} \in \operatorname{argmin} \{ \langle \psi, H_{el} \psi \rangle : \psi \rightarrow \rho \}$$

and define by  $\rho_k$  the universal  $k$ -point density of that minimizer, i.e.

$$\rho_k(x_1, \dots, x_k) = \int |\bar{\psi}(x_1, \dots, x_N)|^2 dx_{k+1} \cdots dx_N$$

$\square$

Of course,  $\rho_k$  may be not unique, since the *ground state* may be degenerate. The existence of a map  $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  which determines the density of  $\rho_2$  from  $\rho$  is highly not trivial and not feasible because still uses high dimension wave function.

**1.4. “Semi-classical limit” and Optimal Transport problem.** A natural approach to understand the behavior of the Hohenberg-Kohn functional  $F^{HK}[\rho]$  (1.12) is to study separately, the contributions of the kinetic energy and the Coulomb interaction electron-electron energy.

A well-know method in DFT literature, known as adiabatic connection, is to neglect the kinetic energy, in which the electron-electron interaction is rescaled with a real parameter  $\lambda$  while keeping the density  $\rho$  fixed [63, 98]

$$F_\lambda^{HK}[\rho] = \min \left\{ \langle \psi, \left( -\frac{\hbar^2}{2} \Delta + \sum_{i=1}^N \sum_{j=i+1}^N \frac{\lambda}{|x_j - x_i|} \right) \psi \rangle : \psi \in H^1((\mathbb{R}^3)^N) : \psi \rightarrow \rho \right\}. \quad (1.15) \quad \text{eq:FHKlambda}$$

The strong-interaction limit ( $\lambda \rightarrow \infty$ ), up to passage on probability measures, was first considered by two papers in physics literature: Seidl [93] and Seidl, Gori-Giorgi & Savin [94]. In [52], Gori-Giorgi, Seidl and Vignale interpreted the strong-interaction regime as a mass transportation problem with a Coulomb cost.

Later, an equivalent limit, so-called “*semi-classical limit*”, was made mathematical rigorously in the two particles case by Cotar, Friesecke & Klupperlberg [33]. They considered the Hosenberg-Kohn functional  $F^{HK} = F_h^{HK}$  as a function of both  $\rho$  and  $\hbar$  ( $\lambda = 1$ ) and proved that - up to passage of the limit  $\hbar \rightarrow 0$  - the  $F_{HK}$  reduces to the following functional obtained by a minimization over pair densities (1.7) instead of wave functions

$$\tilde{F}[\rho] = \inf \left\{ \int_{\mathbb{R}^6} \frac{1}{|x - y|} \rho_2(x, y) dx dy : \rho_2 \in \mathcal{A}_2, \rho_2 \rightarrow \rho \right\}, \quad (1.16) \quad \text{eq:HKpairdensities}$$

where  $\rho_2 \rightarrow \rho$  means that  $\rho_2$  satisfies the equation (1.7), and the set  $\mathcal{A}_2$  of admissible pair densities is defined by the image of  $\mathcal{A}$  under the map  $\psi \rightarrow \rho$ . Instead of the corresponding single particle density case, we do not know any characterization of the space of admissible pair density function  $\mathcal{A}_2$ .

We state in the next theorem the semi-classical limit for the 2-particles case. The general case of  $N$  particles is still an open problem.

CFK

**Theorem 1.5** (“Semi-classical limit” for  $N = 2$ , Cotar-Friesecke-Klupperlberg, [33]). *Let  $\rho : \mathbb{R}^3 \rightarrow [0, \infty]$  be a probability density such that  $\sqrt{\rho} \in H^1(\mathbb{R}^3)$ ,  $N = 2$  and  $f : \mathbb{R} \rightarrow \mathbb{R}$  be the Coulomb or the repulsive harmonic interaction potentials. Then,*

$$\begin{aligned} F_h^{HK}[\rho] &= \min \left\{ \langle \psi, \left( -\frac{\hbar^2}{2} \Delta + f(x_2 - x_1) \right) \psi \rangle : \psi \in H^1(\mathbb{R}^6), \psi \rightarrow \rho \right\} \\ &\xrightarrow{\hbar \rightarrow 0} \min \left\{ \int_{\mathbb{R}^{3N}} f(x_2 - x_1) d\gamma(x_1, x_2) : \gamma \in \Gamma(\mathbb{R}^6, \rho) \right\} =: F^{OT}[\rho] \end{aligned} \quad (1.17) \quad \text{eq.semiclassicallimit}$$

where the set  $\Gamma(\mathbb{R}^6, \rho)$  denotes the set of measures  $\gamma \in \mathcal{P}(\mathbb{R}^6)$  having  $\rho$  as marginals, i.e.  $(e_i)_\# \gamma = \rho$  where, for  $i = 1, \dots, N$   $e_i$  is the projection maps  $e_i : \mathbb{R}^{3N} \rightarrow \mathbb{R}^3$ . The symbol  $\gamma \rightarrow \rho$  means that  $\gamma$  has pair density  $\rho$ .

The main difficulty in proving Theorem 1.5 is that for any transport plan given by a density  $\psi$ ,  $\gamma = |\psi(x_1, \dots, x_N)|^2 dx_1 \dots dx_N$ , we have that  $\psi \notin H^1(\mathbb{R}^{3N})$ ,  $\psi \notin L^2(\mathbb{R}^{3N})$  and  $T[\psi] = \infty$ . Moreover, smoothing the an optimal  $\gamma$  does not work at this level, because this may change the marginals of the problem (1.17). Cotar, Friesecke & Klupperlberg developed a smothing technique in order to deal with this problem without changing the marginals. A complete proof of the previous theorem can be found in [33].

We want to stress a little with the terminology. In Quantum Mechanics, the “*semi-classical limit*” has a precise meaning: it is an asymptotic regime for the Hamiltonian dynamics of a Quantum system defined in a Hilbert space and it is given by a *Weyl-Wigner quantization* (or *quantization by deformation* studied in a more abstract context of Poisson manifolds). In those specific cases, the limit  $\hbar \rightarrow 0$  is called “*semi-classical*” limit, because the first and second order

terms of that asymptotic expansion is given by “classical” terms, functions of the Hamiltonian function in a symplectic manifold [40, 88].

In DFT context, that limit seems to be up to know merely a question of re-scaling. The minimizers of  $F^{OT}[\rho]$  are candidates to *ansatz* to develop approximated methods to compute the ground state energy of the *Electronic Hamiltonian*.

At that point, is natural to define the *DFT-Optimal Transportation ground state*  $E_0^{DFT-OT}$ ,

$$\begin{aligned} E_0^{DFT-OT}[\rho] &= \inf \left\{ T[\psi] + V_{ne}[\rho^\psi] + E_{OT}[\rho^\psi] : \psi \in \mathcal{H} \right\} \\ &= \inf \left\{ T_{QM}[\rho] + V_{ne}[\rho] + E_{OT}[\rho] : \rho \in \mathcal{A} \right\} \end{aligned} \quad (1.18) \quad \text{eq:DFTOT}$$

where  $\rho^\psi$  represents the single particle density (see (1.5)),  $T_{QM}$  and  $E^{OT}$  are defined, respectively, by

$$\begin{aligned} T_{QM}[\psi] &= \inf \left\{ T[\psi] : \psi \in \mathcal{H}, \psi \rightarrow \rho \right\} \quad \text{and} \\ E^{OT}[\rho] &= \inf \left\{ \int_{\mathbb{R}^{3N}} \sum_{i=1}^N \sum_{j=1}^N f(x_j - x_i) d\gamma(x_1, \dots, x_N) : \gamma \in \Pi((\mathbb{R}^d)^N, \rho) \right\}. \end{aligned} \quad (1.19) \quad \text{eq:OTground}$$

where  $\Pi((\mathbb{R}^d)^N, \rho)$  denotes the set of probability measures  $\gamma : (\mathbb{R}^d)^N \rightarrow \mathbb{R}_+$  having  $\rho$  has marginals. We recall the ground state problem (1.4)

$$E_0 = \inf \left\{ T[\psi] + V_{ne}[\psi] + V_{ee}[\psi] : \psi \in \mathcal{H} \right\} = \inf \left\{ T[\psi] + V_{ne}[\rho^\psi] + V_{ee}[\rho_2^\psi] : \psi \in \mathcal{H} \right\}.$$

The problem to minimize  $E_0^{DFT-OT}$  in (1.18) is called DFT-OT problem. Notice that, in the case where  $f$  is the Coulomb potential,

$$V_{ee}[\rho_2] \geq E_{OT}[\rho], \quad \text{for every probability density } \rho \in H^1(\mathbb{R}^3).$$

Finally, taking the infimum in both sides of the previous equation, we have

**Theorem 1.6** (Cotar-Friesecke-Klüpperlberg, [33]). *Consider  $f(|x_j - x_i|) = 1/|x_j - x_i|$ . For every  $N$ , and any potential  $v_R \in L^{3/2} + L^\infty(\mathbb{R}^3)$ , the density functional with electron-electron interaction energy (1.19) is a rigorous lower bound of the Electronic minimization problem (1.4)*

$$E_0 \geq E_0^{DFT-OT}$$

It turn out now that the central question of DFT-OT problem with Coulomb potentials is to characterize the minimizers  $\gamma$  of  $E^{OT}[\rho]$ . Those minimizers could be used an *ansatz* in order to find minimizers of the Hohenberg-Kohn  $F^{HK}[\rho]$ . In [93], the physicist Seidl formulate the following conjecture.

conj:Seidl

**Conjecture 1.7** (Seidl [93]). *There exists a deterministic minimizer of  $E^{OT}[\rho]$  with Coulomb potential/cost*

$$c(x_1, \dots, x_N) = \sum_{i,j=1}^N \frac{1}{|x_i - x_j|}.$$

In other words, for Coulomb-type electron-electron interactions, Seidl conjectures that exist a measure  $\gamma$  that minimizes  $E^{OT}$  of form  $\gamma = (Id \times T_1 \times \dots \times T_{N-1})_\# \rho_1$ , given by maps  $T_1, \dots, T_{N-1} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ . It is natural to ask the same kind of question for more general cost functions. In the past year, similar question was studied in the optimal transportation literature giving different answers for the problem. We are going to mention some of them:

- (i) In the 2-electrons case, the minimizers of  $E_0^{DFT-OT}$  are well-understood [33] and there exists a map which correlates the density of a given minimization plan of  $E^{OT}[\rho]$ , see Theorem 3.6.
- (ii) Gangbo & Swiech (without being aware of Seidl) reinforces the conjecture 1.7 showing in [47, 21] that “*Seidl’s conjecture’s*” is true, respectively, for the attractive harmonic cost  $c(x_1, \dots, x_N) = \sum_{i,j=1}^N |x_i - x_j|^2$ .
- (iii) Colombo, De Pascale & Di Marino [29] answer affirmatively Seidl conjecture in the one dimensional case ( $d = 1$ ), as we will describe in the section 3.
- (v) The authors proved that Seidl’s conjecture is not true for the repulsive harmonic cost  $c(x_1, \dots, x_N) = \sum_{i,j=1}^N -|x_i - x_j|^2$ . Moreover, for the same cost and for a particular single particle density, there exists a “*fractal-like*” optimal transport maps, which are not differentiable almost everywhere. In the see section 4, we are going to state and proof those statements.

## 2. DFT MEETS OPTIMAL TRANSPORTATION THEORY

### 2.1. Couplings and Multi-marginal Optimal Transportation problem.

**Definition 2.1** (Coupling). *Let  $(X, \mu), (Y, \nu)$  be two probability spaces. A coupling  $\gamma$  is a measure on the product  $X \times Y$  such that  $\gamma$  admits  $\mu$  and  $\nu$  as marginals, respectively, on  $X$  and  $Y$ , i.e.*

$$(e_1)_\# \gamma = \mu, \quad (e_2)_\# \gamma = \nu$$

where  $e_1, e_2$  are respectively the projection maps  $(x, y) \in X \times Y \mapsto x \in X$  and  $(x, y) \in X \times Y \mapsto y \in Y$ .

It is easy to show that a coupling of  $\mu$  and  $\nu$  is equivalent to

- (i) For all measurable sets  $A \subset X$  and  $B \subset Y$ , one has  $\gamma[A \times Y] = \mu[A]$  and  $\gamma[X \times B] = \nu[B]$ .
- (ii) For all integrable measurable functions  $u, v$  on  $X, Y$ ,

$$\int_{X \times Y} (u(x) + v(y)) d\gamma(x, y) = \int_X u(x) d\mu(x) + \int_Y v(y) d\nu(y)$$

A trivial coupling of  $(X, \mu)$  and  $(Y, \nu)$  is given by the product measure  $(X \times Y, \mu \times \nu)$ . The symbol  $\rho \rightarrow \psi$  defined in eq. (1.9) means that the measure  $\gamma = |\psi(x_1, \dots, x_N)|^2 dx_1 \dots dx_N$  is a coupling, for instance, between  $\rho_1$  and  $\rho_{k-1}$ . In the classical theory of 2-marginals optimal transport [4, 97] a coupling is also called a transport plan.

In the case 2-molecular system with Coulomb interactions, up to passage to the limit  $\hbar \rightarrow 0$ , the theorem 1.5 states that there exist a special coupling  $\gamma$ , minimizer of (1.17), which is concentrated in the graph of measurable function. That motivates the next definition

**Definition 2.2** (Deterministic Coupling). *A coupling  $\gamma$  of two probability spaces  $(X, \mu)$  and  $(Y, \nu)$  is said to be deterministic if there exists a measurable function  $T : X \rightarrow Y$  such that  $\gamma = (Id, T)_\# \mu$ .*

In that terminology, the main question is to understand when the coupling given by the DFT-Optimal Transport (1.19) is deterministic. In this terminology, Seidl’s conjecture can be rephrased in the following way: *Does exist a deterministic coupling in the Optimal Transportation problem (1.19) for Coulomb costs?*

**2.2. Multimarginal Optimal Transportation Problem.** The “*semi-classical limit*” in the 2 particles case (thm. 1.5) and Seidl’s Conjecture (1.7), motivate us to introduce a more general coupling problem.

We denote by  $N$  by the number of particles (marginals),  $d$  the dimension of the space where those particles live (typically is the physical space  $d = 3$ ) and  $I$  a index set of cardinality  $|I| = N$ . Consider by a function  $x_i : \mathbb{R}^d \rightarrow \mathbb{R}$  the position of a single particle. The configuration of a system of  $N$ -particles is given by a vector  $x = (x_1, x_2, \dots, x_N)$  or, more precisely, by a function  $x : I \rightarrow \mathbb{R}^d$ .

Suppose that we fix the distribution of the  $i$ -th particle, given by a density function  $\rho_i : \mathbb{R}^d \rightarrow [0, 1]$ ,  $i \in I$  and  $c : \mathbb{R}^{Nd} \times \mathbb{R}^{Nd} \rightarrow [0, \infty]$ , a continuous cost function. We set  $\mu_i = \rho_i \mathcal{L}^d, \forall i \in I$  where  $\mathcal{L}^d$  is the Lebesgue measure on  $\mathbb{R}^d$  and we want to characterize the optimal coupling of the Monge-Kantorovich problem

$$(\mathcal{MK}) \quad \inf \left\{ \int_{\mathbb{R}^{dN}} c(x_1, \dots, x_N) d\gamma(x_1, \dots, x_N) : \gamma \in \Pi(\mathbb{R}^{dN}, \mu_1, \dots, \mu_N) \right\}, \quad (2.1)$$

pb:MKM

where  $\Pi(\mathbb{R}^{dN}, \mu_1, \dots, \mu_N)$  denotes the set of couplings  $\gamma(x_1, \dots, x_N)$  having  $\mu_i$  as marginals, for all  $i \in I$ , i.e. the set  $\Pi(\mathbb{R}^{dN}, \mu_1, \dots, \mu_N) = \{\gamma \in \mathcal{P}(\mathbb{R}^{dN}) : (e_i)_\# \gamma = \mu_i, \forall i \in I\}$  and, for all  $i \in I$ ,  $e_i : \mathbb{R}^{dN} \rightarrow \mathbb{R}^d$  is the projection on the  $i$ -th component and  $\mathcal{P}(\mathbb{R}^{dN})$  denotes the set of probability measures in  $\mathbb{R}^{dN}$ . In the DFT-OT problem it was mentioned that a interesting case is when all marginals are the same and equal to a measure  $\mu$  on  $\mathbb{R}^d$ . In that case, we denote the set of transport plans simply by  $\Pi_N(\mu)$ .

Notice that the existence of a such minimum in (2.1) is not a big deal and it is quite standard in Optimal Transport Theory. Indeed,  $\Pi(\mathbb{R}^{dN}, \mu_1, \dots, \mu_N)$  is *trivially non empty*, since the independent coupling  $\mu_1 \times \dots \times \mu_N \in \Pi(\mathbb{R}^{dN}, \mu_1, \dots, \mu_N)$ ; the set  $\Pi(\mathbb{R}^{dN}, \mu_1, \dots, \mu_N)$  is convex and compact for the weak\*-topology thanks to the imposed marginals; and moreover quantity to be minimized  $\gamma \mapsto \int c d\gamma$  is linear with respect to  $\mu$ . Hence, we can guarantee the existence of such a minimum of (2.1) by imposing a very weak hypothesis on the cost function, such as lower-semicontinuity.

However, we are interested in characterize some class of optimal transport plans or, at least, understand when that optimal coupling could be *deterministic*. Our problem so, is to study when minimizers of  $(\mathcal{MK})$  corresponds to Multi-marginal Monge minimizers

$$(\mathcal{M}_N) \quad \min \left\{ \int_{(\mathbb{R}^d)^N} c(x, T_1(x), \dots, T_{N-1}(x)) d\mu_1(x) : T = (T_i)_{i \in I} \in \mathcal{T}_N \right\}, \quad (2.2)$$

pb:MN

where denote by  $\mathcal{T}_N = \{T = (T_i)_{i \in I} : T_i \text{ is a Borelian map such that } T_{i\#} \mu_1 = \mu_i, \forall i \in I\}$ . As usual, it is easy to see that the set of *admissible plans* in (2.1) “includes” the set of *Monge transport maps* in (2.2): in fact, given a transport map  $T = (Id, T_1, \dots, T_{N-1})$ , we can consider a measures  $\gamma_T$  on  $(\mathbb{R}^d)^N$  defined by  $\gamma_T = (Id \times T_1 \times \dots \times T_{N-1})_\# \mu$ . Let  $h : (\mathbb{R}^d)^N \rightarrow \mathbb{R}$  be a  $\gamma_T$ -measurable function, then

$$\int_{(\mathbb{R}^d)^N} h(x_1, \dots, x_N) d\gamma_T(x_1, \dots, x_N) = \int_{\mathbb{R}^d} h(x_1, T_1(x_1), \dots, T_{N-1}(x_1)) d\mu_1(x_1);$$

and if we have  $\gamma_T$ -measurable function  $f : \mathbb{R}^d \rightarrow [0, \infty]$

$$\int_{(\mathbb{R}^d)^N} f(x_i) d\gamma_T(x_1, \dots, x_N) = \int_{\mathbb{R}^d} f(x_i) d\mu_i(x_i),$$

and so,  $\gamma_T$  is a transport plan. In particular, the value of the minimization among Monge-Kantorovich couplings is smaller or equal to the value of the minimization among the deterministic couplings

$$(\mathcal{MK}) \leq (\mathcal{M}_N).$$

Rather than in (2.1), the existence of the optimal deterministic coupling in (2.2) is not obvious and the difficulties do not lie on the “multi-marginality” of the problem. In fact, at that point, it is not possible to apply standard methods in Calculus of Variations even in the classical 2-marginals case in order to guarantee the existence of the minimum (2.2).

Intuitively, the difference between the two problems (2.1) and (2.2) is that in  $\mathcal{M}_N$  almost every point  $x_1 \in \mathbb{R}^d$  is coupled with exactly one point  $x_i \in \mathbb{R}^d$  for each  $i \in \{1, \dots, N-1\}$  whereas throughout the problem (2.1) we allow the splitting of mass in the “transportation” process in between two or more target points in  $\mathbb{R}^d$  for  $i \in \{1, \dots, N-1\}$ .

The problem (2.1), when  $N = 2$ , was studied by L. Kantorovich in 1940's [53, 54], without being aware of the famous article of Monge [75]. In 1975, L. Kantorovich won the Nobel Prize together with C. Koopmans “for their contributions to the theory of optimum allocation of resources”.

A important disadvantage in using the relaxed approach (2.1) in the multi-marginal OT is that the set of optimal transportation plans could be very large and we could need to select some special class of transportation plans.

In the 2-marginal setting, the two formulations (2.1) and (2.2) are proved to be equivalent for general Polish Spaces by Pratelli in [86].

**Theorem 2.3** (Monge equals Monge-Kantorovich in the 2-marginals case, [86]). *Let  $(X, \mu), (Y, \nu)$  be Polish spaces, where  $\mu \in \mathcal{P}(X)$  and  $\nu \in \mathcal{P}(Y)$  are non-atomic probability measures. If  $c : X \times Y \rightarrow \mathbb{R} \cup \{\infty\}$  is a continuous cost, then*

$$\min \left\{ \int_{X \times Y} c(x, y) d\gamma(x, y) : \gamma \in \Gamma(X \times Y) \right\} = \inf \left\{ \int_X c(x, T(x)) d\mu(x), : \begin{array}{l} T\# \mu = \nu, \\ T : X \rightarrow Y \text{ Borel} \end{array} \right\}$$

where  $\Gamma(X \times Y)$  denotes the set of 2-marginals transport plans  $\gamma \subset X \times Y$  having  $\mu$  and  $\nu$  as marginals.

In the particular case of the multi-marginal OT problem in (2.2) and (2.1) when all marginals are equal to  $\rho \mathcal{L}^d$ , we can apply the previous theorem (2.3) on the Polish spaces  $X = \mathbb{R}^d, Y = (\mathbb{R}^d)^{N-1}$  with measures  $\mu = \rho \mathcal{L}^d$  and  $\nu = (e_2, \dots, e_n)\# \gamma$ , for every  $\gamma \in \Gamma(\mathbb{R}^{dN}, \rho \mathcal{L}^d)$ , and obtain the following corollary

**Corollary 2.4.** *Let is  $\mu_1 = \dots = \mu_N = \rho \mathcal{L}^d$  be probability measures on  $\mathbb{R}^d$  with density  $\rho$  and  $c : (\mathbb{R}^d)^N \rightarrow \mathbb{R}$  be a continuous cost function. Then,*

$$(\mathcal{MK}) = (\mathcal{M}_N).$$

Notice that, in general, the equivalence between (2.1) and (2.2) is not a immediate consequence of the theorem 2.3. In particular, remark that the image measure of  $Y = \mathbb{R}^{d(N-1)}$  is not prescribed, but only its marginals.

Finally, we conclude this section giving an example of a particular continuous cost function where the optimal coupling  $\gamma$  is not deterministic, even in the 2 particles case. The nonexistence of an optimal transport maps typically happens when minimizing sequences of  $\gamma$  exhibit strong oscillations.

*Example 2.5* (G. Carlier). Suppose  $\mu$  uniformly distributed in  $[0, 1]$ ,  $\nu$  uniformly distributed in  $[-1, 1]$  and the continuous cost  $c : \mathbb{R}^2 \rightarrow \mathbb{R}$  given by

$$c(x_1, x_2) = (x_1^2 - x_2^2)^2.$$

Since  $c \geq 0$  we have that an optimal plan for this cost is given by  $\bar{\gamma} = \mu \otimes (\frac{1}{2}\delta_x + \frac{1}{2}\delta_{-x})$  since  $\int c d\bar{\gamma} = 0$ . We want to show there doesn't exit a map  $T$  such that the cost attains zero: suppose in fact that  $\int (x^2 - T(x)^2)^2 dx = 0$  but then one should have  $T(x) \in \{x, -x\}$  almost everywhere. Hence,

$$T(x) = (\chi_A - \chi_{[0,1] \setminus A})x$$

for a measurable  $A \subset [0, 1]$ . But then  $T_{\#}\mu = \mathcal{L}^1|_A + \mathcal{L}^1|_{\tilde{A}} \neq \nu$ , where  $x \in \tilde{A}$  iff  $-x \in [0, 1] \setminus A$ , and so we reach an absurd since  $T$  wouldn't be admissible. It is however easy to construct admissible maps  $T_n$ , linear on  $[i/2n, (i+1)/2n]$ , such that  $\int c(x_1, T_n(x_1)) d\mu(x_1)$  tends to zero.

Remark that, we state by simplicity the example in the 2-marginal case, since the same conclusion and arguments apply in the larger interactions case under the cost

$$c(x_1, \dots, x_N) = \sum_{i=1}^N \sum_{j=i+1}^N (x_i^2 - x_j^2)^2.$$

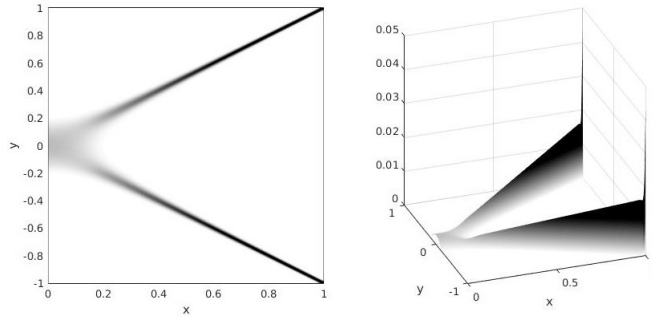


FIGURE 1. (example ??) Left: support of the optimal coupling  $\bar{\gamma}$ . Right: optimal coupling  $\bar{\gamma}$ . The simulation has been performed by using the iterative method described in section 6.

figure:carlier

**2.3. Dual formulation.** We consider the formal convex dual formulation of (2.1)

$$(\mathcal{K}_N) \quad \sup \left\{ \int_{\mathbb{R}^d} \sum_{i=1}^N u_i(x_i) d\mu_i(x_i) : \sum_{i=1}^N u_i(x_i) \leq c(x_1, \dots, x_N) \right\}, \quad (2.3)$$

pb:KN

where  $u_i \in L^1(\mathbb{R}^d, \mu_i)$ , called Kantorovich potentials or simply potentials, are upper semicontinuous for all  $i \in \{1, \dots, N\}$ . This is often called the dual problem of (2.2).

Notice that, when all marginals are equal to  $\mu_1 = \rho(x)\mathcal{L}^d$  and the cost is symmetric, we can assume the Kantorovich potentials  $u_i(x_i)$  are all the same  $u(x_i)$ , and we can rewrite the constraint in (2.3) as

$$\sum_{i=1}^N u(x_i) \leq c(x_1, \dots, x_N).$$

Among all Kantorovich potentials, a particular class is of special interests.

**Definition 2.6** (*c-conjugate function*). Let  $c : (\mathbb{R}^d)^N \rightarrow \mathbb{R} \cup \{\infty\}$  a Borel function. We say that the  $N$ -tuple of functions  $(u_1, \dots, u_N)$  is a  $c$ -conjugate function for  $c$ , if

$$u_i(x_i) = \inf \left\{ c(x_1, \dots, x_N) - \sum_{j=1, j \neq i}^N u_j(x_j) : x_j \in X_j, j \neq i \right\} \quad \forall i = 1, \dots, N$$

As we will see in the next theorem 2.7, the optimal potentials  $u_1, \dots, u_N$  in  $(\mathcal{K}_N)$  are  $c$ -conjugate, and therefore semi-concave. So, they admits super-differentials  $\bar{\partial}u_i(x_i)$  at each point  $x_i \in \mathbb{R}^d$  and then, at least in the compact case, we can expect that for each  $x_i \in \mathbb{R}^d$ , there exists  $x_j \in \mathbb{R}^d, j \neq i$ , such that

$$u(x_i) = c(x_1, \dots, x_N) - \sum_{j \neq i} u_j(x_j).$$

Moreover, if  $c : (\mathbb{R}^d)^N \rightarrow \mathbb{R}$  is differentiable and  $|u_i(x_i)| < \infty$ , for a  $x_i \in \mathbb{R}^d$ ,  $u_i$  is locally Lipschitz. The well-know general result linking both problems (2.1) and (2.3) are proved by Keller in 1984.

er

**Theorem 2.7** (Kellerer, [55]). Let  $(X_1, \mu_1), \dots, (X_N, \mu_N)$  be Polish spaces equipped with Borel probability measures  $\mu_1, \dots, \mu_N$ . Consider  $c : X_1 \dots X_N \rightarrow \mathbb{R}$  be a Borel cost function and assume that  $\bar{c} = \sup_X c < \infty$  and  $\underline{c} = \inf_X c > -\infty$ . Then,

- (i) There exists a solution  $\gamma$  to the Monge-Kantorovich problem (2.1) and a  $c$ -conjugate solution  $(u_1, u_2, \dots, u_N)$  to its dual problem (2.3).
- (ii) “Duality holds”,

$$\inf \left\{ \int_{X_1 \dots X_N} c d\gamma : \gamma \in \Pi(\mu_1, \dots, \mu_N) \right\} = \sup \left\{ \sum_{i=1}^N \int_{X_i} u_i(x_i) d\mu_i : u_i \in \mathcal{D} \right\}, \quad (2.4)$$

where  $\mathcal{D}$  is the set of functions  $u_i : \mathbb{R}^d \rightarrow \mathbb{R}, i = 1, \dots, N$ , such that

$$\sum_{i=1}^N u_i(x_i) \leq c(x_1, \dots, x_N), \quad \text{and} \quad \frac{1}{N} \bar{c} - (\bar{c} - \underline{c}) \leq \sum_{i=1}^N u_i(x_i) \leq \frac{1}{N} \bar{c}.$$

- (iii) For any solution  $\gamma$  of (2.1), any conjugate solution of (2.3) and any  $(x_1, \dots, x_N) \in \text{spt}(\gamma)$ , we have

$$\sum_{i=1}^N u_i(x_i) = c(x_1, \dots, x_N).$$

We remark that despite the great generality of Kellerer’s theorem, it can not be applied directly to Coulomb-type costs, since Coulomb cost does not satisfies the boundedness hypothesis  $\bar{c} = \sup_X c < \infty$  and, in addition, the cost  $c$  can not be bounded by functions in  $L^1(\mathbb{R}^d, \rho \mathcal{L}^d)$

$$\sum_{i=1}^N \sum_{j=i+1}^N \frac{1}{|x_j - x_i|} \leq \sum_{i=1}^N u(x_i).$$

A generalization of Keller’s theorem due to M. Beiglboeck, C. Leonard and W. Schachermayer [10], extends this duality for more general costs, but also cannot be applied in this context.

Recently, De Pascale [36] extended the proof of duality which can be applied also for Coulomb-type costs, see Theorem 3.1. In fact, suppose  $\mu = \rho(x) \mathcal{L}^d$  a probability measure



which does not give mass to sets of cardinality smaller or equal than  $d - 1$ , then

$$\min \left\{ \int_{(\mathbb{R}^d)^N} \sum_{i=1}^N \sum_{j=i+1}^N \frac{1}{|x_i - x_j|} d\gamma : \gamma \in \Pi_N(\mu) \right\} = \max \left\{ N \int_{\mathbb{R}^d} u(x) \rho(x) dx : u \in \tilde{\mathcal{D}} \right\} \quad (2.5) \quad \boxed{\text{eq:DualDeP}}$$

where  $\tilde{\mathcal{D}}$  is the set of potentials  $u \in L^1(\mathbb{R}^d, \mu)$  such that

$$\sum_{i=1}^N u(x_i) \leq \sum_{i=1}^N \sum_{j=i+1}^N \frac{1}{|x_j - x_i|}, \quad \otimes_{i=1}^N \rho - \text{almost everywhere.} \quad (2.6) \quad \boxed{\text{eq:ColConst}}$$

Moreover, we get the conclusions (i) and (iii) of the theorem 2.7. Physically, the constraint (2.6) means that, at optimality, the allowed  $Nd$  configuration space should be a minimum of the classical potential energy [16].

*Remark 2.8.* More generally, the main theorem of [36] show that the Kantorovich duality holds for cost of form

$$\sum_{i=1}^N \sum_{j=i+1}^N \frac{1}{|x_j - x_i|^s}, \quad s \geq 1.$$

**2.4. Geometry of the Optimal Transport sets.** In this section, our goal is to present the relation between the support of a coupling  $\gamma$  and optimality in the Monge-Kantorovich problem (2.1). In the next, rather than give a complete presentation, we will just summarize key results necessary to present recent developments of this theory applied to Coulomb-type and repulsive harmonic-type costs.

We now define a natural generalization of the concept of  $c$ -cyclically monotone sets in the multi-marginal setting of Optimal Transportation Theory, proposed by Brendan Pass [79].

1lymonotoneset

**Definition 2.9** ( $c$ -cyclically monotone set). *Let  $c : \mathbb{R}^{dN} \rightarrow \mathbb{R}$  be a cost function. A subset  $\Gamma \subset \mathbb{R}^{dN}$  is said to be  $c$ -monotone with respect to a partition  $p \subset I$  if, for all  $x = (x_1, \dots, x_N), y = (y_1, \dots, y_N) \in \Gamma$*

$$c(x) + c(y) \leq c(X(x, y, p)) + c(Y(x, y, p))$$

where  $X(x, y, p), Y(x, y, p) \in \mathbb{R}^N$  are functions obtained from  $x$  and  $y$  by exchanging their coordinates on the complement of  $p$ , namely

$$X_i(x, y, p) = \begin{cases} x_i, & \text{if } i \in p \\ y_i, & \text{if } i \in p^c \end{cases} \quad \text{and} \quad Y_i(x, y, p) = \begin{cases} y_i, & \text{if } i \in p \\ x_i, & \text{if } i \in p^c \end{cases} \quad \forall i \in I$$

The notion of  $c$ -cyclical mononotonicity in the multi marginal setting was study by Ghoussoub, Moameni, Maurey, Pass [49, 50, 79, 80] and successfully applied in the study of decoupling PDE systems [51]. Recently, Beiglbock and Griessler [7], presented a new notion called finistic optimality inspired in the martingale optimal transport, which in some cases plays an equivalent role of cyclical monotonicity.

The next proposition gives a necessary condition on the support for an optimal transport plan  $\gamma$ . For the proof, which is based on the 2-marginal result [95], we invite the reader to see the Lemma 3.1.2 of [79].

one

**Proposition 2.10** (Support of transport plan are  $c$ -cyclically monotonic). *Let  $c : (\mathbb{R}^d)^N \rightarrow [0, \infty]$  be a continuous cost and  $\mu_1, \dots, \mu_N$  absolutely continuous probability measure on  $\mathbb{R}^d$ . Suppose that  $\gamma \in \Pi((\mathbb{R}^d)^N, \mu_1, \dots, \mu_N)$  is an optimal transport plan for Multimarginal Monge-Kantorovich Problem (2.1) and assume  $(\mathcal{MK}) < \infty$ . Then  $\gamma$  is concentrated in a  $c$ -monotone set.*

sym

**2.5. Symmetric Case.** We are going to show in this section that we can reduce to study the problem 2.2 and 2.1, respectively, in a suitable class of transport maps (called cyclic maps) and transport plans (symmetric plans). Having in mind the DFT-OT problem, it is natural to suppose that the cost functions does not change if there are some symmetries among the configuration of the particles or, equivalently, we can suppose the measures defined by the densities  $\rho_i$  are cyclically symmetric.

**Definition 2.11** (Symmetric probability measure). *Let  $\mu$  be a probability measure in  $\mathbb{R}^d$ . We say that  $\mu$  is symmetric if  $\mu = (-\text{Id})_{\#}\mu$ .*

**Definition 2.12** (Cyclically symmetric probability measures). *A probability measure  $\gamma \in \mathcal{P}(\mathbb{R}^{dN})$  is cyclically symmetric if*

$$\int_{\mathbb{R}^{dN}} f(x_1, \dots, x_N) d\gamma = \int_{\mathbb{R}^{dN}} f(\sigma(x_1, \dots, x_N)) d\gamma, \quad \forall f \in \mathcal{C}(\mathbb{R}^{dN})$$

where  $\sigma$  is a permutation  $\sigma \in S_N$ . We will denote by  $\Pi^{sym}(\mathbb{R}^{dN}, \mu_1, \dots, \mu_N)$  the set of symmetric transport plans  $\gamma \in \Pi^{sym}(\mathbb{R}^{dN}, \mu_1, \dots, \mu_N)$  and we simply denoted by  $\Pi_N^{sym}(\mu)$ , the space of all symmetric probability measures on  $\mathbb{R}^{dN}$  having the same marginal  $\mu$ .

Notice that if  $(\mu_1, \dots, \mu_N)$  are marginals of  $\gamma \in \Pi^{sym}(\mathbb{R}^{dN}, \mu_1, \dots, \mu_N)$ , then

$$\begin{aligned} \int_{\mathbb{R}^d} f(x_i) d\mu_i(x_i) &= \int_{\mathbb{R}^{dN}} f(x_1, \dots, x_N) d\pi(x_1, \dots, x_N) \\ &= \int_{\mathbb{R}^{dN}} f(\sigma^{N+1-i}(x_1, \dots, x_N)) d\pi(x_1, \dots, x_N) \\ &= \int_{\mathbb{R}^{dN}} f(x_1) d\mu_i(x_1, \dots, x_N) \\ &= \int_{\mathbb{R}^d} f(x_1) d\mu_1(x_1). \end{aligned}$$

Briefly,

$$(e_i)_{\#}\pi = \mu_1, \forall i \in I,$$

where  $e_i$  is the canonical projection.

Analogously in the problem (2.2) and (2.1), we can define the (multimarginal) symmetric Monge-Kantorovich Problem as follows:

$$(\mathcal{MK}_{sym}) \quad \inf \left\{ \int_{\mathbb{R}^{dN}} c(x_1, \dots, x_N) d\pi(x_1, \dots, x_N) : \gamma \in \Pi_N^{sym}(\mu) \right\}, \quad (2.7) \quad \boxed{\text{pb:MKsym}}$$

and the associated multimarginal cyclically Monge Problem

$$(\mathcal{MK}_{cycl}) \quad \inf \left\{ \int_{\mathbb{R}^d} c(x, T_1(x), \dots, T_{N-1}(x)) : T \in \mathcal{T}^{cycl}(\mathbb{R}^{dN}, \mu) \right\},$$

where  $\mathcal{T}^{cycl}(\mathbb{R}^{dN}, \mu)$  denotes the set of maps  $T = (Id, T_1, \dots, T_{N-1})$  such that  $T_i : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a Borel map,  $T_{i\#}\mu = \mu$ ,  $T_{i+1}(x) = T^{(i)}(x)$  for every  $i \in \{1, \dots, N-1\}$  and  $T^{(N)} = Id$ .

**Proposition 2.13.** *Suppose  $\mu \in \mathcal{P}(\mathbb{R}^d)$  an absolutely continuous probability measure with respect to the Lebesgue measure and  $c : (\mathbb{R}^d)^N \rightarrow \mathbb{R}$  be a continuous symmetric cost. Then,*

$$(\mathcal{MK}) = (\mathcal{MK}_{sym}).$$

*Proof.* The fact that the infimum of (2.1) is smaller or equal than (2.7) is obvious, we will simply write  $\mathcal{MK} \leq \mathcal{MK}_{sym}$ . Now, we need to show that for every transport plan  $\gamma \in \Pi_N(\mu)$  we can associate a symmetric plan in  $\Pi_N^{sym}(\mu)$ . Indeed, given  $\gamma \in \Pi_N(\mu)$ , we define

$$\bar{\gamma} = \frac{1}{N!} \sum_{\sigma \in S_N} \sigma_{\#} \gamma,$$

and finally we notice that  $\bar{\gamma}$  has the same cost of  $\gamma$ .  $\square$

We remark that is not so obvious do the same reasoning for the minimum problems  $\mathcal{M}_{cycl}$  and  $\mathcal{M}_N$ . We know only that  $\mathcal{M}_N \leq \mathcal{M}_{cycl}$ . Finally, the fact of  $(\mathcal{M}_{sym} \leq \mathcal{M}_{cycl})$ , was proved by M. Colombo and S. Di Marino [28]. We give the precise statement of the theorem.

CoDMa

**Theorem 2.14** (M. Colombo and S. Di Marino, [28]). *Let  $c : (\mathbb{R}^d)^N \rightarrow \mathbb{R}$  be a continuous function and bounded from below. If  $\mu$  has no atoms, then*

$$(\mathcal{MK}_{sym}) = (\mathcal{M}_{cycl}).$$

Finally, we have the equality between the Multi-marginal cyclic problem and Multi-marginal OT problem.

cor:CoDMa

**Corollary 2.15.** *If  $\mu$  has no atoms and  $c : (\mathbb{R}^d)^N \rightarrow \mathbb{R}$  is a continuous function and bounded from below, then*

$$(\mathcal{MK}_{sym}) = (\mathcal{MK}) = (\mathcal{M}) = (\mathcal{M}_{cycl}).$$

The last result of this section is due Pass [82], which says that optimal transport plans for symmetric costs can be supported on sets of Hausdorff dimension equal or bigger than  $2N - 2$ .

th:sym:Pass

**Theorem 2.16** (Pass, [82]). *Let  $c : (\mathbb{R}^d)^N \rightarrow \mathbb{R}$  be a symmetric cost and  $\{\mu_i\}_{i=1}^N$  be radially symmetric and absolutely continuous with respect to Lebesgue measure in  $\mathbb{R}^d$ . Suppose that for every radii  $(r_1, \dots, r_N)$  the minimizers*

$$(x_1, \dots, x_N) \in \operatorname{argmin}_{|y_i|=r_i} c(y_1, y_2, \dots, y_N) \quad \text{are not all co-linear.}$$

*Then, there exists solutions  $\gamma$  for the Monge-Kantorovich symmetric problem (2.7) whose support is at least  $2N - 2$ -dimensional.*

sec:Coulomb

### 3. MULTI-MARGINAL OT PROBLEM WITH COULOMB COST

This section is devoted to the summary of the results present in literature on the multi marginal optimal transport with Coulomb cost. We already highlighted the fact that the interesting case is when all the marginals are equal, since this is the physical case. We point out that, as always, the focus will be on characterizing the optimal plans as well as looking for cyclical transport maps: in this direction an important conjecture has been made by Siedl in its seminal paper [?], where he gave an explicit construction of a map for measures with radial symmetry in  $\mathbb{R}^d$ . In the last years the conjecture was proven to be right: first in the 1D case when  $\mu$  is the uniform measure on an interval, and then generalized to any diffuse measure on the real line. However quite recently various authors disproved the conjecture in the case  $\mathbb{R}^d$  for  $d \geq 2$  and for  $N = 3$  marginals, yet the Seidl map is still optimal for some class of measures.

We begin by analyzing the general problem, and then we will proceed to look at the radial case.

**3.1. General theory: duality, equivalent formulations and many particles limit.** We begin stating the duality theorem in this case proved by De Pascale [36], by means of  $\Gamma$ -convergence of finite dimensional problems for which the duality is classical.

DeP

**Theorem 3.1.** *Let  $\mu \in \mathcal{P}(\mathbb{R}^d)$  be a measure that is not concentrated on a set of cardinality smaller or equal than  $N - 1$ . Then the duality (2.5) holds and the dual problem has a bounded maximizer.*

The boundedness of the maximizer  $u$  can let us prove also some regularity properties. The first one is the fact that any optimal plan is supported away from the diagonals, while the second one proves second order regularity of  $u$ .

nal

**Lemma 3.2** (Corollary 3.13 in [36]). *Let  $\mu$  as in Theorem 3.1; then there exists  $\alpha > 0$  such that for every optimal plan  $\gamma$  we have  $\text{supp}(\gamma) \subseteq \{(x_1, \dots, x_N) : |x_i - x_j| \geq \alpha \forall i \neq j\}$ .*

pot

**Lemma 3.3.** *Let  $\mu$  as in Theorem 3.1 and  $u$  be a bounded maximizer in the dual problem in (2.5); then  $u$  has a  $\mu$ -representative that is Lipschitz and semi concave.*

*Proof.* As in the classical case we first prove a structural property of the maximizer  $u$ , namely that the constraint (2.6) is saturated, that is  $(u, \dots, u)$  is a  $c$ -conjugate  $N$ -tuple. Let us consider

$$w(x_1) = \underset{(x_2, \dots, x_N) \in \mathbb{R}^{d(N-1)}}{\text{ess-inf}} \left\{ \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} - \sum_{i=2}^N u(x_i) \right\},$$

where the  $\text{ess-inf}$  is made with respect to the measure  $\mu^{\otimes N-1}$ . It is obvious that  $w(x_1) \geq u(x_1)$  for  $\mu$ -a.e  $x_1$ , thanks to (2.6). Suppose that  $w > u$  in a set of positive measure; but then  $(w, u, \dots, u)$  would be a better competitor in the (asymmetrized) dual problem contradicting the fact that  $u$  is a maximizer (we use the fact that the symmetrized dual problem and the asymmetrized one have the same values). Now we have that  $w$  is defined everywhere and so we can talk about its regularity.

Let  $d > 0$  be a number such that there exists points  $p_1, \dots, p_N$  in the support of  $\mu$  such that  $|p_i - p_j| \geq d$ ; this number exists as long as  $\sharp \text{supp}(\mu) \geq N$ . Let us fix  $\varepsilon > 0$  such that  $\frac{1}{2\varepsilon} > \frac{2N(N-1)}{d} + 2N\|u\|_\infty$  and  $\varepsilon \leq d/8$ ; in this way it is true that if we define

$$w_{x_0}(x_1) = \underset{|x_i - x_0| \geq \varepsilon, i=2, \dots, N}{\text{ess-inf}} \left\{ \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} - \sum_{i=2}^N u(x_i) \right\},$$

we have  $w_{x_0} = w$  on  $B(x_0, \varepsilon)$ . In fact in the definition of  $w$  we can choose  $x_i$  among the  $p_i$  (or very close to them, as they belong to the support) such that  $|x_i - x_j| > d/2 - 2\varepsilon \geq d/4$  for every  $i \neq j$ , and so we have that  $w \leq \frac{2N(N-1)}{d} + N\|u\|_\infty$  but then it is clear that for any  $(x_2, \dots, x_N)$  such that  $|x_i - x_0| \leq \varepsilon$  we will have

$$\sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} - \sum_{i=2}^N u(x_i) \geq \frac{1}{2\varepsilon} > w(x_1) \quad \forall x_1 \in B(x_0, \varepsilon);$$

this proves that in fact  $w_{x_0} = w$  on the set  $B(x_0, \varepsilon)$  and so in particular also in  $B(x_0, \varepsilon/2)$ . But in this set  $w_{x_0}$  is Lipschitz and semi concave since it is an infimum of uniformly  $C^\infty$  functions

on  $B(x_0, \varepsilon/2)$ . Moreover the bounds on the first and second derivatives don't depend on  $x_0$  but only on  $\varepsilon$ , that is fixed a priori, and so by a covering argument we obtain the thesis.  $\square$

Another interesting reformulation of the Coulomb-like problem, or more generally when we have only interaction between two particles can be found in [42], where, seeking a dimensional reduction of the problem, the authors use the fact that if  $\gamma \in \Pi_N(\mu)$  is a symmetric plan then

$$\int_{\mathbb{R}^{Nd}} \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} d\gamma = \binom{N}{2} \int_{\mathbb{R}^{2d}} \frac{1}{|x - y|} d(\pi_1, \pi_2)_\# \gamma,$$

that is, since the electron are indistinguishable, it is sufficient to look at the potential energy of a couple of electron and then multiply it by the number of couples of electron. It is clear that  $\gamma_2 = (\pi_1, \pi_2)_\# \gamma$  is a 2-plan whose marginal are still  $\mu$ ; we will say that  $\eta \in \Pi_2(\mu)$  is  $N$ -representable whenever it exists  $\gamma \in \Pi_N(\mu)$  such that  $\eta = (\pi_1, \pi_2)_\# \gamma$ . The equivalent formulation they give is

$$\min_{\gamma \in \Pi_N(\mu)} \int_{\mathbb{R}^{Nd}} \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} d\gamma = \min_{\substack{\eta \in \Pi_2(\mu) \\ \eta \text{ is } N\text{-representable}}} \binom{N}{2} \int_{\mathbb{R}^{2d}} \frac{1}{|x - y|} d\eta.$$

Unfortunately the conditions for being  $N$ -representable are not explicit and they are very difficult to find; this is correlated to the  $N$ -representability for matrices, but here, since we are in the semiclassical limit, we deal with densities instead. In [42] the authors propose, as a method for reduce the dimensionality of the problem, to substitute the condition of being  $N$ -representable with that of being  $k$ -representable with  $k \leq N$ ; the resulting problem will give a lower bound to the real SCE functional.

We present also a general theorem regarding the many particles limit, that embodies the well-known fact that when  $N \rightarrow \infty$  then the solution is the mean field one

**Theorem 3.4** ([34]). *Let  $\mu \in \mathcal{P}(\mathbb{R}^d)$ . Then we have that*

$$\lim_{N \rightarrow \infty} \frac{1}{\binom{N}{2}} \min_{\gamma \in \Pi_N(\mu)} \int_{\mathbb{R}^{Nd}} \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} d\gamma = \int_{\mathbb{R}^{2d}} \frac{1}{|x - y|} d\mu \otimes \mu.$$

*In terms of DFT we are saying that  $E^{OT}(\mu) = \binom{N}{2} E^{MF}(\mu) + o(N^2)$ , where  $E^{MF}$  is the normalized mean field energy  $E^{MF}(\mu) = \int \frac{1}{|x-y|} d\mu(x) d\mu(y)$ .*

*Remark 3.5.* The statement about the Coulomb cost in the physical case is quite classical. In fact, for measure  $\rho \in L^{4/3}(\mathbb{R}^3)$  the Lieb-Oxford bound holds [66, 67]:

$$\binom{N}{2} E^{MF}(\rho) \geq E^{OT}(\rho) \geq N^2 E^{MF}(\rho) - CN^{4/3} \int_{\mathbb{R}^3} \rho^{4/3}(x) dx,$$

and so the conclusion is immediate. However, in [34] the proof is completely different and relies the fact that a measure  $\gamma \in \Pi_2(\mu)$  that is  $N$ -representable for every  $N$  must be in the convex envelope of the measure of the type  $\nu \otimes \nu$ , and then on a direct computation using the Fourier transform. In particular, aside from the Coulomb cost in dimension 3 they prove the theorem for more general costs with binary interaction of the form  $\sum_{1 \leq i < j \leq N} l(x_i - x_j)$ , where  $l \in C_b(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$  satisfies  $l(z) = l(-z)$  and also  $\hat{l} \geq 0$ .

**3.2. The Monge problem: deterministic examples and counterexamples.** In this section we will illustrate the case in which we know that there exists a deterministic solution; in those cases moreover there is also a result of uniqueness. We remark that these are the two extreme case, namely the case  $N = 2$  (and any  $d$ ) and the case  $d = 1$  (and any  $N$ ). We will just sketch the proofs in order to make clear the method used here, and why it is difficult to generalize it. We begin with the 2-particle case, in every dimension  $d$ : this result was proved in [33] by means of standard optimal transport techniques.

tar

**Theorem 3.6.** *Let  $\mu \in \mathcal{P}(\mathbb{R}^d)$  be a probability measure that is absolutely continuous with respect to the Lebesgue measure. Then there exists a unique optimal plan  $\gamma_O \in \Pi_2(\mu)$  for the problem*

$$\min_{\gamma \in \Pi_2(\mu)} \int_{\mathbb{R}^{2d}} \frac{1}{|x_1 - x_2|} d\gamma.$$

*Moreover this plan is induced by an optimal map  $T$ , that is,  $\gamma = (Id, T)_\# \mu$ , and  $T(x) = x + \frac{\nabla \phi}{|\nabla \phi|^{3/2}}$   $\mu$ -almost everywhere, where  $\phi$  is a Lipschitz maximizer for the dual problem.*

*Proof.* Let us consider  $\gamma$  a minimizer for the primal problem and  $\phi$  a bounded and Lipschitz maximizer of the dual problem (it exists thanks to Lemma 3.3). Then we know that

$$F(x_1, x_2) = \phi(x_1) + \phi(x_2) - |x_1 - x_2| \geq 0 \quad \text{for } \mu \otimes \mu\text{-almost every } (x_1, x_2),$$

and we know also that  $F = 0$   $\gamma$ -almost everywhere. But then  $F$  has a minimum on the support of  $\gamma$  and so  $\nabla F = 0$  in this set; in particular we have that  $\nabla \phi(x_1) = -\frac{x_1 - x_2}{|x_1 - x_2|^3}$  on the support of  $\gamma$ . Finding  $x_2$  we have that  $x_2 = x_1 - \frac{\nabla \phi}{|\nabla \phi|^{3/2}}(x_1) = T(x_1)$  on the support of  $\gamma$  and this implies  $\gamma = (Id, T)_\# \mu$  as we wanted to show (see for example).  $\square$

The first positive  $N$ -marginal result for the Coulomb cost is instead shown in [29] where, in dimension  $d = 1$ , the authors can prove that for non-atomic measure an optimal plan is always “induced” by a cyclical optimal map  $T$ .

1DN

**Theorem 3.7.** *Let  $\mu \in \mathcal{P}(\mathbb{R})$  be a diffuse probability measure. Then there exists a unique optimal symmetric plan  $\gamma_O \in \Pi_2^{sym}(\mu)$  that solves*

$$\min_{\gamma \in \Pi_N^{sym}(\mu)} \int_{\mathbb{R}^N} \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} d\gamma.$$

*Moreover this plan is induced by an optimal cyclical map  $T$ , that is,  $\gamma_O = \frac{1}{N!} \sum_{\sigma \in S_N} \sigma_\# \gamma_T$ , where  $\gamma_T = (Id, T, T^{(2)}, \dots, T^{(N-1)})_\# \mu$ . An explicit optimal cyclical map is*

$$T(x) = \begin{cases} F_\mu^{-1}(F_\mu(x) + 1/N) & \text{if } F_\mu(x) \leq (N-1)/N \\ F_\mu^{-1}(F_\mu(x) + 1 - 1/N) & \text{otherwise.} \end{cases}$$

*Here  $F_\mu(x) = \mu(-\infty, x]$  is the distribution function of  $\mu$ , and  $F_\mu^{-1}$  is its lower semicontinuous left inverse.*

*Proof.* We begin observing that if  $\gamma$  is a symmetric optimal plan then a stronger statement than Proposition 2.10 holds, namely we have that for every  $x, y \in \text{supp}(\gamma)$ :

$$c(x) + c(y) = \min\{c(X(x, \sigma(y), p)) + c(Y(x, \sigma(y), p)) : \forall p \subset \{1, \dots, N\}, \forall \sigma \in S_N\},$$

where  $c$  is the Coulomb cost on the  $N$ -tuple of points  $x = (x_1, \dots, x_N)$  and  $\sigma$  acts on the indices. The key point here is that one can show (Proposition 2.4 ??) that this property holds if and

only if a more geometric condition holds true for  $x$  and  $y$ : they are well-ordered. This property amounts to the fact that, up to permute the coordinates of both points, we have

$$x_1 \leq y_1 \leq x_2 \leq \dots \leq x_N \leq y_N,$$

or the other way around. Once this property is proved the rest of proof is rather linear, since we proved that the support of  $(\pi_1, \pi_2)_\# \gamma|_D$ , where  $D = \{x_i \leq x_j \text{ and } y_i \leq y_j, \forall i < j\}$ , is basically monotone in the usual sense where is the set where the coordinates are ordered. We remark that the key property does not easily generalize to  $d > 1$  since it (heavily) uses the ordered structure of  $\mathbb{R}$ .  $\square$

Let us notice that in this case we don't have uniqueness of optimal plan neither of cyclical optimal maps, as pointed out in Remark 1.2 in [29].

For the general case so we have to assume  $d > 1$  and  $N > 2$ : from now on we will look at the case in which the measure  $\mu$  has radial symmetry. We can easily see that in this case we can reduce our problem to a multi marginal problem in  $\mathbb{R}$ .

**Lemma 3.8** (Radial case). *Let  $\mu \in \mathcal{P}(\mathbb{R}^d)$  be a radially symmetric measure. In particular  $\mu$  is determined by  $\mu_r = |\cdot|_\# \mu$ . The for every optimal plan  $\gamma \in \Pi_N(\mu)$ , if we consider  $\gamma_r = R_\# \gamma$ , where  $R : (x_1, \dots, x_N) \mapsto (|x_1|, \dots, |x_N|)$ , we have that  $\gamma_r$  is an optimal plan for the 1D multi marginal problem*

$$\min_{\eta \in \Pi_N(\mu_r)} \int_{\mathbb{R}^N} c_1(r_1, \dots, r_N) d\eta. \quad (3.1) \quad \boxed{\text{eq:red_rad}}$$

where  $c_1$  is the reduced Coulomb cost

$$c_1(r_1, \dots, r_N) = \min \left\{ \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} : |x_i| = r_i \ \forall i = 1, \dots, N \right\}.$$

Moreover  $\int_{\mathbb{R}^{dN}} \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} d\gamma = C \int_{\mathbb{R}^d} c_1(r_1, \dots, r_N) d\gamma_r$ , for some  $C = C(d)$

Given the one dimensional character of this problem, one could expect that the solution is similar to the one depicted in Theorem 3.7. In fact in [94] the authors conjecture a similar structure:

**Conjecture 3.9** (Seidl). *Let  $\mu \in \mathcal{P}(\mathbb{R}^d)$  be an absolutely continuous measure with respect to the Lebesgue measure, with radial symmetry, and let  $\mu_r = |\cdot|_\# \mu$ . Let  $0 = r_0 < r_1 < \dots < r_{N-1} < r_N = \infty$  such that the intervals  $A_i = [r_i, r_{i+1})$  have all the same radial measure  $\mu_r(A_i) = 1/N$ . Then let  $F(r) = \mu_r(0, r]$  be the cumulative radial function and let  $S : [0, \infty) \rightarrow [0, \infty)$  be defined piecewise such that the interval  $A_i$  is sent in the interval  $A_{i+1}$  in an anti monotone way:*

$$S(r) = F^{-1}(2i/N - F(r)) \quad \text{if } r_{i-1} \leq r < r_i \text{ and } i < N$$

$$F(S(r)) = \begin{cases} F^{-1}(F(r) + 1/N - 1) & \text{if } N \text{ is even} \\ F^{-1}(1 - F(r)) & \text{if } N \text{ is odd,} \end{cases} \quad \text{if } r_{N-1} \leq r < r_N.$$

Then  $S$  is an optimal cyclical map for the problem (3.1).

However in the recent papers [31, 37] this conjecture is proved to be wrong, looking at radial measures  $\mu$  concentrated in some thin annulus  $\{1 - \delta \leq |x| \leq 1 + \delta\}$ ; in particular in [37] the 1D problem is proven to be equivalent, when  $\delta \rightarrow 0$ , to the repulsive harmonic one, for which high non-uniqueness holds and the Seidl map is not optimal (see the following section). However in [31] also a positive example is found, namely a class of measures for which the conjecture holds.

These results show that the solution to the multi marginal problem with Coulomb cost is far from being understood: in particular there is no clear condition on the marginal  $\mu$  for the Seidl conjecture to hold.

#### 4. MULTI-MARGINAL OT PROBLEM WITH REPULSIVE HARMONIC COST

This section is devoted to the study of the repulsive harmonic cost. In DFT-OT problem, we replace the electron-electron Coulomb interaction in the section 3, by a weak force which decrease with the square distance of the particles. The term *weak* comes from the fact that, in the repulsive harmonic cost the, interaction cost function has value zero when the particles overlap, instead of  $\infty$  in the Coulomb case.

More precisely, we are interested in characterize the minimizers of the following problem

$$(\mathcal{MK}_{weak}) \quad \min \left\{ \int_{(\mathbb{R}^d)^N} \sum_{i,j=1}^N -|x_j - x_i|^2 d\gamma(x_1, \dots, x_N) : \gamma \in \Pi((\mathbb{R}^d)^N, \mu_1, \dots, \mu_N) \right\}, \quad (4.1)$$

OTHarm:pb:MKweak

where  $\mu_1, \dots, \mu_N$  are absolutely continuous probability measures in  $\mathbb{R}^d$ .

From a mathematical viewpoint this cost has some advantages compared to the Coulomb one, since here we can do explicit examples. We can interpret the solutions of (4.1) as an *ansatz* for DFT problem (Hohenberg-Kohn functional (1.12)) for particles interacting under the repulsive harmonic potential. From a technical aspect, this could be an interesting toy model to approach the case of Coulomb cost. From applications, some minimizers of (4.1) seems to have no particular relevance in physics because, as we will see in the examples 4.10 and 4.11 below, certain optimals  $\gamma$  of (4.1) allows particles to overlap.

We will see that this problem has a very rich structure, that is very different from the classical 2-marginal case. First of all we notice that minimizers of this problem are also minimizers of the problem with the cost  $c(x_1, \dots, x_N) = |x_1 + \dots + x_N|^2$ ; in fact we have that

$$\int_{\mathbb{R}^{dN}} \sum_{i,j=1}^N -|x_i - x_j|^2 d\gamma = \int_{\mathbb{R}^{dN}} c d\gamma - (N-1) \sum_{i=1}^N \int_{\mathbb{R}^d} |x|^2 d\mu_i,$$

but this last additive term depends only on the marginals and not on the specific plan  $\gamma$ . The cost  $c$  is very particular since has a wide class of “trivial” optimal plans, that is the ones that are concentrated on  $x_1 + \dots + x_N = 0$ ; however the structure is very rich, see Lemma 4.3.

Concerning the existence of minimizers of (4.1), we will assume that the measures  $\mu_i$  have finite second moments; then existence follows immediately from the equality

$$\operatorname{argmin}_{\gamma \in \Pi((\mathbb{R}^d)^N, \mu_1, \dots, \mu_N)} \left\{ \int - \sum_{i=1}^N \sum_{j=i+1}^N |x_j - x_i|^2 d\gamma \right\} = \operatorname{argmin}_{\gamma \in \Pi((\mathbb{R}^d)^N, \mu_1, \dots, \mu_N)} \left\{ \int |x_1 + \dots + x_N|^2 d\gamma \right\}.$$

In particular, Corollary 2.15 holds in this case too. Notice that the fact that the repulsive harmonic cost is smooth and has linear gradient does not make the Multi-marginal Optimal Transportation problem easier compared to the Coulomb cost. In fact in this case the problem is that if we write down the optimality conditions for the potentials, in the case described in Lemma 4.3, we simply find the condition  $x_1 + \dots + x_N = \text{const}$  since in this case the potentials are linear functions.

However, we can enjoy the symmetries of this problems and build easy Monge solutions for some particular cases of (4.1), see examples 4.8, 4.10, 4.11 and 4.13 below.



Before stating the main result in the multi-marginal setting, we start analyzing the problem (4.1) in the 2-marginals case, where everything seems to work fine, just as in the square distance case.

**Proposition 4.1.** *Let  $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$  and  $c_w(x, y) = -|x - y|^2$  be the opposite of the square-distance in  $\mathbb{R}^d$ . Suppose that  $\mu$  is an absolutely continuous with respect to the Lebesgue measure in  $\mathbb{R}^d$  and  $\nu$  has no atoms. Then, there exists a unique optimal transport map  $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$  for the problem*

$$\min \left\{ \int c_w(x, y) d\gamma : \gamma \in \Pi_2(\mathbb{R}^{2d}, \mu, \nu) \right\} = \inf \left\{ \int c_w(x, T(x)) d\mu : T_{\#}\mu = \nu \right\}$$

Moreover,  $T = \nabla \phi$ , where  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$  is a concave function and there exists a unique optimal transport plan  $\bar{\gamma}$ , that is  $\bar{\gamma} = (\mathbb{I}_d \times T)_{\#}\mu$ .

*Proof.* This result is a easy consequence of the Brenier's theorem. Indeed, it is enough to verify that, taking  $C = 2 \int |x|^2 d\mu + 2 \int |x|^2 d\nu$ , we have

$$C + \inf \left\{ \int -|x - T(x)|^2 d\mu(x) : T_{\#}\mu = \nu \right\} = \inf \left\{ \int |x - G(x)|^2 d\mu(x) : G_{\#}\mu = \tilde{\nu} \right\}$$

where  $G = -T$  and  $\tilde{\nu} = (-Id)_{\#}\nu$ . Then by Brenier's theorem there exists an unique optimal map  $G$  which can be written as  $G(x) = \nabla \psi(x)$ , and  $\psi : \mathbb{R}^d \rightarrow \mathbb{R}$  is a convex map. In other words,  $T$  is a gradient of a concave function.  $\square$

Notice that if we suppose  $\mu = f(x)dx$  and  $\nu = g(x)dx$  are probability measures with densities  $f, g$  concentrated, respectively, in convex sets  $\Omega_f, \Omega_g$  and assume there exists a constant  $\lambda > 0$  such that  $\lambda \leq f, g \leq 1/\lambda$ . Then,  $T$  is a  $C^{1,\alpha}$  function inside  $\Omega_f$  [17, 18].

Exam2Marginals

*Example 4.2* (2 marginals case, uniform measure in the  $d$ -dimensional cube). Suppose that  $\mu = \nu = \mathcal{L}|_{[0,1]^d}$ . In this case, we can verify easily that the optimal map  $T : [0, 1]^d \rightarrow [0, 1]^d$  is given by the anti-monotone map  $T(x) = (1, \dots, 1) - x$ .

Surprisingly, the next theorem says that we can not always expect Caffarelli's regularity for the optimal transport maps for the repulsive harmonic cost with finitely many marginals  $N > 2$  even when the support of the measures has convex interior (see corollary 4.7 below). Before stating the theorem we will prove a lemma that characterizes many optimal plans:

Harm:stuplemma

**Lemma 4.3.** *Let  $\{\mu_i\}_{i=1}^N$  be probability measures on  $\mathbb{R}^d$  and  $h : \mathbb{R}^d \rightarrow \mathbb{R}$  be a strictly convex function and suppose  $c : ([0, 1]^d)^N \rightarrow \mathbb{R}$  a cost function of the form  $c(x_1, \dots, x_N) = h(|x_1 + \dots + x_N|)$ . Then, if there exists a plan  $\gamma \in \Gamma(\mu_1, \dots, \mu_N)$  concentrated on some hyperplane of the form  $x_1 + \dots + x_N = k$ , this plan is optimal for the multi marginal problem with cost  $c$  and  $\text{supp}(\tilde{\gamma}) \subset \{x_1 + \dots + x_N = k\}$  is a necessary and sufficient condition for  $\tilde{\gamma}$  to be optimal. In this case we will say that  $\gamma$  is a flat optimal plan and  $\{\mu_i\}_{i=1}^N$  is a flat  $N$ -tuple of measures.*

*Proof.* First of all we show that  $k$  is fixed by the marginals  $\mu_i$ , and it is in fact the sum of the barycenters of these measures. Let  $c_i = \int x d\mu_i$ ; then let us suppose that there exists  $\gamma \in \Gamma(\mu_1, \dots, \mu_N)$  that is concentrated on  $\{x_1 + \dots + x_N = k\}$ . Then, using that the  $i$ -th marginal of  $\gamma$  is  $\mu_i$ , we can compute

$$k = \int (x_1 + \dots + x_N) d\gamma = \sum_{i=1}^N \int x_i d\gamma = \sum_{i=1}^N \int x d\mu_i = \sum_{i=1}^N c_i.$$

In particular we notice also that for every admissible plan  $\tilde{\gamma}$  we have  $\int (x_1 + \dots + x_N) d\tilde{\gamma} = k$  and so by Jensen inequality we have

$$\int h(x_1 + \dots + x_N) d\tilde{\gamma} \geq h\left(\int (x_1 + \dots + x_N) d\tilde{\gamma}\right) = h(k) = \int h(x_1 + \dots + x_N) d\gamma.$$

This proves that  $\gamma$  is an optimal plan. Thanks to the strict convexity of  $h$ , this shows also that if  $\tilde{\gamma}$  is optimal then  $\tilde{\gamma}$ -a.e. we should have  $x_1 + \dots + x_N = k$ .  $\square$

This reveals a very large class of minimizers in some cases, as we will see later. However not every  $N$ -tuple of measures is flat: it is clear that we can have marginals such that there is no plan with such property:

*Example 4.4.* Let  $N = 3$  and  $\mu_i = \mu$  for every  $i = 1, 2, 3$  with  $\mu = \nu_1 + \nu_2$  with  $\nu_1 = (-Id)_\# \nu_2$  and  $\nu_1$  concentrated on  $[2, 3]$ . Now it is clear that the barycenter of  $\mu$  is 0 but for every 3 points  $x_1, x_2, x_3$  in the support of  $\mu$  we cannot have  $x_1 + x_2 + x_3 = 0$ : two of them have the same sign, let's say  $x_1$  and  $x_2$ , but then we have  $|x_1 + x_2| \geq 4 > 3 \geq |x_3|$ , which contradicts  $x_1 + x_2 = -x_3$ . So it can't exist an admissible plan concentrated on  $H_0 = \{x_1 + x_2 + x_3 = 0\}$ ; in fact we showed that for every  $\gamma \in \Pi(\mu)$  we have  $\text{supp}(\gamma) \cap H_0 = \emptyset$ .

*Remark 4.5.* In the case  $N = 2$  the condition for which there exists an admissible plan  $\gamma$  concentrated on an hyperplane of the form  $x_1 + x_2 = k$  implies that the two measures  $\mu$  and  $\tilde{\nu} = (-Id)_\# \nu$  are equal up to translation. This condition is in fact very restrictive. However when we think at the DFT problem and in particular we assume that  $\mu = \nu$  then the said condition amounts to have that  $\mu$  is centrally symmetric about its barycenter and in the context of density of 2 electrons around a nucleus this seems a fairly natural condition.

thm

**Theorem 4.6.** Let  $\mu_i = \mu = \mathcal{L}^d|_{[0,1]^d}, \forall i = 1, \dots, N$  be the uniform measure on the  $d$ -dimensional cube  $[0, 1]^d \subset \mathbb{R}^d$ ,  $h : \mathbb{R}^d \rightarrow \mathbb{R}$  a convex function and suppose  $c : ([0, 1]^d)^N \rightarrow \mathbb{R}$  a cost function such that  $c(x_1, \dots, x_N) = h(x_1 + \dots + x_N)$ . Then, there exists a transport map  $T : [0, 1]^d \rightarrow [0, 1]^d$  such that  $T^N(x) = x$  and

$$\min \left\{ \int c d\gamma : \gamma \in \Pi_N(\mu) \right\} = \min \left\{ \int c(x, T(x), \dots, T^{N-1}(x)) d\mu : \begin{array}{l} T_\# \mu = \mu, \\ T^N = I \end{array} \right\}. \quad (4.2)$$

Moreover,  $T$  is not differentiable almost everywhere.

*Proof.* We express every  $z \in [0, 1]$  by its base- $N$  system,  $z = \sum_{k=1}^{\infty} \frac{a_k}{N^k}$  with  $a_k \in \{0, 1, \dots, N-1\}$ . Consider the map by  $S(z) = \sum_{k=1}^{\infty} \frac{\mathcal{S}(a_k)}{N^k}$ , where  $\mathcal{S}$  is defined by the permutation of  $N$  symbols, i.e.  $\mathcal{S}(i) = i + 1, \forall i = 1, \dots, N-1$  and  $\mathcal{S}(N) = 1$ . A straight forward computation shows that

$$z + \sum_{i=1}^N S^i(z) = \frac{N}{2} \quad (4.3) \quad \text{eqn:bary}$$

Let  $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$  be a map defined by  $T(x) = T(z_1, \dots, z_d) = (S(z_1), \dots, S(z_d))$  and denote by  $T^j(x) = (S^{(j)}(z_1), \dots, S^{(j)}(z_d))$ ,  $j = 1, \dots, N-1$ . We will first show that  $S$  is a measure-preserving map. In fact, we can show that there exists functions  $S_k : [0, 1] \rightarrow [0, 1]$  defined recursively by

$$S_0(x) = x, \quad \text{and} \quad S_{k+1}(x) = T_{k+1}(S_k(x))$$

where  $T_k$  acts only of the  $k$ -th digit:  $T_1(x) = x - (N-1) \cdot N^{-k}$  if  $x \in C_k$  and  $T_k(x) = x + N^{-k}$  if  $x \in [0, 1] \setminus C_k$ . The sets  $C_k$  are defined by

$$C_k = \bigcup_{j=1}^{N^{k-1}} \left( \frac{N(j-1) + (N-1)}{N^k}, \frac{j}{N^{k-1}} \right].$$

Moreover, it is easy to see that

$$(T_k)_\# \mathcal{L}_{[0,1]} = \mathcal{L}_{[0,1]} \quad \forall k \in \mathbb{N}, \quad \text{and} \quad S_k \rightarrow S \text{ uniformly.}$$

Hence

$$\int f(x) S_\# \mathcal{L}^d = \int f(x) dx, \quad \forall f \in C_0([0, 1]^d).$$

Now, it remains to show that  $T$  is optimal. But this is true thanks to the fact that (4.3) implies that the plan induced by  $T$  satisfies the hypothesis of Lemma 4.3.  $\square$

OTHARM:maincor

**Corollary 4.7.** *Let  $\mu = \mathcal{L}^d|_{[0,1]^d}$  be the uniform measure on the  $d$ -dimensional cube  $[0, 1]^d$  in  $\mathbb{R}^d$  and suppose  $c : ([0, 1]^d)^N \rightarrow \mathbb{R}$  the  $N$ -dimensional repulsive harmonic cost.*

$$c(x_1, \dots, x_N) = - \sum_{i=1}^N \sum_{j=i+1}^N |x_j - x_i|^2, \quad (x_1, \dots, x_N) \in ([0, 1]^d)^N$$

*Then, there exists an optimal cyclical transport map  $T : [0, 1]^d \rightarrow [0, 1]^d$ : in particular*

$$\min \left\{ \int c d\gamma : \gamma \in \Pi_N(\mu) \right\} = \min \left\{ \int c(x, T(x), \dots, T^{N-1}(x)) d\mu : \begin{array}{l} T_\# \mu = \mu, \\ T^N = I \end{array} \right\}. \quad (4.4)$$

*Moreover,  $T$  is not differentiable almost everywhere.*

*Proof.* As we already observed, the problem with the cost  $c$  is equivalent to the problem with the cost  $|x_1 + \dots + x_N|^2$  and the result follows from the Theorem 4.6.  $\square$

We remark the construction of  $T$  in the proof of the Theorem 4.6 also works if  $N = 2$  and  $T$  is exactly the optimal transport map described in the example (4.2).

The unexpected aspect of Corollary 4.7 is the existence - for  $N \geq 3$  - of an optimal transport map which is not differentiable almost everywhere. It turns out that this optimal map  $T$  could be not unique if  $d > 1$  and  $N \geq 3$  and, in that case, we can construct explicitly a regular optimal map. This kind of richness appears already in the case  $d = 1$  with the Coulomb cost; however in that case we have uniqueness if we restrict ourselves to the symmetric optimal plan. Here this is not the case as it is easy to see that modifying the action  $\mathcal{S}$  on the digits (the important thing is that when we see  $\mathcal{S}$  as a permutation, it is a cycle), we obtain another map, and the symmetrized plan is not equal to the one generated by the map described before.

In the following, we are going to present some concrete examples where we have other explicit solutions for such kind of optimal maps for  $(\mathcal{MK}_{weak})$  in (4.1). For these example the goal is to show that there can be smooth optimal maps but they don't necessarily satisfy the "group rule", that is, we can find maps  $T_2, \dots, T_N$  such that  $(Id, T_2, \dots, T_N)_\# \mu_1$  is an optimal plan but  $T_2 \circ T_2 \neq T_i$  for any index  $i$ .

*Example 4.8* (3 particles, asymmetric). Consider the case when three particles are distributed in  $\mathbb{R}^3$  as Gaussians,  $\mu_1 = \mu_2 = \frac{1}{(2\pi)^{3/2}} \exp(-\frac{1}{2}(x_1^2 + x_2^2 + x_3^2))$  and  $\mu_3 = \frac{1}{2\pi^{3/2}} \exp(-(x_1^2 + x_2^2 + x_3^2))$ .

OTHarm:ex:3p

In this case, we can verify that the couple  $(T, S)$  of maps  $T, S : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $T(x) = x$  a.e. and  $S(x) = -2x$  a.e. is admissible and it is optimal since  $x + T(x) + S(x) = 0$ .

als

*Example 4.9* (2N particles on  $S^1$ , ). Suppose  $\mu_1, \dots, \mu_{2N}$  uniform probability measures on the circle  $S^1$ . The rotation map  $R_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  with angle  $\theta = \pi/N$  is an optimal transport map. Also, the maps  $R_{k\theta}$ ,  $k = 2, \dots, N$ , are optimal transport maps for the repulsive harmonic cost.

als

*Example 4.10* (2N particles, breathing map). Suppose  $\mu_1, \dots, \mu_{2N}$  uniform probability measures on  $S^2 \subset \mathbb{R}^3$ . Consider the vector  $v$ , the  $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  the antipodal map  $A(x) = -x$  and  $R_\theta^v : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  the rotation of angle  $\theta = \pi/N$  and direction  $v$ . Then,  $T = A \circ R_\theta^v$  is a cyclic optimal transport map.

The optimal solution  $\gamma = (x, T, T^2, \dots, T^{N-1})_\# \mu_1$  is called “breathing” solution [94]: the coupling  $\gamma$  represent the configuration where the  $2N$  electrons are always at the same distance from the center, opposite to each other in the equilibrium configuration. Notice that, the map  $G = T^2$  is also an cyclical optimal transportation map. Moreover,  $T$  is a  $C^\infty$  function and its a gradient of a convex-concave function  $\phi(x_1, x_2, \dots, x_{2N-1}, x_{2N}) = \frac{1}{2}(x_1^2 + x_3^2 + \dots + x_{2N-1}^2) - \frac{1}{2}(x_2^2 + x_4^2 + \dots + x_{2N}^2)$ .

2Np

*Example 4.11* (2N particles in  $\mathbb{R}^d$ , symmetric  $\rho$ ). In this case, there exists maps  $T(x) = x$  a.e. and  $S(x) = -x$  a.e. such that  $c(T(x), S(x), \dots, T(x), S(x)) = 0$ . We notice that in this case we have  $S^{(2)}(x) = x$  and so in particular this solution has the cyclic structure

$$(T(x), S(x), \dots, T(x), S(x)) = (x, S(x), S^{(2)}(x), \dots, S^{(2N-1)}(x)).$$

1an

*Example 4.12* (Optimal Maps which doesn't satisfy a group law). Let us consider the case  $d = 1$  and  $N = 3$  where the measures are  $\mu_1 = \mu_2 = \mu_3 = \frac{1}{2}\mathcal{L}|_{[-1,1]}$ .

Then we define the maps

$$T(x) = \begin{cases} x+1 & \text{if } x \leq 0 \\ x-1 & \text{if } x > 0, \end{cases} \quad S(x) = \begin{cases} -1-2x & \text{if } x \leq 0 \\ 1-2x & \text{if } x > 0. \end{cases}$$

We have that  $T_\# \mu = \mu$  and  $S_\# \mu = \mu$ , and moreover  $x + T(x) + S(x) = 0$ ; in particular  $(Id, S, T)_\# \mu$  is a flat optimal plan and so the thesis.

1an

*Example 4.13* (Optimal Transport diffuse plan). Let us consider the same problem as in Example 4.12. Now we consider a general symmetric plan  $\gamma = \frac{1}{2}\mathcal{H}^2|_H f(\max\{|x|, |y|, |z|\})$ , where  $H$  is defined as  $H = \{x + y + z = 0\} \cap \{|x| \leq 1, |y| \leq 1, |z| \leq 1\}$ . This is a symmetric flat optimal plan; now we compute the marginals. Since it is clear that  $\gamma$  is invariant under  $x \mapsto -x$ , it is sufficient to consider the marginal on the set  $x > 0$ . But then we can make the computation

$$\begin{aligned} \int_{x \geq 0} \phi(x) d\gamma(x, y, z) &= \sqrt{3} \int_{|x|, |y|, |x+y| \leq 1, x \geq 0} \phi(x) f(\max\{|x|, |y|, |x+y|\}) dx dy \\ &= \sqrt{3} \int_0^1 \int_0^{1-x} \phi(x) f(x+y) dy dx + \sqrt{3} \int_0^1 \int_{-x}^0 \phi(x) f(x) dy dx \\ &\quad + \sqrt{3} \int_0^1 \int_{-1}^{-x} \phi(x) f(|y|) dy dx \\ &= \sqrt{3} \int_0^1 \phi(x) \left( x f(x) + 2 \int_x^1 f(t) dt \right) dx. \end{aligned}$$

In particular the choice  $f(x) = \frac{\sqrt{3}}{6}x$  gives the marginals equal to  $\frac{1}{2}\mathcal{L}|_{[-1,1]}$ . We notice also that any even density  $\rho = h(|x|)$  for some decreasing function  $h : [0, 1] \rightarrow [0, \infty)$  can be represented in this way: in fact it is sufficient to choose  $f(x) = \frac{1}{\sqrt{3}}\left(\frac{h(x)}{x} - 2x \int_x^1 \frac{h(t)}{t^3} dt\right)$ .

*Example 4.14* (A Counterexample on the uniqueness of Multimarginal Optimal Maps for  $N > 3$ ) Passcounterexample  
As mentioned in the theorem 2.16 and in [42], it was already understood by Pass that in these high dimensional cases, the solution of repulsive costs may also be non unique, differently on the two marginals case. On a higher dimensional surface there can be enough wiggle room to construct more than one measure with common marginals, as shows in the examples 4.8, 4.10, 4.11, 4.12, 4.13. In most of the cases, the non-uniqueness seems to be given by the symmetries of the problem, but in Example 4.13 and Corollary 4.7 this is not the case, as we exploit the fact that the dimension of the set  $c(x + y + z) - \phi(x) - \phi(y) - \phi(z) = 0$ , is greater then the minimal one.

Finally, the last proposition of this section says that when  $d = 1$  and odd  $N$  we have no hope in general to find piecewise regular cyclic optimal transport maps.

**Proposition 4.15.** *Let  $\mu = \mathcal{L}|_{[0,1]}$  and  $N \geq 3$  be an odd number. Then, the infimum*

$$\inf \left\{ \int c(x, T(x), T^2(x), \dots, T^{N-1}(x)) d\mu : \begin{array}{l} T_{\#}\mu = \mu, \\ T^N = I \end{array} \right\}.$$

*is not achieved by a map  $T$  which is differentiable almost everywhere. In particular, the non differentiable transport map  $T$  given by the theorem 4.6 is a Monge minimizer for (4.1)*

*Proof.* First of all we notice that if  $T$  is differentiable almost everywhere then also  $T^{(2)}$  has the same property, thanks to the fact that  $T_{\#}\mu = \mu$ . In particular, since the Lusin property holds true for  $T^{(i)}$  for every  $i = 1, \dots, N$  and  $T_{\#}\mu = \mu$ , the change of variable formula holds and in particular we have that  $(T^{(i)})'(x) = \pm 1$  for almost every  $x$  (notice also that  $T$  is bijective almost everywhere since  $T^{(N)}(x) = x$ ).

Since  $\mu$  is  $N$ -flat we have that the condition on  $T$  in order to be an optimal cyclical map is  $x + T(x) + T^{(2)}(x) + \dots + T^{(N-1)}(x) = N/2$ ; now we can differentiate this identity and so we will get

$$1 + T'(x) + (T^{(2)})'(x) + \dots + (T^{(N-1)})'(x) = 0 \quad \text{for a.e. } x.$$

But this is absurd since on the left hand side we have an odd number of  $\pm 1$  and their sum will be an odd number. □

In conclusion, also in the case of the repulsive harmonic cost, the picture is far from being clear: an interesting structure appears when  $\{\mu_i\}_{i=1}^N$  is a flat  $N$ -tuple of measures but we still can't characterize this property. Moreover, in the flat case in which  $\mu_i = \mu$ , for example when  $\mu = \frac{1}{2}\mathcal{L}|_{[-1,1]}$ , we have both diffuse optimal plan and a cyclical optimal map.

An interesting open problem is whether if for any  $N$ -flat measure  $\mu$ , say absolutely continuous with respect to the Lebesgue measure, we have a cyclical optimal map and a diffuse plan.

## 5. MULTI-MARGINAL OT FOR THE DETERMINANT

We are going to give a short overview of the main results in [21], where Carlier and Nazaret consider the following optimal transport problems for the determinant:

$$(\mathcal{MK}_{Det}) \quad \sup \left\{ \int_{(\mathbb{R}^d)^d} \det(x_1, \dots, x_d) d\gamma(x_1, \dots, x_d) : \gamma \in \Pi((\mathbb{R}^d)^d, \mu_1, \dots, \mu_d) \right\}, \quad (5.1) \quad \text{Det:pb:MKDet}$$

and

$$(\mathcal{MK}_{|Det|}) \quad \sup \left\{ \int_{(\mathbb{R}^d)^d} |\det(x_1, \dots, x_d)| d\gamma(x_1, \dots, x_d) : \gamma \in \Pi((\mathbb{R}^d)^d, \mu_1, \dots, \mu_d) \right\}, \quad (5.2) \quad \text{Det:pb:MKDetm}$$

where  $\mu_1, \dots, \mu_d$  are the uniform probability measures in  $\mathbb{R}^d$ . In addition, in order to guarantee existence of a solution, we assume that exist  $p_1, \dots, p_d \in [1, \infty[$  such that

$$\sum_{i=1}^d \frac{1}{p_i} = 1, \quad \text{and} \quad \sum_{i=1}^d \frac{|x_i|_i^p}{p_i} < +\infty.$$

Notice here the problem (2.1) makes sense only when  $N = d$ . In the following, we will focus in discuss the problem (5.1) and exhibit explicit minimizers  $\gamma$  in the radial case. Clearly, the difference between (5.1) and (5.2) is that the second one admits positively and negatively oriented basis of vectors, while the first one “chooses” only the positive ones. Moreover, if we assume that among the marginal  $\mu_1, \dots, \mu_d$  there exists two symmetric probability measures  $\mu_i, \mu_j, i \neq j$ , *i.e.*  $\mu_i = (-\text{Id})_{\#} \mu_i$  and  $\mu_j = (-\text{Id})_{\#} \mu_j$ , then any solution  $\gamma$  of (5.1) satisfies  $\det(x_1, \dots, x_d) \geq 0$   $\gamma$ -almost everywhere and so solves also (5.2) (Proposition 6, [21]).

Similarly to the Gangbo-Świąch cost [47], the Monge-Kantorovich problem for the determinant (5.1) can be seen as a natural extension of classical optimal transport problem with two marginals and so, it is equivalent to the 2-marginals repulsive harmonic cost (4.1).

Indeed, we can write in the two marginals case,  $\det(x_1, x_2) = \langle x_1, R x_2 \rangle$ , where  $R : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is the rotation of angle  $-\pi/2$ . Hence, since  $\mu_1$  and  $\mu_2$  has finite second moments, up to a change of variable  $\tilde{x}_2 = R x_2$ , the problem  $(\mathcal{MK}_{Det})$  in (5.1) is equivalent to the classical Brenier's optimal transportation problem, as shows the following computation.

$$\begin{aligned} \operatorname{argmax}_{\gamma \in \Pi(\mu_1, \mu_2)} \int_{\mathbb{R}^2} \det(x_1, x_2) d\gamma(x_1, x_2), &= \operatorname{argmax}_{\gamma \in \Pi(\mu_1, \mu_2)} \int_{\mathbb{R}^2} \langle x_1, \tilde{x}_2 \rangle d\gamma(x_1, x_2), \\ &= \operatorname{argmax}_{\tilde{\gamma} \in \Pi(\mu_1, \tilde{\mu}_2)} \int_{\mathbb{R}^2} \langle x_1, x_2 \rangle d\tilde{\gamma} - C, \\ &= \operatorname{argmin}_{\gamma \in \Pi(\mu_1, \tilde{\mu}_2)} \int_{\mathbb{R}^2} \frac{|x_1 - x_2|^2}{2} d\gamma(x_1, x_2) \end{aligned}$$

where  $C = 1/2(\int |x_1|^2 d\mu_1 + \int |x_2|^2 d\mu_2)$  and  $\tilde{\mu}_2 = R_{\#} \mu_2$ .

In the sequel, we are going to construct maximizers for (5.1), thanks to some properties of the Kantorovich potentials in the dual problem associated to (5.1) (see theorem 5.1 below),

$$(\mathcal{K}_N^{Det}) \quad \inf \left\{ \int_{\mathbb{R}^d} \sum_{i=1}^d u_i(x_i) d\mu_i(x_i) : \sum_{i=1}^N u_i(x_i) \leq \det(x_1, \dots, x_d) \right\}. \quad (5.3) \quad \text{pb:KNDet}$$

In [21] it is provided an useful characterization of optimal transport plans through the potentials  $u_i$ , given by the following theorem 5.1 below. In addition, by means of a standard convexification trick we obtain regularity results on the Kantorovich potentials.

Det:carcgamma

**Theorem 5.1.** *A coupling  $\gamma \in \Pi((\mathbb{R}^d)^d, \mu_1, \dots, \mu_d)$  is optimal in (5.1) if and only if there exists lower semi-continuous convex functions  $u_i : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{\infty\}$  such that for all  $i \in \{1, \dots, d\}$ ,*

$$\begin{aligned} \sum_{i=1}^d u_i(x_i) &\leq \sum_{i=1}^d u_i^*((-1)^{i+1} \bigwedge_{i \neq j} x_j), \quad \text{on } (\mathbb{R}^d)^d; \\ \sum_{i=1}^d u_i(x_i) &= \sum_{i=1}^d u_i^*((-1)^{i+1} \bigwedge_{i \neq j} x_j), \quad \gamma - \text{almost everywhere}; \\ (-1)^{i+1} \bigwedge_{i \neq j} x_j &\in \partial u_i(x_i), \quad \gamma - \text{almost everywhere}. \end{aligned}$$

where  $\bigwedge_{i=1}^d x_j$  denotes the wedge product and, for every  $i$ ,  $u_i^*$  is the convex dual of the Kantorovich potential  $u_i$ .

Now, the main idea is to use the geometrical constraints on the Kantorovich potentials  $u_i$  (5.3), given by the theorem (5.1), in order to construct an explicit solutions.

We illustrate the theorem (5.1) and explain how to construct a particular optimal  $\gamma$  for (5.1) by an example in the three marginals case in the unit ball in  $\mathbb{R}^3$ . Let  $\mu_i = \rho_i \mathcal{L}^3$ ,  $i = 1, 2, 3$  radially symmetric probability measures on the 3-dimensional ball  $B$ .

In this particular situation, the optimizers of (5.1) and (5.2) has a natural geometric interpretation: what is the best way to place three random vectors  $x, y, z$ , distributed by probability measures  $\mu_1, \mu_2, \mu_3$  on the sphere, such that the simplex generated by those three vectors  $(0, x, y, z)$  has maximum average volume?

Suppose  $\gamma \in \Pi(B, \mu_1, \mu_2, \mu_3)$  optimal in (5.1) when  $d = 3$ . From optimality of  $\gamma$ , we have

$$u_1(x) + u_2(y) + u_3(z) = \det(x, y, z), \quad \gamma - \text{almost everywhere}.$$

Applying the theorem (5.1), we get

$$\begin{cases} u_2(y) + u_3(z) = u_1^*(y \wedge z) \\ u_1(x) + u_3(z) = u_2^*(-x \wedge z) \\ u_1(x) + u_2(y) = u_3^*(x \wedge y) \end{cases}, \quad \gamma - \text{almost everywhere},$$

and,

$$\begin{cases} \nabla u_1(x) = y \wedge z \\ \nabla u_2(y) = -x \wedge z \\ \nabla u_3(z) = x \wedge y \end{cases}, \quad \gamma - \text{almost everywhere}. \quad (5.4)$$

sys:perp

Again, from the optimality of  $\gamma$ , given a vector  $x$  on in the ball, the conditional probability of  $y$  given  $x$  is supported in a “meridian”  $M(x)$

$$M(x) = \{y \in S^2 : \langle \nabla u_1(x), y \rangle = 0\},$$

where  $S^2$  is the 2-sphere. Finally, assuming that  $\langle x, \nabla u_1(x) \rangle \neq 0$ , the conditional probability of  $z$  given the pair  $(x, y)$  is simply given by a delta function on  $z$

$$z = \frac{\nabla u_1(x) \wedge \nabla u_2(y)}{\langle x, \nabla u_1(x) \rangle};$$

In particular, we have

$$\langle x, \nabla u_1(x) \rangle + \langle y, \nabla u_2(y) \rangle + \langle z, \nabla u_3(z) \rangle = \det(x, y, z) = \det(\nabla u_1(x), \nabla u_2(y), \nabla u_3(z)).$$

*Example 5.2* (An explicit solution in the ball  $B \subset \mathbb{R}^3$ ). Suppose  $\mu_i = \mathcal{L}_B^3$ ,  $i = 1, 2, 3$ , the 3-dimensional Lebesgue measure in the ball  $B \subset \mathbb{R}^3$ . The following coupling  $\gamma^*$

$$\int_{B^3} f d\gamma^* = \frac{1}{\mathcal{L}^3(B)} \int_B \left( \int_{M(x)} f(x, |x|y, x \wedge y) \frac{d\mathcal{H}^1(y)}{2\pi} \right) dx, \quad \forall f \in C(B^3, \mathbb{R}). \quad (5.5)$$

Det: examplegamma

is an optimizer for (5.1) with  $d = 3$ . Indeed, from we can show explicit potentials  $u_1^*(x) = u_2^*(x) = u_3^*(x) = |x|^3/3$ ; clearly we have

$$\det(x, y, z) \leq |x||y||z| \leq |x|^3/3 + |y|^3/3 + |z|^3/3, \quad \forall (x, y, z) \in B,$$

with equality when  $|x| = |y| = |z|$  and  $x, y, z$  are orthogonal. Since  $\gamma^*$  is concentrated on this kind of triples of vector we have the optimality. Finally, by a suitable change of variable, it is easy to see that  $\gamma^* \in \Pi_3(\mathcal{L}_B^3)$ .

*Some comments on the radially symmetric  $d$ -marginals case:* In [21], for  $d$  radially symmetric probability measures, the authors exhibit explicit optimal couplings  $\gamma^*$ . In their proof, two aspects were crucial: the first one is remark that if  $\{\mu_i\}_{i=1}^d$  are radially symmetric measures in  $\mathbb{R}^d$ , then the optimal Kantorovich potentials  $u_i(x_i) = u_i(|x_i|)$  are also radially symmetric; in particular the system (5.4) for general  $d$  implies that the support of  $\gamma^*$  is contained in the set of orthogonal basis. The second observation is to notice that in the support of  $\gamma^*$  we have  $H_i(|x_1|) = |x_i|$ , where  $H_i$  is the unique monotone increasing map such that  $(H_i^*)_{\#}\mu_1 = \mu_i$ , where  $H_i^*(x) = \frac{x}{|x|}H_i(|x|)$ . This is done analyzing the correspondent radial problem (with cost  $c(r_1, \dots, r_d) = r_1 \cdots r_d$ ), using the optimality condition  $\phi'_i(r_i) = \partial_i c$  and the fact that in this case  $r_i \phi'_i(r_i) = r_1 \phi'_1(r_1) = c \geq 0$ ; then, exploiting the convexity of  $\phi_i$  we get  $r_i = H_i(r_1)$  for some increasing function  $H_i$ , that is uniquely determined.

*Existence of Monge-type solutions:* In the 3-marginals case in the unit ball, by construction of the the coupling  $\gamma^*$  in (5.5) or, more generally, the optimal coupling in the  $d$ -marginal case (see theorem 4 in [21]), we can see that their support are not concentrated in the graph of cyclic maps  $T, T^2, \dots, T^{d-1}$  or simply on the graph of maps  $T_1, \dots, T_{d-1}$  as we could expect from corollary (2.15). In other words,  $\gamma^*$  in (5.5) is not Monge-type solution.

The existence of Monge type solutions for the determinant cost is still an open problem for odd number of marginals. From the geometric conditions we discussed above, in the case in which  $\mu_i = \mu$  a radial measure, if Monge solutions exists then, for every  $x \in \mathbb{R}^d$ ,  $(x, T_1(x), \dots, T_{d-1}(x))$  should be a orthogonal basis, and  $|T_i(x)| = |x|$ ,  $i = 1, \dots, d-1$ .

For the interesting even dimensional case, we can observe a similar phenomena remarked in the repulsive harmonic costs and build easy Monge solutions for (5.1).

*Example 5.3. The even dimensional phenomena:* As in the repulsive harmonic cost, it is easy to construct Monge minimizers for the determinant cost for even number of marginals  $\geq 4$ . For instance, suppose  $c(x_1, x_2, x_3, x_4) = \det(x_1, x_2, x_3, x_4)$  and define maps  $T_1, T_2, T_3 : B \rightarrow \mathbb{R}^4$  by

$$T_1(x) = \begin{pmatrix} -x_2 \\ x_1 \\ -x_4 \\ x_3 \end{pmatrix}, \quad T_2(x) = \begin{pmatrix} -x_3 \\ x_4 \\ x_1 \\ -x_2 \end{pmatrix}, \quad T_3(x) = \begin{pmatrix} -x_4 \\ -x_3 \\ x_2 \\ x_1 \end{pmatrix}, \quad x = (x_1, x_2, x_3, x_4) \in B.$$

We can see that  $\gamma_T = (Id, T_1, T_2, T_3)_{\#}\gamma$  is a Monge-type optimal transport plan for (5.1) and (5.2).



## 6. NUMERICS

numerics

Numerics for the multi-marginal problems have so far not been extensively developed. Discretizing the multi-marginal problem leads to a linear program where the number of constraints grows exponentially in the number of marginals. In [22] Carlier, Oberman and Oudet studied the matching of teams problem and they were able to reformulate the problem as a linear program whose number of constraints grows only linearly in the number of marginals. More recently a numerical method based on an entropic regularization has been developed and it has been applied to various optimal transport problem in [11, 12, 35]. Let us mention that the Coulomb cost has been treated numerically in various paper as in [73, 52], where the solutions are based on the analytical form of transport maps given in [52], in [74], where parameterized functional forms of the Kantorovich potential are used, in [25, 26], where a linear programming approach has been developed to solve the 1-dimensional problem, and, as already mentioned, in [12]. In this section we focus on the regularized method proposed in [11, 12, 35] and we, finally, present some numerical experiments for the costs studied above.

subIPFP

**6.1. The regularized problem and the Iterative Proportional Fitting Procedure.** Let us consider the problem

$$(\mathcal{MK}) \quad \inf \left\{ \int_{\mathbb{R}^{dN}} c(x_1, \dots, x_N) \gamma(x_1, \dots, x_N) dx_1, \dots, dx_N : \gamma \in \Pi(\mathbb{R}^{dN}, \mu_1, \dots, \mu_N) \right\}, \quad (6.1)$$

pb:MKN2

where  $N$  is the number of marginals  $\mu_i$  which are probability distributions over  $\mathbb{R}^d$ ,  $c(x_1, \dots, x_N)$  is the cost function,  $\gamma$  the coupling, is the probability distribution over  $\mathbb{R}^{dN}$  and  $\Pi := \bigcap_{i=1}^N \mathcal{C}_i$  with  $\mathcal{C}_i := \{\gamma \in \mathcal{P}(\mathbb{R}^{dN}) : e_{i\#} \gamma = \mu_i\}$ .

*Remark 6.1.* From now on the marginals  $\mu_i$  and the coupling  $\gamma$  are densities and when the optimal coupling  $\gamma$  is induced by maps  $T_i$ , we write it as  $\gamma = \mu_1(x_1) \delta(x_2 - T_2(x_1)) \cdots \delta(x_N - T_N(x_1))$ .

In order to discretize (6.1), we use a discretisation with  $M_d$  points of the support of the  $k$ th marginal as  $\{x_{j_k}\}_{j_k=1, \dots, M_d}$ . If the densities  $\mu_k$  are approximated by  $\mu_k \approx \sum_{j_k} \mu_{j_k} \delta_{x_{j_k}}$ , we get

$$\min_{\gamma \in \mathcal{C}} \sum_{j_1, \dots, j_N} c_{j_1, \dots, j_N} \gamma_{j_1, \dots, j_N}, \quad (6.2)$$

pb:MKdiscrete

where  $c_{j_1, \dots, j_N} = c(x_{j_1}, \dots, x_{j_N})$  and the coupling support for each coordinate is restricted to the points  $\{x_{j_k}\}_{k=1, \dots, M_d}$  thus becoming a  $(M_d)^N$  matrix again denoted  $\gamma$  with elements  $\gamma_{j_1, \dots, j_N}$ . The marginal constraints  $\mathcal{C}_i$  becomes

$$\mathcal{C}_i := \left\{ \gamma \in \mathbb{R}_+^{(M_d)^N} : \sum_{j_1, \dots, j_{i-1}, j_{i+1}, \dots, j_N} \gamma_{j_1, \dots, j_N} = \mu_{j_i}, \forall j_i = 1, \dots, M_d \right\}. \quad (6.3)$$

constraint

As in the continuous framework the problem (6.1) admits a dual formulation

$$\begin{aligned} \max_{u_{j_k}} & \sum_{k=1}^N \sum_{j=1}^{M_d} u_{j_k} \mu_{j_k} \\ \text{s.t.} & \sum_{k=1}^N u_{j_k} \leq c_{j_1, \dots, j_N} \quad \forall j_k = 1, \dots, M_d, \end{aligned} \quad (6.4)$$

DualDiscret

where  $u_{j_k} = u_k(x_{j_k})$  is the  $k$ th Kantorovich potential. One can notice that the primal (6.1) has  $(M_d)^N$  unknown and  $M_d \times N$  linear constraints and the dual problem (6.4) has  $M_d \times N$

unknown, but  $(M_d)^N$  constraints. This actually makes the problems computationally unsolvable with standard linear programming methods even for small cases.

*Remark 6.2.* We underline that in many applications we have presented (as in DFT) the marginals  $\mu_i$  are all equals ( $\mu_i = \mu, \forall i \in \{1, \dots, N\}$ ). Thus the dual problem can be re-written in a more convenient (but still computationally unfeasible for many marginals) way

$$\begin{aligned} \max_{u_j} \quad & \sum_{j=1}^M N u_{j_k} \mu \\ \text{s.t.} \quad & \sum_{k=1}^N u_{j_k} \leq c_{j_1 \dots j_N} \quad \forall j_k = 1, \dots, M_d, \end{aligned} \tag{6.5} \quad \boxed{\text{DualDiscreteBis}}$$

where  $u_j = u(x_{j_k})$ . Now the dual problem has  $M_d$  unknown, but  $(M_d)^N$  linear constraints.

A different approach consists in computing the problem (6.1) regularized by the entropy of the joint coupling. This regularization dates to E. Schrödinger [92] and, as mentioned above, it has been recently introduced in many applications involving optimal transport [11, 12, 35, 45]. Thus, we consider the following discrete regularized problem

$$\min_{\gamma \in \mathcal{C}} \sum_{j_1, \dots, j_N} c_{j_1, \dots, j_N} \gamma_{j_1, \dots, j_N} + \epsilon E(\gamma), \tag{6.6}$$

where  $E(\gamma)$  is defined as follows

$$E(\gamma) = \begin{cases} \sum_{j_1, \dots, j_N} \gamma_{j_1, \dots, j_N} \log(\gamma_{j_1, \dots, j_N}) & \text{if } \gamma \geq 0 \\ +\infty & \text{otherwise.} \end{cases} \tag{6.7}$$

After elementary computations, we can re-write the problem as

$$\min_{\gamma \in \mathcal{C}} KL(\gamma | \bar{\gamma}) \tag{6.8} \quad \boxed{\text{eq22}}$$

where  $KL(\gamma | \bar{\gamma}) = \sum_{i_1, \dots, i_N} \gamma_{i_1, \dots, i_N} (\log(\frac{\gamma_{i_1, \dots, i_N}}{\bar{\gamma}_{i_1, \dots, i_N}}))$  is the Kullback-Leibler distance and

$$\bar{\gamma}_{i_1, \dots, i_N} = e^{-\frac{c_{j_1, \dots, j_N}}{\epsilon}}. \tag{6.9} \quad \boxed{\text{barg}}$$

As explained in section 2, when the transport plan  $\gamma$  is concentrated on the graph of a transport map which solves the Monge problem, after discretisation of the densities, this property is lost along but we still expect the  $\gamma$  matrix to be sparse. The entropic regularization spreads the support and this helps to stabilize the computation as it defines a strongly convex program with a unique solution  $\gamma^\epsilon$ . Moreover the solution  $\gamma^\epsilon$  can be obtained through elementary operations. The regularized solutions  $\gamma^\epsilon$  then converge to  $\gamma^*$  (see figure 2), the solution of (6.1) with minimal entropy, as  $\epsilon \rightarrow 0$  (see [32] for a detailed asymptotic analysis and the proof of exponential convergence).

In order to introduce the Iterative Proportional Fitting Procedure (IPFP), we consider the two marginals problem

$$\min_{\gamma \in \mathcal{C}} \sum_{i,j} c_{ij} \gamma_{ij}. \tag{6.10} \quad \boxed{\text{pd:MKdiscrete2}}$$

The aim of the IPFP is to find the KL projection of  $\bar{\gamma}$  on the set  $\mathcal{C} = \{\gamma_{ij} \in \mathbb{R}^{M_d} \times \mathbb{R}^{M_d} : \sum_j \gamma_{ij} = \mu_i\} \cap \{\gamma_{ij} \in \mathbb{R}^{M_d} \times \mathbb{R}^{M_d} : \sum_i \gamma_{ij} = \nu_j\}$ . By writing down the lagrangian associated

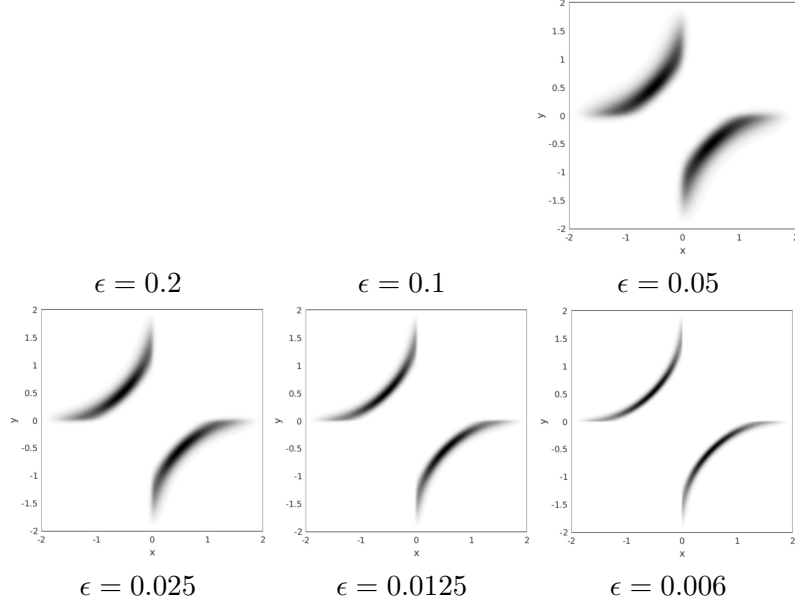


FIGURE 2. Support of the coupling  $\gamma^\epsilon$  for the Coulomb cost and all marginals  $\mu_i = (1 + \cos(\pi x/2))/2$ . The simulation has been performed on a discretization of  $[-2, 2]$  with  $M_d = 1000$ .

regularization

to (6.10) and computing the optimality condition, we find that  $\gamma_{ij}$  can be written as

$$\gamma_{ij} = a_i b_j \bar{\gamma}_{ij} \quad \text{with} \quad a_i = e^{u_i/\epsilon}, \quad b_j = e^{v_j/\epsilon}, \quad (6.11)$$

where  $u_i$  and  $v_j$  are the regularized Kantorovich potential. Then,  $a_i$  and  $b_j$  can be uniquely determined by the marginal constraint

$$a_i = \frac{\mu_i}{\sum_j b_j \bar{\gamma}_{ij}}, \quad b_j = \frac{\nu_j}{\sum_i a_i \bar{\gamma}_{ij}}. \quad (6.12)$$

Thus, we can now define the following iterative method

$$b_j^{n+1} = \frac{\nu_j}{\sum_i a_i^n \bar{\gamma}_{ij}}, \quad a_i^{n+1} = \frac{\mu_i}{\sum_j b_j^{n+1} \bar{\gamma}_{ij}}. \quad (6.13)$$

IPFP

*Remark 6.3.* In [89] Rüschendorf proves that the iterative method (6.13) converges to the KL-projection of  $\bar{\gamma}$  on  $\mathcal{C}$ .

*Remark 6.4.* Rüschendorf and Thomsen (see [91]) proved, in the measure continuous framework, that a unique KL-projection exists and takes the form  $\gamma(x, y) = a(x) \otimes b(y) \bar{\gamma}(x, y)$  (where  $a(x)$  and  $b(y)$  are non-negative functions).

The extension to the multi-marginal framework is straightforward but cumbersome to write. It leads to a problem set on  $N M_d$ -dimensional vectors  $a_{j, i_{(\cdot)}}$ ,  $j = 1, \dots, N$ ,  $i_{(\cdot)} = 1, \dots, M_d$ . Each update takes the form

$$a_{j, i_j}^{n+1} = \frac{\rho_{i_j}}{\sum_{i_1, i_2, \dots, i_{j-1}, i_{j+1}, \dots, i_N} \bar{\gamma}_{i_1, \dots, i_N} a_{1, i_1}^{n+1} a_{2, i_2}^{n+1} \dots a_{j-1, i_{j-1}}^{n+1} a_{j+1, i_{j+1}}^n \dots a_{N, i_N}^n}. \quad (6.14)$$

ip3

**6.2. Numerical experiments: Coulomb cost.** We present now some results for the multi-marginal problem with Coulomb cost in the real line. We consider the  $N$  marginals equals to a density  $\mu$  (we work in a *DFT* framework so the marginals represent the electrons which are indistinguishable). We recall that splitting  $\mu$  into  $N$   $\tilde{\mu}_i$  with equal mass ( $\int \tilde{\mu}_i(x)dx = \frac{1}{N} \int \mu(x)dx$ ), we will have  $\tilde{\mu}_i \rightarrow \tilde{\mu}_{i+1}$   $i = 1, \dots, N-1$  and  $\tilde{\mu}_N \rightarrow \tilde{\mu}_1$ .

The simulations in figure 3 are all performed on a discretization of  $[-5, 5]$  with  $M_d = 200$ , with marginals  $\mu_i = \mu(x) = \frac{N}{10}(1 + \cos(\frac{\pi}{5}x))$   $i = 1, \dots, N$  and  $\epsilon = 0.02$ . If we focus on the support of the coupling  $\tilde{\gamma}_{12}(x, y)$  we can notice that the numerical solution correctly reproduces the prescribed behavior: the transport plan is induced by a cyclical optimal maps (see section 3).

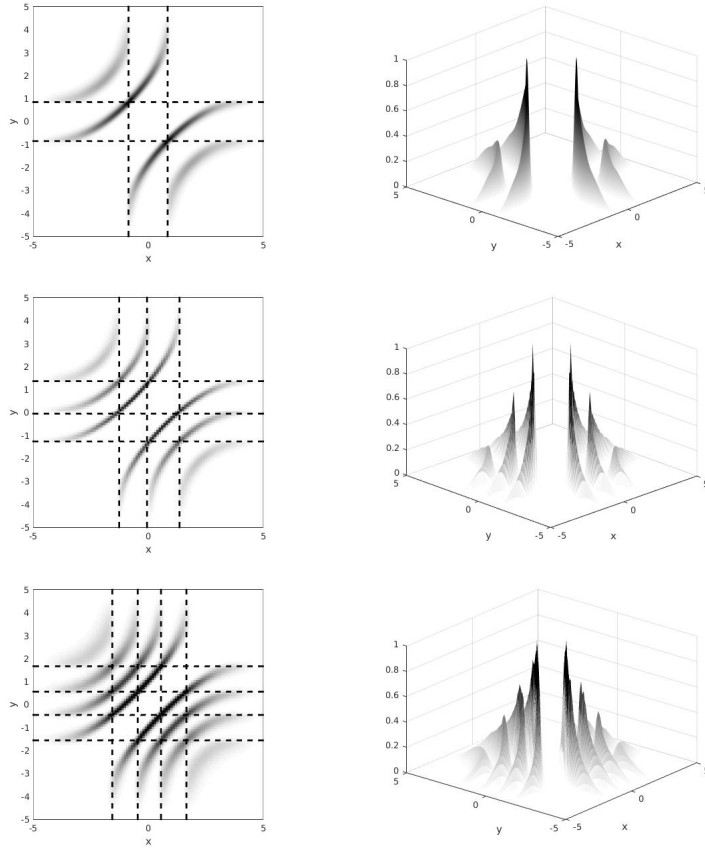


FIGURE 3. (1st Row) Left: support of the optimal coupling  $\tilde{\gamma}_{12}(x, y)$  for  $N = 3$ . Right: support of the optimal coupling  $\tilde{\gamma}_{12}(x, y)$  for  $N = 3$  (2nd Row) Left: support of the optimal coupling  $\tilde{\gamma}_{12}(x, y)$  for  $N = 4$ . Right: optimal coupling  $\tilde{\gamma}_{12}(x, y)$  for  $N = 4$ . (3rd Row) Left: support of the optimal coupling  $\tilde{\gamma}_{12}(x, y)$  for  $N = 5$ . Right: optimal coupling  $\tilde{\gamma}_{12}(x, y)$  for  $N = 5$ . The dashed lines delimit the intervals where  $\tilde{\mu}_i$ , with  $i = 1, \dots, N$ , are defined.

*Remark 6.5.* Theorem 3.7 actually works also for other costs functions as  $c(x_1, \dots, x_N) = \sum_{i < j}^N -\log(|x_i - x_j|)$ . This means that if we use the density  $\mu(x) = \frac{N}{10}(1 + \cos(\frac{\pi}{5}x))$ , we expect to obtain the same solution of the Coulomb cost. Thus if we now consider the 3-marginals case and the discretized cost  $c_{j_1, j_2, j_3} = -\log(|x_{j_1} - x_{j_2}|) - \log(|x_{j_1} - x_{j_3}|) - \log(|x_{j_2} - x_{j_3}|)$ , we can

notice (see figure 4) that we recover the same optimal coupling of the corresponding case for Coulomb (first row of figure 3). The simulation is performed on a discretization  $[-5, 5]$  with  $M_d = 200$ , with marginals  $\mu(x) = \frac{N}{10}(1 + \cos(\frac{\pi}{5}x))$   $N = 3$  and  $\epsilon = 0.01$ .

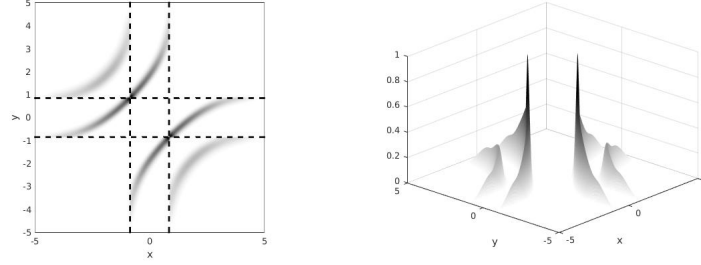
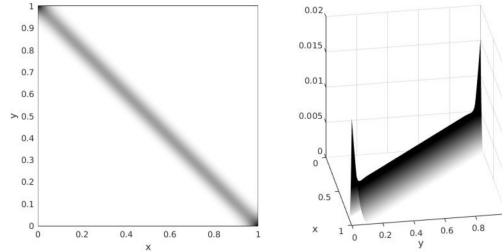


FIGURE 4. (logarithmic cost) Left: support of the optimal coupling  $\tilde{\gamma}_{12}(x, y)$  for  $N = 3$ . Right: support of the optimal coupling  $\tilde{\gamma}_{12}(x, y)$  for  $N = 3$ . The dashed lines delimit the intervals where  $\tilde{\mu}_i$ , with  $i = 1, \dots, 3$ , are defined.

figure:logNcos

**6.3. Numerical experiments: repulsive Harmonic cost.** As shown in section 4 the minimizers of the OT problem with the harmonic cost are also minimizers of the problem with  $c(x_1, \dots, x_N) = |x_1 + \dots + x_N|^2$ , so for the discrete problem (6.2) we take  $c_{j_1 \dots j_N} = |x_{j_1} + \dots + x_{j_N}|^2$ . Let us firstly consider the two marginal case and the uniform density on  $[0, 1]$  (see example 4.2), and, as expected, we find a deterministic coupling given by  $\gamma(x, y) = \mu(x)\delta(y - T(x))$  where  $T(x) = 1 - x$ , see figure 5. The simulation in figure 5 has been performed on a discretization of  $[0, 1]$  with  $M_d = 1000$  gridpoints and  $\epsilon = 0.005$ .



onic2marginals

FIGURE 5. Left: support of the optimal coupling. Right: optimal coupling.

The multi-marginals case is more delicate to treat: the original OT problem does not admit a unique solution where as the regularized problem does. As we have explained in section 6.1 the regularized problem is a strongly convex problem which admits a unique solution and the resulting coupling is the one with the minimal entropy. However we are able to make the IPFP algorithm converge to a selected coupling among the optimal ones for the original problem. Let us focus on the example 4.12. In this case we have all marginals equals to  $\mu = \frac{1}{2}\mathcal{L}|_{[-1,1]}$  and we can find a deterministic coupling given by  $\gamma(x, y, z) = \mu(x)\delta(y - T(x))\delta(z - S(x))$  (see example 4.12 for the maps  $T(x)$  and  $S(x)$ ). In order to select the desired coupling, the idea is to modify the cost function by adding a penalization term  $p(x, y) = \frac{\tau}{|x-y|}$  which makes  $\mu_1$  and  $\mu_2$  be *furthest as possible*. Thus, the discretized cost now reads as

$$c_{j_1, j_2, j_3} = |x_{j_1} + x_{j_2} + x_{j_3}|^2 + \frac{\tau}{|x_{j_1} - x_{j_2}|}. \quad (6.15)$$

In figure 6 we visualize the projection  $\tilde{\gamma}_{12}(x, y) = \pi_{12}\gamma(x, y, z)$  and  $\tilde{\gamma}_{13}(x, z) = \pi_{13}\gamma(x, y, z)$  of the computed coupling. The simulation has been performed on a discretization of  $[-1, 1]$  with  $M_d = 1000$  gridpoints,  $\epsilon = 0.0005$  and  $\eta = 0.1$ .

We, finally, take the same marginals as in the previous example, but we do not add the penal-

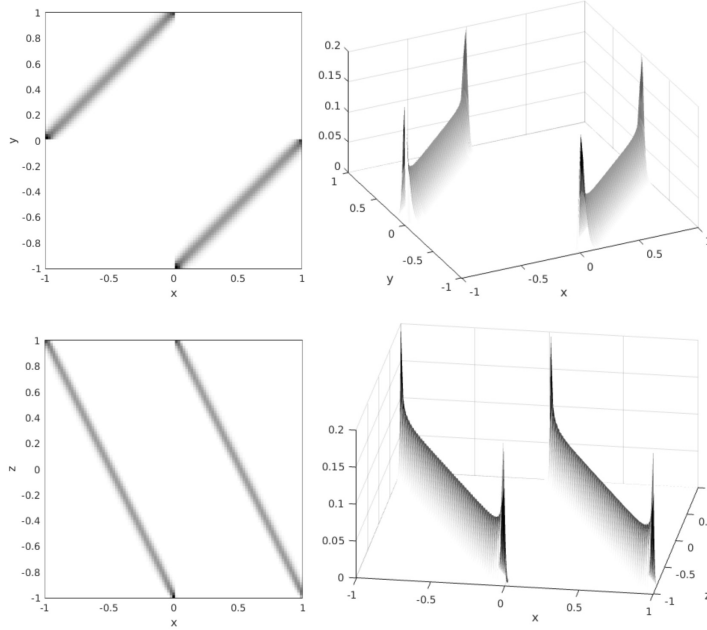


FIGURE 6. *Top-Left: support of the optimal coupling  $\tilde{\gamma}_{12}(x, y)$ . Top-Right: optimal coupling  $\tilde{\gamma}_{12}(x, y)$ . Bottom-Left: support of the optimal coupling  $\tilde{\gamma}_{13}(x, z)$ . Bottom-Right: optimal coupling  $\tilde{\gamma}_{13}(x, z)$ .*

ization. In this case we expect a diffuse plan since, if we forget about the marginal constraint, it is simple to show that the second order  $\Gamma$ -limit of the regularized problem as  $\epsilon \rightarrow 0$  is the relative entropy of  $\gamma$  with respect to  $\mathcal{H}^2|_{\{x+y+z=0\}}$ ; we expect that the  $\Gamma$ -limit is the same also when we add the marginal constraint.

In figure 7 we present the projection  $\tilde{\gamma}_{12}(x, y) = \pi_{12}\gamma(x, y, z)$  of the computed coupling (it is enough to visualize only this projection because of the symmetries of the problem). The simulation has been performed on a discretization of  $[-1, 1]$  with  $M_d = 1000$  gridpoints,  $\epsilon = 0.0005$ .

**6.4. Numerical experiments: Determinant cost.** Numerical simulations for the multi-marginal problem with the determinant cost present an obvious computational difficulty: in order to compute the solution just for the 3-marginals case, we have to introduce a discretization of  $\mathbb{R}^3$ .

*Remark 6.6.* We remind the reader that  $c(x_1, \dots, x_N) = \det(x_1, \dots, x_N)$  where  $x_i \in \mathbb{R}^N$

However if we take radially symmetric densities  $\mu_i$  as marginals (see section 5 for a more complete description of the problem in the measure framework), Carlier and Nazaret show that  $(\mathcal{MK})$  can be reduced to a 1-dimensional problem: the only unknown are the relations between the norm of each vector  $r_i = \|x_i\|$ . These relations can be obtained by solving the following

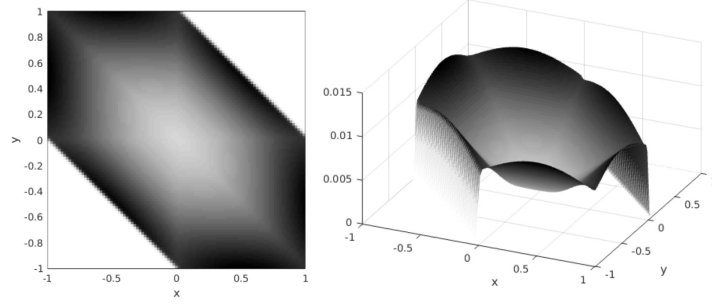


FIGURE 7. *Left: support of the optimal coupling  $\tilde{\gamma}_{12}(x, y)$ . Right: optimal coupling  $\tilde{\gamma}_{12}(x, y)$ .*

figure:diffplan

problem

$$\min_{\gamma \in \Pi(\lambda_1(r), \dots, \lambda_N)} - \int \left( \prod_{i=1}^N r_i \right) \gamma(r_1, \dots, r_N) dr_1 \cdots dr_N, \quad (6.16)$$

radialPb

where  $\lambda_i(r) = \sigma(r)\mu_i(\mathbf{r})$  (i.e.  $N = 3$  then  $\mathbf{r} = (r, \theta, \varphi)$  and  $\sigma(r) = 4\pi r^2$ ). Moreover, for this problem we know that there exists a unique (deterministic) optimal coupling (see Proposition 5 in [21]). The discretized cost now reads  $c_{j_1 \dots j_N} = \prod_{k=1}^N r_{j_k}$ . Let us consider the 3-marginals case and all densities  $\lambda_i$  equals to the uniform on the ball  $\mathcal{B}_{0.5} = \{x \in \mathbb{R}^3 \mid \|x\| \leq 0.5\}$ . As proved in [21], the optimal coupling is actually supported by the graph of maps which rearrange measure (see figure 8 ). The simulation in figure 8 has been performed on a discretization of  $[0, 0.5]$  with  $M_d = 100$  and  $\epsilon = 0.01$ . In the same way we can take all marginals  $\lambda_i(r) = 4\pi r^2 e^{-4r}$  and

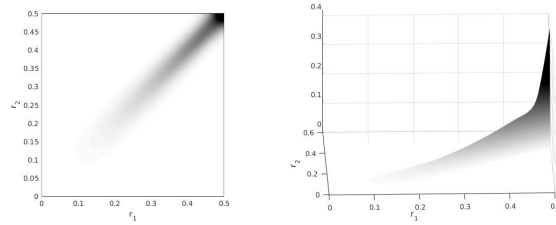


FIGURE 8. *(Uniform density) Left: support of the optimal coupling  $\tilde{\gamma}_{12}(r_1, r_2)$ . Right: optimal coupling  $\tilde{\gamma}_{12}(r_1, r_2)$ .*

ure:detUniform

we obtain again a deterministic coupling, figure 9. The simulation has been performed on a discretization of  $[0, 3]$  with  $M_d = 300$  and  $\epsilon = 0.05$ .

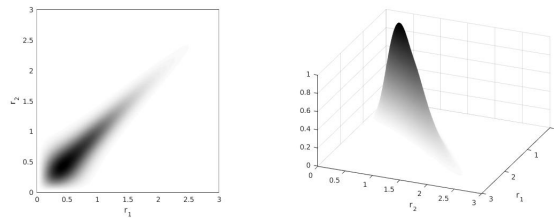


FIGURE 9. *(Exponential density) Left: support of the optimal coupling  $\tilde{\gamma}_{12}(r_1, r_2)$ . Right: optimal coupling  $\tilde{\gamma}_{12}(r_1, r_2)$ .*

detExponential

We, finally, present a simultaion (see figure 11) with three different marginals  $\lambda_i$  (see figure 10). The simulation has been performed on a discretization of  $[0, 4]$  with  $M_d = 300$  and  $\epsilon = 0.07$ .

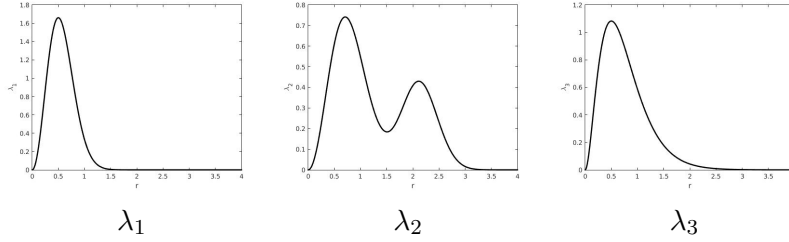


FIGURE 10. *Densities  $\lambda_1$ ,  $\lambda_2$  and  $\lambda_3$ .*

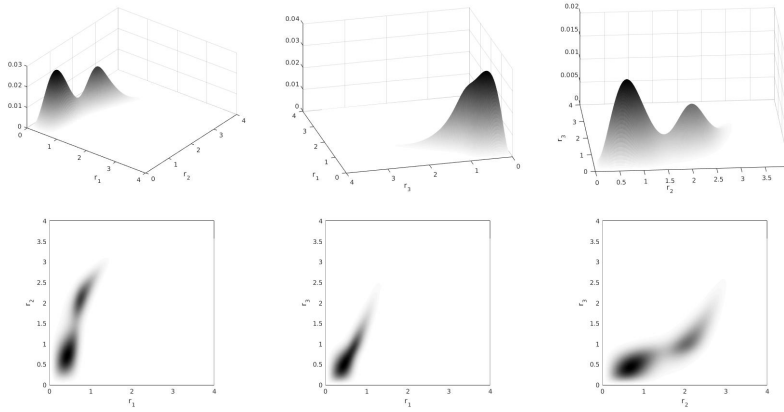


FIGURE 11. *Top-Left: optimal coupling  $\tilde{\gamma}_{12}(r_1, r_2)$ . Top-Center: optimal coupling  $\tilde{\gamma}_{13}(r_1, r_3)$ . Top-Right: optimal coupling  $\tilde{\gamma}_{23}(r_2, r_3)$ . Bottom-Left: support of the optimal coupling  $\tilde{\gamma}_{12}(r_1, r_2)$ . Bottom-Center: support of the optimal coupling  $\tilde{\gamma}_{13}(r_1, r_3)$ . Bottom-Right: support of the optimal coupling  $\tilde{\gamma}_{23}(r_2, r_3)$ .*

## CONCLUSION

blalblalalal

## ACKNOWLEDGEMENTS

We would like to thank Paola Gori-Giorgi for many helpful and stimulating discussions.

Luca Nenna gratefully acknowledges the support of the ANR, through the project ISOTACE (ANR-12-MONU-0013) and INRIA through the “action exploratoire” MOKAPLAN.

## REFERENCES

- [1] I. Abraham, R. Abraham and M. Bergounioux and G. Carlier, *Tomographic reconstruction from a few views: a multi-marginal optimal transport approach*. preprint, available at <https://www.ceremade.dauphine.fr/~carlier/aabc.pdf>;
- [2] M. Agueh and G. Carlier, *Barycenters in the Wasserstein space*. SIAM J. Math. Anal., 43(2) (2011);
- [3] G. Alberti, *On the structure of singular sets of convex functions*. Calc. Var. Partial Differential Equations 2, (2000);



- [4] L. Ambrosio, N. Gigli and G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lectures in Mathematics. ETH Zrich (2008);
- [5] A. Becke, *Density-functional thermochemistry III. The role of exact exchange*. J. Chem. Phys. 98, 5648 (1993);
- [6] M. Beiglbock, A. Cox and M. Huesmann, *Optimal Transport and Skorokhod Embedding*. preprint;
- [7] M. Beiglbock and C. Griessler, *An optimality principle with applications in optimal transport*. Preprint available at <http://arxiv.org/abs/1404.7054>;
- [8] M. Beiglbock, P. Henry-Labordere, and F. Penkner, *M. Beiglbock, P. Henry-Labordere, and F. Penkner. Model independent bounds for option prices: a mass transport approach*. Finance Stoch.,17:477501 (2013);
- [9] M. Beiglbock and N. Juillet, *On a problem of optimal transport under marginal martingale constraints*. Finance Stoch.,17:477501 (2013);
- [10] M. Beiglboeck, C. Leonard and W. Schachermayer, *A General Duality Theorem for the Monge–Kantorovich Transport Problem*. Studia Math. 209, no. 2 (2012);
- [11] JD Benamou, G. Carlier, M. Cuturi, L. Nenna, and G. Peyré, *Iterative Bregman Projections for Regularized Transportation Problems* SIAM Journal on Scientific Computing 37:2, A1111-A1138 (2015);
- [12] JD Benamou, G. Carlier and L. Nenna, *A Numerical Method to solve Optimal Transport Problems with Coulomb Cost*, preprint at arXiv:1505.01136 (2015).
- [13] J. Bigot and T. Klein, *Consistent estimation of a population barycenter in the Wasserstein space*. Proceedings of the International Conference Statistics and its Interaction with Other Disciplines, pages 153-157 (2013);
- [14] E. Bodo, *Applicazioni di meccanica quantistica: appunti per le lezioni*. Lecture notes for a course in quantum chemistry. Available (April/2015) at <http://w3.uniroma1.it/bodo/dispense/appl.pdf>, in italian (2008);
- [15] Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*. Comm. Pure Appl. Math. 44 (1991);
- [16] G. Buttazzo, L. De Pascale and P. Gori-Giorgi, *Optimal-transport formulation of electronic density-functional theory*. Phys. Rev. A 85, 062502;
- [17] L. Caffarelli, *A localization property of viscosity solutions to the Monge-Ampère equation and their strict convexity*. Ann. of Math. (2), 131 (1990);
- [18] L. Caffarelli, *Some regularity properties of solutions to Monge-Ampère equations*. Comm. Pure Appl. Math., 44 (1991);
- [19] G. Carlier and I. Ekeland, *Matching for teams*. Econom. Theory, 42(2) (2010);
- [20] G. Carlier, A. Galichon and V. Chernozhukov, *Vector Quantile Regression*. Available at <https://www.ceremade.dauphine.fr/~carlier/VQR.pdf>;
- [21] G. Carlier and B. Nazaret, *Optimal transportation for the determinant*. ESAIM Control Optim. Calc. Var., 14(4) (2008);
- [22] G. Carlier, A. Oberman and E. Oudet, *Numerical methods for matching for teams and Wasserstein Barycenters*. pre-print, available on arXiv:1411.3602v1 (2014);
- [23] P.A. Chiappori, A. Galichon, and B. Salanie, *The roommate problem is more stable than you think*. Available at <http://www.columbia.edu/~bs2237/RoommatePbm.pdf>;
- [24] P-A. Chiapporri, R. McCann, and L. Nesheim, *Hedonic price equilibria stable matching and optimal transport; equivalence, topology and uniqueness*. Econom. Theory., 42(2):317?354 (2010);
- [25] H. Chen, G. Friesecke, and C. Mendl. *Numerical methods for a Kohn-Sham density functional model based on optimal transport*. Preprint at <http://arxiv.org/abs/1405.7026>, (2014);
- [26] H. Chen and G. Friesecke. *Pair densities in density functional theory*. Preprint at arXiv:1503.07539, (2015);
- [27] A. J. Cohen, P. Mori-Shez and W. Yang, *Challenges for Density Functional Theory*. Chem. Rev., 2012, 112 (1), (2011);
- [28] M. Colombo and S. Di Marino, *Equality between Monge and Kantorovich multimarginal problems with Coulomb cost*. Ann. Mat. Pura Appl. (2013);
- [29] M. Colombo, L. De Pascale and S. Di Marino, *Multimarginal Optimal Transport maps for 1-dimensional Repulsive costs*. Canadian Journal of Mathematics (2013);
- [30] R. Cominetti, J.S. Martin, *Asymptotic analysis of the exponential penalty trajectory in linear programming*. Mathematical Programming 67(1-3), 169?187 (1994);
- [31] M. Colombo, F. Stra, *Counterexamples to multimarginal optimal transport maps with Coulomb cost and radial measures*. in preparation;

- [32] R. Cominetti, J.S. Martin, *Asymptotic analysis of the exponential penalty trajectory in linear programming*. Mathematical Programming 67(1-3), 169-187 (1994);
- [33] C. Cotar, G. Friesecke and C. Klupperlberg, *Density functional theory and optimal transportation with Coulomb cost*. Comm. Pure Appl. Math. 66, (2013);
- [34] C. Cotar, G. Friesecke and B. Pass, *Infinite-body optimal transport with Coulomb cost*. To appear in Calc. Var. Partial Differential Equations (2015);
- [35] M. Cuturi, *Sinkhorn distances: Lightspeed computation of optimal transport*. Advances in Neural Information Processing Systems (NIPS) 26, pp. 2292-2300 (2013);
- [36] L. De Pascale, *Optimal Transport with Coulomb cost. Approximation and duality*. preprint available on arXiv:1503.07063;
- [37] S. Di Marino, P. Gori-Giorgi, L. Nenna, *The strictly-correlated electron functional for spherically symmetric systems revisited*. in preparation;
- [38] Y. Dolinsky and H. M. Soner, *Martingale optimal transport and robust hedging in continuous time*. submitted to Probability Theory and Related Fields (2013);
- [39] Y. Dolinsky and H. M. Soner, *Robust hedging with proportional transaction costs*. preprint (2013);
- [40] L. C. Evans and M. Zworski, *Lectures on Semiclassical Analysis* Version 0.2. available (april/2015) on <https://math.berkeley.edu/~evans/semiclassical.pdf> (2003);
- [41] U. Frisch, S. Matarrese, R. Mohayaee, and A. Sobolevski, *Monge-Ampere-Kantorovitch (MAK) reconstruction of the early universe*. Nature ,417(260), (2002).
- [42] G. Friesecke, C. B. Mendl, B. Pass, C. Cotar and C. Klupperlberg, *N-density representability and the optimal transport limit of the Hohenberg-Kohn functional*. J. Chem. Phys. 139 (2013);
- [43] G. Friesecke, *The multi configuration equation for atoms and molecules: charge quantization and existence of solutions*. Arch. Rat. Mech. Analysis 169, (2003);
- [44] I. Ekeland, *An optimal matching problem*. ESAIM Control. Optim. Calc. Var. (2005);
- [45] A. Galichon, B. Salanié, *Matching with trade-offs: Revealed preferences over competing characteristics*. Tech. rep., Preprint SSRN-1487307 (2009);
- [46] A. Galichon, P. Henry-Labordere, and N. Touzi, *A stochastic control approach to non-arbitrage bounds given marginals, with an application to Lookback options*. The Annals of Applied Probability (2014);
- [47] W. Gangbo and A. Swiech, *Optimal Maps for the Multidimensional Monge-Kantorovich Problem*. Commu. Pure Applied Math. (1998);
- [48] W. Gangbo and V. Oliker, *Existence of Optimal Maps in the Reflector-type Problems*. Commu. Pure Applied Math. (1998); COCV, ESAIM 13, no 1, 93-106 (2007);
- [49] N. Ghoussoub and B. Maurey, *Remarks on multi-marginal symmetric Monge-Kantorovich problems*. Discrete and Continuous Dynamical Systems-A, Issue 4, (2014);
- [50] N. Ghoussoub and A. Moameni, *Symmetric Monge-Kantorovich problems and polar decompositions of vector fields*. Geometric and Functional Analysis, Volume 24, Issue 4 (2014);
- [51] N. Ghoussoub and B. Pass, *Decoupling of DeGiorgi-type systems via multimarginal optimal transport*. Comm. Partial Differential Equations, 6:1032-1047 (2014).
- [52] P. Gori-Giorgi, M. Seidl, and G. Vignale, *Density-Functional Theory for Strongly Interacting Electrons*. Physical Review Letters 103, 166402 (2009);
- [53] L. V. Kantorovich, *On the translocation of masses*. J. Math. Sci., 2006, vol. 133, no. 4, pp. 1381-1382, doi:10.1007/s10958-006-0049-2. Translated from Dokl. Akad. Nauk SSSR, vol. 37, no. 7-8, pp. 227-229 (1942);
- [54] L. V. Kantorovich, *On a problem of Monge*. J. Math. Sci., 2006, vol. 133, no. 4, pp. 1383, doi:10.1007/s10958-006-0050-9. Translated from Uspekhi Mat. Nauk, vol. 3, no. 2, 225-226 (1948);
- [55] H.G. Kellerer, *Duality theorems for marginal problems*. Z. Wahrsch. Verw. Gebiete, 67:399-432 (1984);
- [56] Y.-H. Kim and B. Pass, *Multi-marginal optimal transport on a Riemannian manifold*. Preprint available at arXiv:1303.6251;
- [57] Y.-H. Kim and B. Pass, *A general condition for Monge solutions in the multi-marginal optimal transport problem*. SIAM J. Math. Anal, 46 (2014);
- [58] Y.-H. Kim and B. Pass, *Wasserstein Barycenters over Riemannian manifolds*. preprint available on arXiv:1412.7726;
- [59] J. Kitagawa and B. Pass, *The multi-marginal optimal partial transport problem*. Preprint available at <http://arxiv.org/abs/1401.7255>;

- [60] M. Knott and C.S. Smith, *On a generalization of cyclic monotonicity and distances among Random vectors.* linear Algebra Appl. 199, 363-371 (1994);
- [61] W. Kohn, *Nobel Lecture: Electronic structure of matter – wave functions and density functionals.* Review Modern Physics 71, 1253 (1999);
- [62] A. V. Kolesnikov, D. A. Zhev, *Optimal transportation of processes with infinite Kantorovich distance. Independence and symmetry.* preprint, available on arXiv:1303.7255;
- [63] D. C. Langreth and J. P. Perdew, *The Exchange Correlation Energy of a Metallic Surface.* Solid State Communications. Vol 17, pp 1425-1429 (1975);
- [64] C. Lee, W. Yang and R. G. Parr, *Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density.* Physics Review B 37, 785-789 (1988);
- [65] E.H. Lieb, *Density Functionals for Coulomb systems.* International Journal of Quantum Chemistry 24 (1983);
- [66] E.H. Lieb, *A lower bound for Coulomb energies.* Phys. Lett. A 70, 444-446 (1979);
- [67] E.H. Lieb and S. Oxford, *Improved lower bound on the indirect Coulomb energy.* Int. J. Quant. Chem., 19, 427-439 (1981);
- [68] E.H. Lieb and B. Simon, *The Hartree-Fock Theory for Coulomb Systems.* Commun Math Physics, 53 (1977);
- [69] P-L. Lions, *Solutions of The Hartree-Fock equations for Coulomb Systems.* Commun Math Physics, 109 (1987);
- [70] P-L. Lions, *Solutions of The Hartree-Fock equations for Coulomb Systems.* Commun Math Physics, 109 (1987);
- [71] A.O. Lopes, J.K. Mengue, *Duality Theorems in Ergodic Transport.* Journ Stat. Phys., 149(5), 921-942 (2012);
- [72] A.O. Lopes, E. O. Oliveira, P. Thieullen *The Dual Potential, the involution kernel and Transport in Ergodic Optimization.* preprint available at arXiv: 1111.0281;
- [73] F. Malet, P. Gori-Giorgi *Strong correlation in kohn-sham density functional theory.* Phys. Rev. Lett. 109, 246,402 . DOI 10.1103/PhysRevLett.109.246402. URL <http://link.aps.org/doi/10.1103/PhysRevLett.109.246402>,(2012);
- [74] C. Mendl and L. Lin, *Towards the Kantorovich dual solution for strictly correlated electrons in atoms and molecules.* Physical Review B, 87:125106, (2013);
- [75] G. Monge, *Mémoire sur la théorie des déblais et des remblais.* In Histoire de l'Académie royale des sciences de Paris (année 1781), pp. 666-704. Imprimerie royale (1784);
- [76] I. Olkin and S.T. Rachev, *Maximum submatrix traces for positive definite matrices.* SIAM J. Matrix Ana. Appl. 14, 390-397 (1993);
- [77] B. Pass, *Multi-marginal optimal transport: theory and applications.* To appear in ESAIM: Math. Model. Numer. Anal. (Special issue on "Optimal transport in applied mathematics.") available on arXiv:1406.0026;
- [78] B. Pass, *Multi-marginal optimal transport and multi-agent matching problems: uniqueness and structure of solutions.* Discrete Contin. Dyn. Syst., 34:1623-1639 (2014);
- [79] B. Pass, *Structural results on optimal transportation plans.* PhD Thesis at the University of Toronto;
- [80] B. Pass, *On the local structure of optimal measures in the multi-marginal optimal transportation problem.* Calc. Var. (2012);
- [81] B. Pass, *Uniqueness and Monge solutions in the multimarginal optimal transportation problem.* SIAM J. Math. Anal., 43 (2011);
- [82] B. Pass, *Remarks on the semi-classical Hohenberg-Kohn functional.* Nonlinearity 26 (2013);
- [83] B. Pass and Y-H Kim, *A general condition for Monge solutions in the multi-marginal optimal transport problem* SIAM J. Math. Anal. 46 (2014) 1538-1550.
- [84] B. Pass and Y-H Kim, *Multi-marginal optimal transport on Riemannian manifolds.* Pre-print, available on arXiv:1412.7726 (2014);
- [85] S. Pazzani, S. Moroni, P. Gori-Giorgi, and G.B. Bachelet, *Local-spin-density functional for multideterminant density functional theory.* Phys. Rev. B 73, 155111 (2006);
- [86] A. Pratelli, *On the equality between Monge's infimum and Kantorovich's minimum in optimal mass transportation.* Ann. Inst. Henri Poincaré Probab. Statist., 43 (2007), 1-13;
- [87] J. Rabin, G. Peyre, J. Delon, and M. Bernot, *Wasserstein barycenter and its application to texture mixing.* In Scale Space and Variational Methods in Computer Vision, pages 435-446 (2012);
- [88] D. Robert, *Autour de l'approximation semi-classique.* Progress in Mathematics 68, Birkhäuser (1987);

- [89] L. Rüschendorf, *Convergence of the iterative proportional fitting procedure*. The Annals of Statistics 23(4), 1160?1174 (1995);
- [90] L. Rüschendorf and L. Uckelmann, *On optimal multivariate couplings*. In Proceedings of Prague 1996 conference on marginal problems. Kluwer Acad. Publ. (1997);
- [91] L. Rüschendorf., W. Thomsen, *Closedness of sum spaces and the generalized Schrodinger problem*. Theory of Probability and its Applications 42(3), 483?494 (1998)
- [92] E. Schrödinger, E., *Über die umkehrung der naturgesetze*. Sitzungsberichte Preuss. Akad. Wiss. Berlin. Phys. Math. 144, 144?153 (1931);
- [93] M. Seidl, *Strong-interaction limit of density-functional theory*. Physical Review A 60, 4387 (1999).
- [94] M. Seidl, P. Gori-Giorgi and A. Savin, *Strictly correlated electrons in density-functional theory: A general formulation with applications to spherical densities*. Physical Review A 75, 042511 (2007);
- [95] C. Smith and M. Knott, *On Hoeffding-Fréchet bounds and cyclic monotone relations*. Journal of multivariable Analysis, 40 (1992);
- [96] J. Toulouse, F. Colonna, A. Savin, *Short-range exchange and correlation energy density functionals: Beyond the local-density approximation*. J. Chem. Phys. 122, 14110 (2005);
- [97] C. Villani, *Optimal Transport: Old and New*. Grundlehren der mathematischen Wissenschaften. Springer. ISBN-10: 3540710493 (2009);
- [98] W. Yang, *Generalized adiabatic connection in density functional theory*. Journal of Chemical Physics, Volume 109, Issue 23, pp. 10107-10110 (1998);
- [99] G.-S. Xia, S. Ferradans, G. Peyré, and J.-F. Aujol, *Synthesizing and mixing stationary Gaussian texture models*. SIAM J. Imaging Sci., 7 , pp 476-508,(2014).
- [100] D. Zaev, *On the Monge-Kantorovich problem with additional linear constraints*. Preprint available on arXiv:1404.4962;
- [101] G.M. Zhislin, *Discussion of the spectrum of Schrödinger operators for systems of many particles*. Trudy Moskovskogo matematicekogo obscestva 0, in russian (1960);

UNIVERSITÉ PARIS-SUD,

Email address: `simone.dimarino@math.u-psud.fr`

DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DI PISA, LARGO B. PONTECORVO 5, 56126 PISA, ITALY

Email address: `gerolin@mail.dm.unipi.it`

INRIA, MOKAPLAN, DOMAINE DE VOLUCEAU LE CHESNAY, FRANCE

Email address: `luca.nenna@inria.fr`