

Optimal transportation theory with repulsive costs

Simone Di Marino, Augusto Gerolin, Luca Nenna

▶ **To cite this version:**

Simone Di Marino, Augusto Gerolin, Luca Nenna. Optimal transportation theory with repulsive costs. 2015. hal-01163737v2

HAL Id: hal-01163737

https://inria.hal.science/hal-01163737v2

Preprint submitted on 1 Dec 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

OPTIMAL TRANSPORTATION THEORY WITH REPULSIVE COSTS

SIMONE DI MARINO, AUGUSTO GEROLIN, AND LUCA NENNA

Contents

Introduction 1
Organization of this Survey 2
What is new? 3
1. Why Multi-marginal Transport Theory for repulsive costs? 3
1.1. Brief introduction to Quantum Mechanics of N-body systems 3
1.2. Probabilistic Interpretation and Marginals 6
1.3. Density Functional Theory (DFT) 6
1.4. "Semi-classical limit" and Optimal Transport problem 8
2. DFT meets Optimal Transportation Theory 10
2.1. Couplings and Multi-marginal Optimal Transportation problem 10
2.2. Multimarginal Optimal Transportation Problem 10
2.3. Dual formulation 12
2.4. Geometry of the Optimal Transport sets 14
2.5. Symmetric Case 15
3. Multi-marginal OT with Coulomb Cost 16
3.1. General theory: duality, equivalent formulations and many particles limit 16
3.2. The Monge problem: deterministic examples and counterexamples 18
4. Multi-marginal OT with repulsive Harmonic Cost 19
5. Multi-marginal OT for the Determinant 24
6. Numerics 27
6.1. The regularized problem and the Iterative Proportional Fitting Procedure 27
6.2. Numerical experiments: Coulomb cost 30
6.3. Numerical experiments: repulsive Harmonic cost 30
6.4. Numerical experiments: Determinant cost 32
7. Conclusion 34
Acknowledgements 34
References 34

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. We are interested in characterizing, 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 $c : (\mathbb{R}^d)^N \longrightarrow \mathbb{R}\cup\{\infty\}$ 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 μ_i as marginals, for all $i=1,\dots,N$. c is, typically, a sum of convex decreasing functions, as the Coulomb interaction cost [12, 27, 57, 45, 57, 82]

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

which is the case of most interest in DFT (see section 1). Also related to Quantum Mechanics, in the so-called weak interaction regime [12, 38, 73, 85], 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 focus on the multi-marginal problem for the determinant cost

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

studied by Carlier and Nazaret in [17].

Multi-marginal optimal transport showed to be powerful in many applications as, for instance, in economics [15, 21, 22, 35, 71], mathematical finance [5, 8, 7, 33, 34, 39], image processing [87], tomography [1] and statistics [16].

Although very natural in such applied problems, optimal transport problems with 3 or more marginals was only considered in literature during the 90's by Knott and Smith, Olkin and Rachev and Rüschendorf and Uckelmann [53, 66, 79], and later by Gangbo and Święch [41]. In [41], 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 [67, 68, 69] and Kim and Pass [49, 50]. Several theoretical developments were made last years, as the study of partial multi-marginal transport [52], symmetric Monge-Kantorovich problem [24, 31, 42, 43, 44, 54, 61, 62, 89] and duality [6, 30, 48, 89]. Applications in Geometry and Analysis as the Wasserstein Barycenter problem and decoupling a system of PDE's can be found in [2, 44, 51].

For a complete presentation of the multimarginal optimal transport theory with general costs functions, including a discussion of many applications, we refer to the survey [72].

In this paper, 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 computing accurately the ground state of a 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 the interesting case is when all marginals of the coupling γ 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. But when a cost can be considered *repulsive*? It is difficult to find a good notion for this concept since usually it is based on the intuition that comes solving the particular problem. However a possible formal definition for repulsive cost is to require a mild condition, namely that we don't want any two particles in the *same* position; it should be cheaper to have them in different position. For example can write the inequality for the particles 1 and 2:

$$\begin{aligned} c(x_1,x_1,x_3,\dots,x_N) + c(x_2,x_2,x_3,\dots,x_N) \geq \\ c(x_1,x_2,x_3,\dots,x_N) + c(x_2,x_1,x_3,\dots,x_N). \end{aligned}$$

Then, also all other inequalities for $i \neq j$ should be considered. This condition is satisfied by all the considered costs, even the ones that doesn't seem so "repulsive" at first, such as those of the form $f(x_1 + \ldots + x_n)$ with f convex function.

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

Section 2 introduces 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 sections 3, 4 and 5, respectively, we discuss the recent development for the specific cases of transportation problem with Coulomb, repulsive harmonic and determinant cost.

Finally, in section 6 we introduce an iterative method, recently used in many applications of optimal transport [9, 10], and we present some numerical experiments for the problems studied in previous sections.

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

$$c(x_1,\dots,x_N)=f(x_1+\cdots+x_N),\quad f:\mathbb{R}^d\to\mathbb{R}\ \textit{is convex},$$

which, in particular, include the case of repulsive harmonic costs (see corollary 4.8). In addition, we show:

(i) *Existence of Fractal-like cyclical Monge solutions:* In Theorem 4.6, we show the existence of couplings $\gamma=(Id,T,T^{(2)},\dots,T^{(N-1)})_\sharp\mu$ concentrated on the graphs of functions $T,T^{(2)},\dots,T^{(N-1)}$, where μ 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 [13, 14].

(ii) *Optimal Transport plans supported on sets of Hausdorff dimension* >1*:* In standard optimal transportation (the 2 marginals case) the target is to understand if solutions of $(\mathcal{MK})$ are well-known to be Monge-type solutions: optimal transport plans are supported by the graph of a function T. In Section 4, we give a concrete example (see example 4.14) 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 from being satisfactory, mostly because they are not quantitative.

Classical behavior of atoms and molecules is described accurately, at least from 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 masses and the positions of the α-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 function is antisymmetric or *fermionic* if it changes 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,

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

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

In physics literature, the *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 it 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. ψ 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 ψ boils down to a (time-dependent) Schrödinger Equation

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

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

$$T_n = -\hbar \sum_{\alpha=1}^{M} \frac{1}{2m_\alpha} \Delta_{R_\alpha} \quad \textit{(Nuclei Kinetic Energy)}$$

$$T_e = -\hbar \sum_{i=1}^{N} \frac{1}{2m_e} \Delta_{x_i} \quad \textit{(Electron Kinetic Energy)}$$

$$V_{ne} = -\sum_{i=1}^{N} v_R(x_i) = -\sum_{i=1}^{N} \left(\sum_{\alpha=1}^{M} \frac{Z_\alpha}{|x_i - R_\alpha|} \right) \qquad \left(\begin{array}{c} \textit{Potential Energy of Inte-} \\ \textit{raction Nuclei-Electron} \end{array} \right)$$

$$V_{ee} = \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} \frac{1}{|x_i - x_j|} \quad \textit{(Interaction Electron-Electron)}$$

$$V_{nn} = \sum_{\alpha=1}^{M} \sum_{\beta=1, \beta\neq\alpha}^{M} \frac{Z_\alpha Z_\beta}{|R_\alpha - R_\beta|} \quad \textit{(Interaction Nuclei-Nuclei)}$$

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 is 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 $\psi(x, R, t) = \psi_e(x, R)\chi(R, t)$, which physically means that the dynamics of the electron ψ_e are decoupled of the dynamics of the atomic nuclei χ. Substituting that *ansatz* in (1.1), we can show that the electronic part $\psi_e(x, R)$ solves the following eigenvalue problem, so-called *Eletronic Schrödinger Equation*

$$H_e\psi_e = \lambda_e\psi_e, \qquad H_e = T_e + V_{ne} + V_{ee} + V_{nn} \tag{1.2}$$

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

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

We refer to [11] 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 nuclei are fully localized at single points representing classical point particles.

In Quantum Chemistry it is interesting to study the so called *Geometric Optimization Problem*.

The Geometric Optimization Problem: Compute the following minimizer

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

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* bellow).

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 eingenvalue (*"ground state energy"*) E_0 of the following linear operator, called *Electronic Hamiltonian*

Definition 1.1 (Eletronic Hamiltonian)**.** *The Eletronic Hamiltonian is a linear operator* $H_{el} : L^2_{anti}(\mathbb{R}^{3N}) \to L^2_{anti}(\mathbb{R}^{3N})$,

$$H_{el} = -\frac{\hbar^2}{2}\sum_{i=1}^{N}\Delta_{x_i} + \sum_{i=1}^{N} v_R(x_i) + \sum_{i=1}^{N}\sum_{j=i+1}^{N} f(x_j - x_i),$$

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

The first term of the Eletronic 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 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, it 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*

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

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

$$E_0 = \min\Big\{T_e[\psi] + V_{ne}[\psi] + V_{ee}[\psi] : \psi \in \mathcal{H}\Big\} \tag{1.4}$$

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

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

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

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

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

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

Here, the ground state quantum N-body problem refers to the problem of finding equilibrium states for system of type (1.1). A theorem proved first by Zhislin [90], guarantees the existence of the minimum in (1.4). Some variants can be found in the literature due, for instance, to Lieb & Simon [59], Lions [60] and Friesecke [37].

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 square norm of the wave function $\psi(x_1, s_1, \cdots, x_N, s_N)$ can be interpreted as a N-point probability density distribution for the electrons are in the points x_i with spins s_i.

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

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

$$\rho(x_1) = N \int_{\mathbb{R}^{3(N-1)}} \gamma_N(x_1, s_1, \cdots, x_N, s_N) dx_2, \cdots dx_N, \tag{1.5}$$

where γ_N represents the N-point position density

$$\gamma_N(x_1, \cdots, x_N) = \sum_{s_1, \cdots, s_N \in \mathbb{Z}_2} |\psi(x_1, s_1, \cdots, x_N, s_N)|^2. \tag{1.6}$$

Analogously, we can define the k-density

$$\gamma_k(x_1, \dots, x_k) = \binom{N}{k} \int_{\mathbb{R}^{3(N-k)}} \gamma_N(x_1, s_1, \cdots, x_N, s_N) dx_{k+1} \cdots dx_N \tag{1.7}$$

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

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

The relevance of ρ and γ_2 to compute the ground state energy (1.4) is due to the fact that the interation nuclei-nuclei energy $V_{ne}[\psi]$ and electron-electron energy $V_{ee}[\psi]$ depends only on these.

Lemma 1.2 ([27]). *If $\psi, \nabla\psi \in L^2((\mathbb{R}^3 \times \mathbb{Z}^2)^N, \mathbb{R})$, ψ antisymmetric and $\|\psi\|_2 = 1$, then*

$$V_{ne}[\psi] = \int_{\mathbb{R}^3} v(x)\rho(x)dx, \quad V_{ee}[\psi] = \int_{\mathbb{R}^6} \frac{1}{|x-y|} \gamma_2(x,y) dxdy.$$

.

It is natural to wonder about the space of those densities arising from such a density ψ is defined, i.e, the space $\mathcal{A}$ of densities $\rho : \mathbb{R}^3 \to \mathbb{R}$ verifying (1.5) for an antisymmetric ψ 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 (due to Lieb [57]):

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

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 [23]. DFT theory approximates quantum mechanics via variational principles for the marginal density

$$\rho(x_1) := \int_{\mathbb{R}^{d(N-1)}} |\psi(x_1, \dots, x_N)|^2 dx_2 \dots dx_N, \tag{1.9}$$

where ψ 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

$$\gamma_2(x_1, x_2) := \int_{\mathbb{R}^{3(N-2)}} |\psi(x_1, \dots, x_N)|^2 dx_3 \dots dx_N. \tag{1.10}$$

in terms of its marginal ρ. We simply write $\psi \to \rho$ and $\psi \to \gamma_2$ to denote the relation between ψ and ρ, and ψ and γ_2. This means, respectively, that ψ has single particle density ρ and pair density γ_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

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

with

$$F^{HK}[\rho] = \inf \left\{ T[\psi] + V_{ee}[\gamma_2] : \rho \in \mathcal{A}, \psi \to \rho \right\}. \tag{1.12}$$

where F_{HK} is the so called *Hohenberg-Kohn* or *Levy-Lieb functional*. For a mathematical sake of completeness we also consider a weak repulsive interaction given by

$$\hat{V}_{ee}[\rho] = \int_K \sum_{i=1}^N \sum_{j=1}^N -|x_j - x_i|^2 \gamma_N(x_1, \cdots, x_N) dx_1 \dots dx_N, \tag{1.13}$$

in a compact set $K \subset \mathbb{R}^{3N}$ and $+\infty$ otherwise.

The next theorem is the main result in Density Functional Theory.

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 $f : \mathbb{R} \to \mathbb{R}$ be a continuous function. There exists a functional $F^{HK} : \mathcal{P}(\mathbb{R}^3) \to \mathbb{R}$ depending only on the single-particle density ρ such that for any potential V_{nn}, the exact Quantum Mechanics ground state energy (1.4) satisfies*

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

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

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

where, $\psi \to \rho$ means that ψ has single-particle density ρ and $\mathcal{A}$ is the set of $\rho : \mathbb{R}^3 \to \mathbb{R}$ 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 *universal* in the sense that it does not depend on the molecular system under consideration. From a Physics point of view, it garantees that in a molecular system of N electrons, the single electron density ρ determines the pair density of the system (see corollary 1.4). From a mathematical perspective, the proof of this theorem is a functional analysis exercise that, for sake of completeness, we are going to present in a version we learnt from Gero Friesecke.

Proof. Firstly we remark that the non-universal part of the energy only depends on ρ:

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

Then, we rewrite problem (1.4) as a double inf problem

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

where $F^{HK}[\rho] = \inf_{\psi \mapsto \rho} \left(\langle \psi, \mathrm{H}_{el}\, \psi \rangle \right)$. □

The present version stated in 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).

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 $\gamma_2(x_1, x_2)$ - or even k-density $\gamma_k(x_1, \dots, x_k)$ - 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

$$\overline{\psi} \in \operatorname{argmin}\{\langle \psi, H_{el}\psi \rangle : \psi \to \rho\}$$

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

$$\gamma_k(x_1, \cdots, x_k) = \int |\overline{\psi}(x_1, \cdots, x_N)|^2 d_{k+1} \cdots dx_N$$

□

Of course, γ_k may be not unique, since the *ground state* may be degenerate. The existence of a map $T : \mathbb{R}^3 \to \mathbb{R}^3$ which determines the density of γ_2 from ρ 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 λ while keeping the density ρ fixed [55, 88]

$$F_\lambda^{HK}[\rho] = \min_{\substack{\psi\in\mathcal{H}\\ \psi\to\rho}} \left\{ \langle\psi, \big(-\frac{\hbar^2}{2}\Delta + \sum_{i=1}^{N}\sum_{j=i+1}^{N} \frac{\lambda}{|x_j - x_i|}\big)\psi\rangle \right\}. \tag{1.15}$$

The strictly correlated electron limit ($\lambda \to \infty$), up to relaxing in the space of probability measures, was first considered by two papers in physics literature: Seidl [82] and Seidl, Gori-Giorgi & Savin [83]. In [45], 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 mathematically rigorous in the two particles case by Cotar, Friesecke & Klüppelberg [27]. They considered the Hohenberg-Kohn functional $F^{HK} = F_\hbar^{HK}$ as a function of both ρ and $\hbar$ ($\lambda = 1$) and proved that - up to passage of the limit $\hbar \to 0$ - the F_{HK} reduces to the following functional

$$\tilde{F}[\rho] = \inf\left\{ \int_{\mathbb{R}^6} \frac{1}{|x-y|}\gamma_2(x,y)dxdy : \gamma_2 \in \mathcal{A}_2, \gamma_2 \to \rho \right\}, \tag{1.16}$$

where $\gamma_2 \to \rho$ means that γ_2 satisfies the equation (1.7), and the set $\mathcal{A}_2$ of admissible pair density functions is defined by the image of $\mathcal{A}$ under the map $\psi \to \gamma_2$. As pointed out in [27], 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.

Theorem 1.5 ("Semi-classical limit" for $N = 2$, Cotar-Friesecke-Klüppelberg, [27])**.** *Let $\rho : \mathbb{R}^3 \to [0,\infty]$ be a probability density such that $\sqrt{\rho} \in H^1(\mathbb{R}^3)$, $N = 2$ and $f : \mathbb{R} \to \mathbb{R}$ be the Coulomb or the repulsive harmonic interaction potentials. Then,*

$$\begin{aligned} F_\hbar^{HK}[\rho] \quad &= \min_{\psi\in H^1(\mathbb{R}^6),\psi\to\rho} \langle\psi, \big(-\tfrac{\hbar^2}{2}\Delta + f(x_2 - x_1)\big)\psi\rangle \\ &\underset{\hbar\to 0}{\to} \min_{\gamma\in\Pi_2(\mathbb{R}^6,\rho)} \int_{\mathbb{R}^{3N}} f(x_2 - x_1)d\gamma(x_1,x_2) =: F^{OT}[\rho], \end{aligned} \tag{1.17}$$

where the set $\Pi_2(\mathbb{R}^6,\rho)$ denotes the set of measures $\gamma \in \mathcal{P}(\mathbb{R}^6)$ having ρ as marginals, i.e. $(e_i)_\sharp\gamma = \rho$ where, for $i = 1, 2$, $e_i : \mathbb{R}^6 = \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ are the projection maps.

The main difficulty in proving Theorem 1.5 is that for any transport plan given by a density ψ, $\gamma = |\psi(x_1,\dots,x_N)|^2 dx_1 \dots x_N$, it may be that $\psi \notin H^1(\mathbb{R}^{3N}), \psi \notin L^2(\mathbb{R}^{3N})$ and $T[\psi] = \infty$. Moreover, smoothing the optimal γ does not work at this level, because this may change the marginals of the problem (1.17). Cotar, Friesecke & Klüppelberg developed a smoothing technique in order to deal with this problem without changing the marginals. A complete proof of the previous theorem can be found in [27].

Let us precise somewhat 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 the more abstract context of Poisson manifolds). In those specific cases, the limit $\hbar \to 0$ is called *"semi-classical"* limit, because the first and second order terms of an asymptotic expansion of the Hamiltonian operator is given by *"classical"* terms, functions of the Hamiltonian function in a symplectic manifold [36, 76].

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

At that point, it is natural to define the *DFT-Optimal Transportation ground state* E_0^{DFT-OT} [27],

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

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

$$T_{QM}[\rho] = \inf\left\{T[\psi] : \psi\in\mathcal{H}, \psi\to\rho\right\} \quad \text{and}$$

$$E^{OT}[\rho] = \inf\left\{\int_{\mathbb{R}^{3N}}\sum_{i=1}^{N}\sum_{i=1}^{N} f(x_j-x_i)d\gamma(x_1,\dots,x_N) : \gamma\in\Pi((\mathbb{R}^d)^N,\rho)\right\}, \tag{1.19}$$

where $\Pi((\mathbb{R}^d)^N,\rho)$ denotes the set of probability measures $\gamma : (\mathbb{R}^d)^N \to \mathbb{R}_+$ having ρ has marginals.

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

$$V_{ee}[\gamma_2] \geq E_{OT}[\rho], \quad \text{for every probability density } \rho\in H^1(\mathbb{R}^3).$$

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

Theorem 1.6 (Cotar-Friesecke-Klüppelberg, [27]). *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 turns out now that the central question of DFT-OT problem with Coulomb potentials is to characterize the minimizers γ of $E^{OT}[\rho]$. Those minimizers could be used as *ansatz* in order to find minimizers of the Hohenberg-Kohn $F^{HK}[\rho]$. In [82], the physicist Seidl formulated the following conjecture.

Conjecture 1.7 (Weak Seidl Conjecture [82]). *There exists a deterministic minimizer of $E^{OT}[\rho]$ with Coulomb potential/cost*

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

In other words, for Coulomb-type electron-electron interactions, Seidl conjectures[1] state that, at least for radially symmetric measures ρ, there exists an optimal measure γ for E^{OT} of form $\gamma = (Id\times T_1\times\cdots\times T_{N-1})_\sharp\rho$, with $T_1,\dots,T_{N-1}:\mathbb{R}^3\to\mathbb{R}^3$. This maps T_i are called co-motion functions and, due to the symmetries of the problem, a natural group law is required for them, namely that for every i,j there exists k such that $T_i\circ T_j = T_k$; this is satisfied if for example $T_i = T^{(i)}$ (cyclical case, see Equation (2.9)).

It is natural to ask the same kind of question for other cost functions. In the following we mention some interesting results:

(i) In the 2-electrons case, the minimizers of E_0^{DFT-OT} are well-understood [27] and there exists a map which correlates the density of a given minimization plan of $E^{OT}[\rho]$, see Theorem 3.6.

(ii) Gangbo & Święch (without being aware of Seidl) reinforces the conjecture 1.7 showing in [41] that for the attractive harmonic cost $c(x_1,\dots,x_N) = \sum_{i,j=1}^N |x_i-x_j|^2$ there exists a unique optimal plan, and it is a deterministic one.

(iii) Colombo, De Pascale & Di Marino [31] answer affirmatively (strong) Seidl conjecture in the one dimensional case ($d=1$), as we will describe in section 3.

[1]There is also a stronger version of the conjecture, the Strong Seidl Conjecture, where the author describes explicitly a possible optimal map in the radially symmetric case: for the precise statement see the conjecture 3.9 and the lemma 3.8 in section 3.

(iv) The authors proved that *"Seidl's conjecture"* is not true for the repulsive harmonic cost $c(x_1,\dots,x_N)=\sum_{i,j=1}^N -|x_i-x_j|^2$ (see theorem 4.6). Moreover, for the same cost and for a particular single particle density, there exists *"fractal-like"* optimal transport maps, i.e. an optimal map T which are not differentiable at every point. We are going to state and prove those statements, see section 4.

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 γ is a measure on the product $X\times Y$ such that γ admits μ and ν as marginals, respectively, on X and Y, i.e.*

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

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 μ and ν 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,μ) and (Y,ν) is given by the product measure $\mu\times\nu$. The symbol $\rho\to\psi$ defined in eq. (1.9) means that the measure $\gamma=|\psi(x_1,\dots,x_N)|^2dx_1\dots dx_N$ is a coupling, for instance, between ρ_1 and ρ_{k-1}. In the classical theory of 2-marginals optimal transport [3, 86] a coupling is also called a transport plan.

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

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

In that terminology, the main question is to understand when the coupling given by the DFT-Optimal Transport (1.19) is deterministic. Moreover, 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 the number of particles (marginals), d the dimension of the space where those particles live (typically $d=3$) and $I=\{1,\dots,\mathbb{N}\}$ an index set. Take $x_i:\mathbb{R}^d\to\mathbb{R}$ as the position of a single particle, then 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 $x:I\to\mathbb{R}^d$.

Let us consider the distribution of the i-th particle given by a density function $\rho_i:\mathbb{R}^d\to[0,1]$, $i\in I$ and $c:\mathbb{R}^{Nd}\times\mathbb{R}^{Nd}\to[0,\infty]$, 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

$$(\mathcal{MK})\qquad \inf_{\gamma\in\Pi(\mathbb{R}^{dN},\mu_1,\dots,\mu_N)}\int_{(\mathbb{R}^d)^N}c(x_1,\dots,x_N)d\gamma(x_1,\dots,x_N),\tag{2.1}$$

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

Notice that the existence of a minimum for (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 $\mu_1 \times \cdots \times \mu_N$ belongs to $\Pi(\mathbb{R}^{dN}, \mu_1, \dots, \mu_N)$; 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 the quantity to be minimized $\gamma \mapsto \int c\, d\gamma$ is linear with respect to μ. Hence, we can guarantee the existence of a minimum for (2.1) by imposing a very weak hypothesis on the cost function, such as lower-semicontinuity.

However, we are interested in characterizing some class of optimal transport plans or, at least, understand when that optimal coupling is *deterministic*. Thus, we would like to study when minimizers for $(\mathcal{MK})$ correspond to minimizers the multi-marginal Monge problem

$$(\mathcal{M}_N) \qquad \min_{T=(T_i)_{i=1}^N \in \mathcal{T}_N} \int_{(\mathbb{R}^d)^N} c(x, T_2(x), \dots, T_N(x)) d\mu_1(x), \tag{2.2}$$

where $\mathcal{T}_N = \big\{T = (T_i)_{i=1}^N : T_i$ *is a Borelian map such that* $T_{i\sharp}\mu_1 = \mu_i, \forall i \in \{1, \dots, N\}$ and $T_1 = Id\big\}$. 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 $T = (Id, T_2, \dots, T_N)$, we can consider a measures γ_T on $(\mathbb{R}^d)^N$ defined by $\gamma_T = (\mathrm{Id} \times T_2 \times \cdots \times T_N)_\sharp \mu$. Let $h : (\mathbb{R}^d)^N \to \mathbb{R}$ be a γ_T-measurable function, then

$$\int_{(\mathbb{R}^d)^N} h(x_1, \dots, x_N) d\gamma_T(x_1 \dots, x_N) = \int_{\mathbb{R}^d} h(x_1, T_2(x_1), \dots, T_N(x_1)) d\mu_1(x_1);$$

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

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

and so, γ_T is a transport plan. In particular, the value of the minimization over Monge-Kantorovich couplings is less than or equal to the value of the minimization over the deterministic couplings

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

Contrary to 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 (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 $i \in \{2, \dots, N\}$ whereas throughout the problem (2.1) we allow the splitting of mass in the "transportation" process.

The problem (2.1), when $N = 2$, was studied by L. Kantorovich in 1940's [47, 46], without being aware of the famous article of Monge [65]. In 1975, L. Kantorovich won the Nobel Prize together with C. Koopmans *"for their contributions to the theory of optimum allocation of resources"*.

An 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 classes of transportation plans.

In the 2-marginal setting, the two formulations (2.1) and (2.2) are proved to have the same value for general Polish Spaces by Pratelli in [74].

Theorem 2.3 (Monge equals Monge-Kantorovich in the 2-marginals case, [74])**.** *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 $c : X \times Y \to \mathbb{R} \cup \{\infty\}$ is a continuous cost, then*

$$\min\left\{ \int_{X\times Y} c(x,y) d\gamma(x,y) \ : \ \gamma \in \Pi(X \times Y, \mu, \nu) \right\} =$$

$$= \inf\left\{ \int_X c(x, T(x)) d\mu(x), \ : \ \begin{array}{l} T_\sharp \mu = \nu, \\ T : X \to Y \ \textit{Borel} \end{array} \right\} \tag{2.3}$$

where $\Pi(X \times Y)$ denotes the set of 2-marginals transport plans $\gamma \subset X \times Y$ having μ and ν 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 $\nu = (e_2, \dots, e_n)_\sharp\gamma$, for every $\gamma \in \Pi(\mathbb{R}^{dN}, \rho\mathcal{L}^d)$, and obtain the following corollary

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

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

Notice that, in general, the equivalence between (2.1) and (2.2) is not an 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 by giving an example of a particular continuous cost function where the optimal coupling γ is not deterministic, even in the 2 particles case. The nonexistence of an optimal transport maps typically happens when minimizing sequences of γ exhibit strong oscillations.

Example 2.5 (G. Carlier). Suppose μ uniformly distributed in $[0, 1]$, ν uniformly distributed in $[-1, 1]$ and the continuous cost $c : \mathbb{R}^2 \to \mathbb{R}$ 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 $\overline{\gamma} = \mu \otimes (\frac{1}{2}\delta_x + \frac{1}{2}\delta_{-x})$ since $\int c\, d\overline{\gamma} = 0$. 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 $T_\sharp\mu = \mathcal{L}^1|_A + \mathcal{L}^1|_{\tilde{A}} \neq \nu$, where $x \in \tilde{A}$ iff $-x \in [0, 1] \setminus A$, and so we obtain a contradiction since T wouldn't 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, but the same conclusion and arguments apply in the larger interactions case under the cost

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

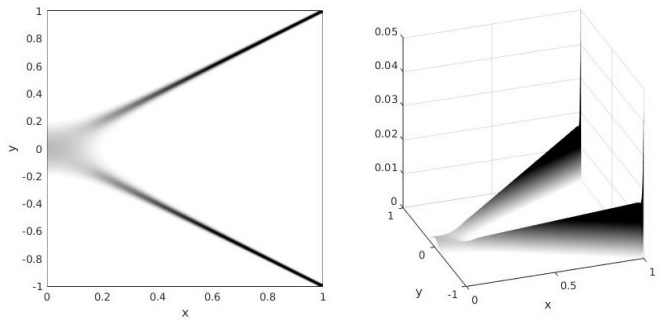


FIGURE 1. *(example 2.5) Left: support of the optimal coupling $\overline{\gamma}$. Right: optimal coupling $\overline{\gamma}$. The simulation has been performed by using the iterative method described in section 6.*

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

$$(\mathcal{K}_N) \quad \sup\left\{ \int_{\mathbb{R}^d} \sum_{i=1}^{N} u_i(x_i) d\mu_i(x_i) : \sum_{i=1}^{N} u_i(x_i) \leq c(x_1, \dots, x_N) \right\}, \tag{2.4}$$

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 (2.4) 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 interest.

Definition 2.6 (c-conjugate function)**.** *Let $c : (\mathbb{R}^d)^N \to \mathbb{R} \cup \{\infty\}$ 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-conjugates, and therefore semi-concave. So, they admit super-differentials $\overline{\partial} u_i(x_i)$ 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 $(x_j)_{j\neq i} \in \mathbb{R}^{d(N-1)}$, such that

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

Moreover, if $c : (\mathbb{R}^d)^N \to \mathbb{R}$ is differentiable and $|u_i(x_i)| < \infty$, for some $x_i \in \mathbb{R}^d$, u_i is locally Lipschitz. The well-known general result linking both (2.1) and (2.4) has been proven by Kellerer in 1984.

Theorem 2.7 (Kellerer, [48])**.** *Let $(X_1, \mu_1), \dots, (X_N, \mu_N)$ be Polish spaces equipped with Borel probability measures $\mu_1, \dots, \mu_N$. Consider $c : X_1 \dots X_N \to \mathbb{R}$ a Borel cost function and assume that $\overline{c} = sup_X c < \infty$ and $\underline{c} = \inf_X c > -\infty$. Then,*

(i) *There exists a solution γ of the Monge-Kantorovich problem* (2.1) *and a c-conjugate solution $(u_1, u_2, ..., u_N)$ of the dual problem (2.4).*

(ii) *"Duality holds",*

$$\inf \left\{ \int_{X_1 \dots X_N} c d\gamma \ : \ \gamma \in \Pi(\mu_1, \dots, \mu_N) \right\} = \sup \left\{ \sum_{i=1}^{N} \int_{X_i} u_i(x_i) d\mu_i : u_i \in \mathcal{D} \right\}, \tag{2.5}$$

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

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

(iii) *For any solution γ of* (2.1)*, any conjugate solution of* (2.4) *and γ a.e. $(x_1, \dots, x_N)$, we have*

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

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

A generalization of Kellerer's theorem due to M. Beiglboeck, C. Leonard and W. Schachermayer [6], extends this duality for more general costs, but also cannot be applied in this context.

Recently, De Pascale [30] extended the proof of duality which can be applied also for Coulomb-type costs, see Theorem 3.1. In fact, suppose $\mu = \rho(x)dx$ a probability measure on $\mathbb{R}^d$ which does not give mass to sets of cardinality less than or equal to $d-1$, then

$$\min_{\gamma \in \Pi_N(\mu)} \int_{(\mathbb{R}^d)^N} \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{1}{|x_i - x_j|} d\gamma = \max_{u \in \tilde{\mathcal{D}}} \int_{\mathbb{R}^d} N u(x) \rho(x) dx \tag{2.6}$$

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

$$\sum_{i=1}^{N} u(x_i) \leq \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{1}{|x_j - x_i|}, \quad \otimes_{i=1}^{N} \rho - \text{ almost everywhere.} \tag{2.7}$$

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

Remark 2.8. More generally, the main theorem in [30] shows that the Kantorovich duality holds, for instance, for costs 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.** Our goal in this section is to present the relation between the support of a coupling γ and optimality in the Monge-Kantorovich problem (2.1). In the following, we will just summarize key results necessary to present recent developments of optimal transportation with Coulomb and repulsive harmonic-type costs.

Roughly speaking, it is well-known in optimal transport theory with 2-marginals that, for a wide class of costs c, a coupling γ is optimal if, and only if, the support of γ is concentrated in a c-cyclically monotone set [3, 75, 80]:

Definition 2.9 (c-cyclically monotone, 2-marginal case)**.** *Let X, Y be Polish spaces, $c : X\times Y\to\mathbb{R}\cup\{\infty\}$ be a cost function and Γ be subset of the product space $X\times Y$. We say that Γ is a c-cyclically monotone set if, for any finite couples of points $\{(x_i,y_i)\}_{i=1}^{M}\subset\Gamma$,*

$$\sum_{i=1}^{M}c(x_i,y_i)\leq\sum_{i=1}^{M}c(x_i,y_{\sigma(i)}).$$

In [67], Pass suggested a possible extension of the concept of c-cyclically monotone sets for the multi-marginal case. For our proposes, it will be enough to consider this notion in $\mathbb{R}^d$.

Definition 2.10 (c-cyclically monotone set)**.** *Let $c:\mathbb{R}^{dN}\to\mathbb{R}$ be a cost function. A subset $\Gamma\subset\mathbb{R}^{dN}$ is said to be c-monotone with respect to a partition $(p,p^c), p\subset I$, 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

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

We say that $\Gamma\subset\mathbb{R}^{dN}$ is c-cyclically monotone, if it is c-monotone with respect to any partition (p,p^c), $p\subset I$.

The notion of c-cyclical monotonicity in the multi marginal setting was studied by Ghoussoub, Moameni, Maurey, Pass [42, 43, 67, 69] and successfully applied in the study of decoupling PDE systems [44]. For Coulomb costs, it was useful in order to characterize the optimal transport maps in the one-dimensional case [31].

More recently, Beiglboeck and Griessler [4], presented a new notion called *finistic optimality* inspired by the martingale optimal transport, which is equivalent, in the 2-marginals case, to the definition 2.9.

The next proposition gives a necessary condition on the support for an optimal transport plan γ. For the proof, which is based on the 2-marginal result [84], we refer the reader to Lemma 3.1.2 of [67].

Proposition 2.11 (Support of transport plan are c-cyclically monotonic)**.** *Let $c:(\mathbb{R}^d)^N\to[0,\infty]$ 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 γ is concentrated in a c-cyclically monotone set.*

2.5. **Symmetric Case.** We are going to show in this section that we can reduce our study of the Monge-Kantorovich (2.1) and the Monge problems (2.2), respectively, to 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 system of particles or, equivalently, we can suppose the couplings $\gamma \in \Pi_N(\mu)$ of those particles are cyclically symmetric.

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

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

where $\overline{\sigma}$ is the cyclical permutation

$$\overline{\sigma}(x_1,x_2,\dots,x_N) = (x_2,x_3,\dots,x_N,x_1).$$

We will denote by $\Pi_N^{sym}(\mu_1)$, the space of all symmetric probability measures on $\mathbb{R}^{dN}$ having the same first marginal μ_1.

Notice that if $(\mu_1,\dots,\mu_N)$ are the marginals of $\gamma \in \Pi^{sym}(\mathbb{R}^{dN},\mu_1)$, then all marginals $\mu_1,\dots,\mu_N$ are necessarily the same: $\forall\ f \in \mathcal{C}(\mathbb{R}^d)$,

$$\begin{aligned}\forall i \in I,\quad \int_{\mathbb{R}^d} f(x_i)d\mu_i(x_i) &= \int_{\mathbb{R}^{dN}} \tilde{f}(x_1,\dots,x_N)d\gamma(x_1,\dots,x_N)\\ &= \int_{\mathbb{R}^{dN}} \tilde{f}(\overline{\sigma}^{1-i}(x_1,\dots,x_N))d\gamma\\ &= \int_{\mathbb{R}^d} f(x_1)d\mu_i(x_1),\end{aligned}$$

where $\tilde{f}(x_1,\dots,x_N) = f(x_i)$. Briefly,

$$(e_i)_\sharp\gamma = \mu_1, \forall i \in I,$$

where e_i is the canonical projection.

As previously, we can define the multi marginal symmetric Monge-Kantorovich Problem as follows:

$$(\mathcal{MK}_{sym}) \qquad \inf_{\gamma\in\Pi_N^{sym}(\mu)} \int_{\mathbb{R}^{dN}} c(x_1,\dots,x_N)d\gamma(x_1,\dots,x_N), \tag{2.8}$$

and the associated multi marginal cyclically Monge Problem

$$(\mathcal{M}_{cycl}) \qquad \inf_{\substack{T_\sharp\mu=\mu,\\ T^{(N)}=I}} \int_{\mathbb{R}^d} c(x,T^{(1)}(x),T^{(2)}(x),\dots,T^{(N-1)}(x)), \tag{2.9}$$

Notice that, $(\mathcal{M}_{cycl})$ is related to $(\mathcal{M})$ by choosing $T_\sharp\mu = \mu$, $T_{i+1}(x) = T^{(i)}(x)$ for every $i \in \{1,\dots,N-1\}, T^{(N)} = Id$.

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

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

Proof. The fact that the infimum of (2.1) is less than or equal to (2.8) is obvious. 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

$$\overline{\gamma} = \frac{1}{N!}\sum_{\sigma\in\mathscr{S}_N} \sigma_\sharp\gamma,$$

and finally we notice that $\overline{\gamma}$ has the same cost as γ. □

We remark that prove $\inf(\mathcal{M}_{cycl}) \geq \inf(\mathcal{M}_N)$ is not obvious . However, we can avoid this problem in order to prove the equivalence between the Multi-marginal cyclic problem and Multi-marginal OT problem, by noticing that $\inf(\mathcal{M}_{sym} \leq \inf(\mathcal{M}_{cycl})$. The last part was proved by M. Colombo and S. Di Marino [24]. We give the precise statement of this theorem.

Theorem 2.14 (M. Colombo and S. Di Marino, [24]). *Let $c : (\mathbb{R}^d)^N \to \mathbb{R}$ be a continuous function and bounded from below. If μ 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.

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

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

The last result of this section is due Pass [70], which states that optimal transport plans for symmetric costs can be supported on sets of Hausdorff dimension equal or bigger than $2N-2$.

Theorem 2.16 (Pass, [70]). *Let $c : (\mathbb{R}^d)^N \to \mathbb{R}$ 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 $(r_1, ...r_N)$ the minimizers*

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

Then, there exists solutions γ for the Monge-Kantorovich symmetric problem (2.8) *whose support is at least $(2N-2)$-dimensional.*

3. Multi-marginal OT 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 Seidl in its seminal paper [82], 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 μ 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 by stating the duality theorem in this case proved by De Pascale [30], by means of Γ-convergence of finite dimensional problems for which the duality is classical.

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* (2.6) *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.

Lemma 3.2 (Corollary 3.13 in [30]). *Let μ as in Theorem 3.1; then there exists $\alpha > 0$ such that for every optimal plan γ we have* $\operatorname{supp}(\gamma) \subseteq \{(x_1, \dots, x_N) \,:\, |x_i - x_j| \geq \alpha,\ \forall\, i \neq j\}$.

Lemma 3.3. *Let μ as in Theorem 3.1 and u be a bounded maximizer in the dual problem in* (2.6)*; then u has a μ-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 (2.7) is saturated, that is $(u, \dots, u)$ is a c-conjugate N-tuple. Let us consider

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

where the 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 μ-a.e x_1, thanks to (2.7). Suppose that $w > u$ in a set of positive measure; but then $(w, u, \dots, u)$

would be a better competitor in the (not symmetric) dual problem contradicting the fact that u is a maximizer (we use the fact that the symmetric dual problem and the not symmetric 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 μ such that $|p_i - p_j| \geq d$; this number exists as long as $\sharp\,\mathrm{supp}(\mu) \geq N$. 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

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

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

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

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^∞ 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 ε, that is fixed a priori, and so by a covering argument we obtain the thesis. □

Another interesting reformulation of the Coulomb-like problem, or more generally when we have only interaction between two particles, can be found in [38] 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

$$\int_{\mathbb{R}^{Nd}} \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|}\, d\gamma = \binom{N}{2} \int_{\mathbb{R}^{2d}} \frac{1}{|x - y|}\, d(e_1, e_2)_\sharp \gamma,$$

that is, since the electron are indistinguishable, it is sufficient to look at the potential energy of a couple of electrons and then multiply it by the number of couples of electron. It is clear that $\gamma_2 = (e_1, e_2)_\sharp \gamma$ is a 2-plan whose marginal are still μ; we will say that $\eta \in \Pi_2(\mu)$ is N-representable whenever it exists $\gamma \in \Pi_N(\mu)$ such that $\eta = (e_1, e_2)_\sharp \gamma$. 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 [38] the authors propose, as a method for reducing the dimension 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 SCE functional .

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

Theorem 3.4 ([28])**.** *Let $\mu \in \mathcal{P}(\mathbb{R}^d)$. Then we have that*

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

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 a measure $\rho \in L^{4/3}(\mathbb{R}^3)$ the Lieb-Oxford bound holds [56, 58]:

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

and so the conclusion is immediate. However, in [28] the proof is completely different and relies on the fact that a measure $\gamma \in \Pi_2(\mu)$ that is N-representable for every N must be in the convex

envelope of measures 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 cases 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-particles case, in every dimension d: this result was proved in [27] by means of standard optimal transport techniques.

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, $\gamma = (Id, T)_\sharp\mu$, and $T(x) = x + \frac{\nabla\phi}{|\nabla\phi|^{3/2}}$ μ-almost everywhere, where ϕ is a Lipschitz maximizer for the dual problem.

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

$$F(x_1, x_2) = \phi(x_1) + \phi(x_2) - \frac{1}{|x_1 - x_2|} \leq 0 \qquad \text{for } \mu\otimes\mu\text{-almost every } (x_1, x_2),$$

and we know also that $F = 0$ γ-almost everywhere. But then F has a maximum on the support of γ 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 γ. 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 γ and this implies $\gamma = (Id, T)_\sharp\mu$ as we wanted to show. □

The first positive N-marginal result for the Coulomb cost is instead shown in [31] 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.

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*

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

Moreover this plan is induced by an optimal cyclical map T, that is, $\gamma_O = \frac{1}{N!}\sum_{\sigma\in\mathscr{S}_N} \sigma_\sharp\gamma_T$, where $\gamma_T = (Id, T, T^{(2)}, \dots, T^{(N-1)})_\sharp\mu$. 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 μ, and F_μ^{-1} is its lower semicontinuous left inverse.

Proof. We begin by observing that if γ is a symmetric optimal plan then a stronger statement than Proposition 2.11 holds, namely we have that for every $x, y \in \mathrm{supp}(\gamma)$:

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

where c is the Coulomb cost on the N-tuple of points $x = (x_1, \dots, x_N)$ and σ acts on the indices. The key point here is that one can show (Proposition 2.4 [31]) 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

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

or the other way around. Once this property is proven the rest of proof is rather straightforward, since we proved that the support of $(e_1, e_2)_\sharp\gamma|_D$, where $D = \{x_i \leq x_j \text{ and } y_i \leq y_j, \forall\, i < j\}$, is

monotone in the usual sense. We remark that the key property does not easily generalize to $d > 1$ since it (heavily) uses the ordered structure of $\mathbb{R}$. □

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 [31].

For the general case we thus have to assume $d > 1$ and $N > 2$: from now on we will look at the case in which the measure μ 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 μ is determined by $\mu_r = |\cdot|_\sharp\mu$. Then for every optimal plan $\gamma \in \Pi_N(\mu)$, if we consider $\gamma_r = R_\sharp\gamma$, where $R : (x_1, \dots, x_N) \mapsto (|x_1|, \dots, |x_N|)$, we have that γ_r is an optimal plan for the $1D$ multi marginal problem*

$$\min_{\eta \in \Pi_N(\mu_r)} \int_{\mathbb{R}^N} c_1(r_1, \dots, r_N)\, d\eta. \tag{3.1}$$

where c_1 is the reduced Coulomb cost

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

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

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 [83] the authors conjecture a similar structure:

Conjecture 3.9 (Strong Seidl Conjencture)**.** *Let $\mu \in \mathcal{P}(\mathbb{R}^d)$ be an absolutely continuous measure with respect to the Lebesgue measure, with radial symmetry, and let $\mu_r = |\cdot|_\sharp\mu$. 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 $S : [0, \infty) \to [0, \infty)$ be defined piecewise such that the interval A_i is sent in the interval A_{i+1} in an anti monotone way:*

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

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

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

However in the recent papers [25, 32] this conjecture is proven to be wrong, looking at radial measures μ concentrated in some thin annulus $\{1 - \delta \le |x| \le 1 + \delta\}$; in particular in [32] the $1D$ problem is proven to be equivalent, when $\delta \to 0$, to the repulsive harmonic one, for which high non-uniqueness holds and the Seidl map is not optimal (see the following section). However in [25] 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 μ for the strong Seidl conjecture to hold.

4. Multi-marginal OT with repulsive Harmonic Cost

This section is devoted to the study of the repulsive harmonic cost. In DFT-OT problem (see section 1), we replace the electron-electron Coulomb interaction by a weak force which decreases 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 ∞ in the Coulomb case.

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

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

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) seem to have no particular relevance in physics because, as we will see in the examples 4.10, 4.11 and 4.12 below, certain optimal γ of (4.1) allow 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

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

but this last additive term depends only on the marginals and not on the specific plan γ. The cost c is very particular since it 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 μ_i have finite second moments; then existence follows immediately from the equality

$$\operatorname*{argmin}_{\gamma} \int -\sum_{i=1}^{N} \sum_{j=i+1}^{N} |x_j - x_i|^2 d\gamma = \operatorname*{argmin}_{\gamma} \int |x_1 + \dots + x_N|^2 d\gamma.$$

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 $x_1 + \dots + x_N = const$, 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 4.9, 4.10, 4.11, 4.12 and 4.14 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 μ is an absolutely continuous with respect to the Lebesgue measure in $\mathbb{R}^d$ and ν has no atoms. Then, there exists a unique optimal transport map $T : \mathbb{R}^d \to \mathbb{R}^d$ for the problem*

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

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

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

$$C + \inf_{T_\sharp \mu = \nu} \int -|x - T(x)|^2 d\mu(x) = \inf_{G_\sharp \mu = \tilde{\nu}} \int |x - G(x)|^2 d\mu(x)$$

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

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 Ω_f, Ω_g and assume there exits a constant $\lambda > 0$ such that $\lambda \leq f, g \leq 1/\lambda$. Then, T is a $C^{1,\alpha}$ function inside Ω_f [13, 14].

Example 4.2 (2 marginals case, uniform measure in the d-dimensional cube). Suppose that $\mu = \nu = \mathcal{L}\big|_{[0,1]^d}$. In this case, we can verify easily that the optimal map $T : [0,1]^d \to [0,1]^d$ 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 4.8 below). Before stating the theorem we will prove a lemma that characterizes many optimal plans:

Lemma 4.3. *Let $\{\mu_i\}_{i=1}^N$ be probability measures on $\mathbb{R}^d$ and $h : \mathbb{R}^d \to \mathbb{R}$ be a strictly convex function and suppose $c : ([0,1]^d)^N \to \mathbb{R}$ a cost function of the form $c(x_1, \dots, x_N) = h(x_1 + \cdots + x_N)$. Then, if there exists a plan $\gamma \in \Pi(\mu_1, \dots, \mu_N)$ 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 $\mathrm{supp}(\tilde{\gamma}) \subset \{x_1 + \dots + x_N = k\}$ is a necessary and sufficient condition for $\tilde{\gamma}$ to be optimal. In this case we will say that γ 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 μ_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 γ is μ_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 γ 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 $x_1 + \dots + x_N = k$. □

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:

Remark 4.4. Let $N = 3$ and $\mu_i = \mu$ for every $i = 1, 2, 3$ with $\mu = \nu_1 + \nu_2$ with $\nu_1 = (-Id)_\sharp \nu_2$ and ν_1 concentrated on $[2, 3]$. Now it is clear that the barycenter of μ is 0 but for every 3 points x_1, x_2, x_3 in the support of μ 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 there is no 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 $\mathrm{supp}(\gamma) \cap H_0 = \emptyset$.

Remark 4.5. In the case $N = 2$ the condition for which there exists an admissible plan γ concentrated on an hyperplane of the form $x_1 + x_2 = k$ implies that the two measures μ and $\tilde{\nu} = (-Id)_\sharp \nu$ 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 μ is centrally symmetric about its barycenter and in the context of density of 2 electrons around a nucleus this seems a fairly natural condition.

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

$$\min_{\gamma \in \Pi_N(\mu)} \int c d\gamma = \min_{\substack{T_\sharp \mu = \mu, \\ T^{(N)} = I}} \int c(x, T(x), \dots, T^{(N-1)}(x)) d\mu$$

Moreover, T is not differentiable at any point and it is a fractal map, meaning that it is the unique fixed point of a "self-similar" linear transformation $\mathcal{F}$ acting on $L^\infty([0,1]^d; [0,1]^d)$.

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 given by $S(z) = \sum_{k=1}^\infty \frac{\mathcal{S}(a_k)}{N^k}$, where $\mathcal{S}$ is the permutation of N symbols such that $\mathcal{S}(i) = i + 1, \forall \{i = 0, 1, \dots, N-2\}$ and $\mathcal{S}(N-1) = 0$. A straightforward computation shows that

$$z + \sum_{i=1}^N S^i(z) = \frac{N}{2} \tag{4.2}$$

Let $T : \mathbb{R}^d \to \mathbb{R}^d$ 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 exist functions $S_k : [0,1] \to [0,1]$ 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: $T_k(x) = x - (N-1) \cdot N^{-k}$ 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

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

Moreover, it is easy to see that

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

Hence

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

Now, it remains to show that T is optimal. But this is true thanks to the fact that (4.2) implies that the plan induced by T satisfies the hypothesis of Lemma 4.3.

It is clear that we can reduce to prove the non-differentiability and the fractal properties only for the map S. The non-differentiability comes from the fact that for each $z \in [0,1]$, we have that the base-N representation $z = \sum_{k=1}^{\infty} \frac{a_k}{N^k}$ is such that a_k can't be definitively $N-1$. Now it is sufficient to choose those k_j such that $a_{k_j} \neq N-1$ and consider the numbers $z_j = z + (N - 1 - a_{k_j})/N^{k_j}$ and $z'_j = z \pm 1/N^{k_j}$ (depending on whether $a_{k_j} = 0$ or not) which have the same digits as z apart from the j_k-th digit. Then it is straightforward to see that $S(z_j) - S(z) = -(a_{k_j} + 1)/N^{k_j}$ while $S(z'_j) - S(z) = z'_j - z$; in particular we have $\frac{S(z'_j) - S(z)}{z'_j - z} = 1$ while $\frac{S(z_j) - S(z)}{z_j - z} \leq -\frac{1}{N}$ and so, letting $j \to \infty$, we get that S is not differentiable at the point z.

As for the fractal property: we consider the transformation

$$\mathcal{F}(g)(x) = \begin{cases} \frac{1}{N} g(Nx - i) + \frac{i+1}{N} & \text{for } \frac{i}{N} \leq x < \frac{i+1}{N}, \text{ with } i = 0, \dots, N-2 \\ \frac{1}{N} g(Nx - N + 1) & \text{for } \frac{N-1}{N} \leq x < 1. \end{cases}$$

In order to see what the construction is doing we imagine to divide $[0,1]^2$ in a grid $N \times N$ of squares and then we are putting scaled copies (by a factor N) of the original function in the above diagonal squares and in the rightmost bottom one. Now it is easy to see that S is the unique[2] fixed point of $\mathcal{F}$, and this gives the property of self-similarity. □

Remark 4.7. Another proof of Theorem 4.6 can be done noticing that the transformation $\mathcal{F}$ (the one defined in the proof) acts as a $1/N$-contraction in the space of measure preserving L^∞-bijections from $[0,1)$ to $[0,1)$.

Corollary 4.8. *Let $\mu = \mathcal{L}^d\big|_{[0,1]^d}$ be the uniform measure on the d-dimensional cube $[0,1]^d$ in $\mathbb{R}^d$ and suppose $c : ([0,1]^d)^N \to \mathbb{R}$ the N-dimensional repulsive harmonic cost*

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

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

$$\min_{\gamma \in \Pi_N(\mu)} \int c d\gamma = \min_{\substack{T_\sharp \mu = \mu, \\ T^{(N)} = I}} \int c(x, T(x), T^{(2)}(x), \dots, T^{(N-1)}(x)) d\mu.$$

Moreover, T is not differentiable at any point.

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

[2]We notice that $\mathcal{F}$ is a contraction: $\|\mathcal{F}(g) - \mathcal{F}(g')\|_\infty \leq \frac{1}{N} \|g - g'\|_\infty$

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 4.8 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 examples 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 $(Id, T_1, T_2, \dots, T_{N-1})_\sharp \mu_1$ is an optimal plan but $T_1 \circ T_1 \neq T_i$ for any index $i \in \{1, \dots, N-1\}$.

Example 4.9 (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 $\mu_3 = \frac{1}{2\pi^{3/2}} \exp(-(x_1^2 + x_2^2 + x_3^2))$. In this case, we can verify that the couple (T, S) of maps $T, S : \mathbb{R}^d \to \mathbb{R}^d$, $T(x) = x$ a.e. and $S(x) = -2x$ a.e. is admissible and it is optimal since $x + T(x) + S(x) = 0$.

Example 4.10 (2N particles on S^1,). Suppose $\mu_1, \dots, \mu_{2N}$ uniform probability measures on the circle S^1. The rotation map $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ 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.

Example 4.11 (2N particles, breathing map). Suppose $\mu_1, \dots, \mu_{2N}$ uniform probability measures on $S^2 \subset \mathbb{R}^3$. Consider the vector v, $A : \mathbb{R}^3 \to \mathbb{R}^3$ the antipodal map $A(x) = -x$ and $R_\theta^v : \mathbb{R}^3 \to \mathbb{R}^3$ 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 $\gamma = (x, T, T^{(2)}, \dots, T^{(N-1)})_\sharp \mu_1$ is called *"breathing" solution* [83]: the coupling γ 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^∞ function and it is a gradient of the convex-concave function $\phi(x_1, x_2, \dots, x_{2N-1}, x_{2N}) = \frac{1}{2}(x_1^2 + x_3^2 + \dots + x_{2N-1}^2) - \frac{1}{2}(x_2^2 + x_4^2 + \dots + x_{2N}^2)$.

Example 4.12 (2N particles in $\mathbb{R}^d$, symmetric ρ). In this case, on can consider the maps $T(x) = x$ a.e. and $S(x) = -x$ a.e. which are 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

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

Example 4.13 (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

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

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

Example 4.14 (Optimal Transport diffuse plan). Let us consider the same problem as in Example 4.13. 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 $H = \{x + y + z = 0\} \cap \{|x| \leq 1, |y| \leq 1, |z| \leq 1\}$ and f is a function to be chosen later. This is a symmetric flat optimal plan; now we compute the marginals. Since it is clear that γ 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

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

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 $h:[0,1]\to[0,\infty)$ can be represented in this way: in fact it is sufficient to choose $f(x)=\frac{1}{\sqrt{3}}\Big(\frac{h(x)}{x}-2x\int_x^1\frac{h(t)}{t^3}\,dt\Big)$.

Example 4.15 (A Counterexample on the uniqueness $N>3$). As mentioned in theorem 2.16 and in [38], it was already understood by Pass that in these high dimensional cases, the solution of repulsive costs may also be non unique, as opposite to 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 shown in the examples 4.9–4.14. In most of the cases, the non-uniqueness seems to be given by the symmetries of the problem, but in Example 4.14 and Corollary 4.8 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 than the minimal one.

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

Proposition 4.16. *Let $\mu=\mathcal{L}|_{[0,1]}$ and $N\geq 3$ be an odd number. Then, the infimum*

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

is not attained by a map T which is differentiable almost everywhere.

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 $T_\sharp\mu=\mu$. In particular, since the Lusin property holds true for $T^{(i)}$ for every $i=1,\dots,N$ and $T_\sharp\mu=\mu$, 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 μ is N-flat we have that the condition on T in order to be an optimal cyclical map is $x+T(x)+T^{(2)}(x)+\ldots+T^{(N-1)}(x)=N/2$; now we can differentiate this identity and so we will get

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

But this is absurd since on the left hand side we have an odd number of ± 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 $\mu=\frac{1}{2}\mathcal{L}^1|_{[-1,1]}$, we have both diffuse optimal plan and a cyclical optimal map.

An interesting open problem is whether for any N-flat measure μ, 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 [17], where Carlier and Nazaret consider the following optimal transport problems for the determinant:

$$(\mathcal{MK}_{Det})\qquad \sup_{\gamma\in\Pi((\mathbb{R}^d)^d,\mu_1,\dots,\mu_d)}\int_{(\mathbb{R}^d)^d}\det(x_1,\dots,x_d)d\gamma(x_1,\dots,x_d) \tag{5.1}$$

and

$$(\mathcal{MK}_{|Det|}) \qquad \sup_{\gamma\in\Pi((\mathbb{R}^d)^d,\mu_1,\dots,\mu_d)} \int_{(\mathbb{R}^d)^d} |\det(x_1,\dots,x_d)| d\gamma(x_1,\dots,x_d), \tag{5.2}$$

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 there exist $p_1,\dots,p_d \in [1,\infty[$ such that

$$\sum_{i=1}^d \frac{1}{p_i} = 1, \quad \text{and} \quad \sum_{i=1}^d \int_{\mathbb{R}^d} \frac{|x_i|^{p_i}}{p_i} d\mu(x_i) < +\infty.$$

Notice that for this particular cost the problem (2.1) makes sense only when $N = d$. In the following, we will focus on problem (5.1) and exhibit explicit minimizers γ 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 marginals $\mu_1,\dots,\mu_d$ there exist two symmetric probability measures $\mu_i,\mu_j, i\neq j$, *i.e.* $\mu_i = (\text{-Id})_\sharp\mu_i$ and $\mu_j = (\text{-Id})_\sharp\mu_j$, then any solution γ of (5.1) satisfies $\det(x_1,\dots,x_d) \geq 0$ γ-almost everywhere and so solves also (5.2) (Proposition 6, [17]).

Similarly to the Gangbo-Święch cost [41], 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, $\det(x_1,x_2) = \langle x_1, Rx_2\rangle$, where $R:\mathbb{R}^2\to\mathbb{R}^2$ is the rotation of angle $-\pi/2$. Hence, since μ_1 and μ_2 have finite second moments, up to a change of variable $\tilde{x}_2 = Rx_2$, the problem $(\mathcal{MK}_{Det})$ in (5.1) is equivalent to the classical Brenier's optimal transportation problem:

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

where $C = 1/2(\int |x_1|^2\, d\mu_1 + \int |x_2|^2\, d\mu_2)$ and $\tilde\mu_2 = R_\sharp\mu_2$.

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

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

In [17], the authors provide a useful characterization of optimal transport plans through the potentials u_i, given by theorem 5.1 . In addition, by means of a standard convexification trick we obtain regularity results on the Kantorovich potentials.

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 $u_i:\mathbb{R}^d\to\mathbb{R}\cup\{\infty\}$ such that for all $i\in\{1,\dots,d\}$,*

$$\sum_{i=1}^d u_i(x_i) \leq \sum_{i=1}^d u_i^*((-1)^{i+1}\bigwedge_{i\neq j} x_j), \qquad on \quad (\mathbb{R}^d)^d;$$

$$\sum_{i=1}^d u_i(x_i) = \sum_{i=1}^d u_i^*((-1)^{i+1}\bigwedge_{i\neq j} x_j), \quad \gamma - almost\ everywhere;$$

$$(-1)^{i+1}\bigwedge_{i\neq j} x_j \in \partial u_i(x_i), \quad \gamma - almost\ everywhere.$$

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 solution.

We illustrate the theorem (5.1) and explain how to construct a particular optimal γ for (5.1) by an example in the three marginals case. 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) have a natural geometric interpretation: what is the best way to place three random vectors x, y, z, distributed by probability measures μ_1, μ_2, μ_3 on the sphere, such that the simplex generated by those three vectors (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 γ, we have

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

Applying the theorem (5.1), we get

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

and,

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

It follows from (5.4), 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 $\mu_i = \mathcal{L}^3_B,\ i = 1, 2, 3$, the 3-dimensional Lebesgue measure in the ball $B \subset \mathbb{R}^3$. The following coupling γ^*

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

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 γ^* 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 $\gamma^* \in \Pi_3(\mathcal{L}^3_B)$.

Some comments on the radially symmetric d-marginals case: In [17], for d radially symmetric probability measures, the authors exhibit explicit optimal couplings γ^*. 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 γ^* is contained in the set of an orthogonal basis. The second observation is to notice that in the support of γ^* we have $H_i(|x_1|) = |x_i|$, where H_i is the unique monotone increasing map such that $(H_i^*)_\sharp \mu_1 = \mu_i$, 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 ϕ_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 γ^* in (5.5) or, more generally, the optimal coupling in the d-marginal case (see theorem 4 in [17]), 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, γ^* 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 an 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, concerning the existence of trivial even dimensional solutions for the Monge problem in (5.1). **We expect the existence of non-regular optimal transport map also to this case.**

Example 5.3 (Carlier & Nazaret, [17]). *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 ≥ 4. For instance, suppose $c(x_1, x_2, x_3, x_4) = \det(x_1, x_2, x_3, x_4)$, define transport maps $T_1, T_2, T_3 : B \to \mathbb{R}^4$ by, for $x = (x_1, x_2, x_3, x_4) \in B$

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

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

6. 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 [18] 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 rencently a numerical method based on an entropic regularization has been developed and it has been applied to various optimal transport problem in [9, 10, 29]. Let us mention that the Coulomb cost has been treated numerically in various paper as in [63, 45], where the solutions are based on the analytical form of transport maps given in [45], in [64], where parameterized functional forms of the Kantorovich potential are used, in [20, 19], where a linear programming approach has been developed to solve the 1-dimensional problem, and, as already mentioned, in [10]. In this section we focus on the regularized method proposed in [9, 10, 29] and we, finally, present some numerical experiments for the costs studied above.

6.1. The regularized problem and the Iterative Proportional Fitting Procedure.

Let us consider the problem

$$(\mathcal{MK}) \qquad \inf_{\gamma \in \Pi(\mathbb{R}^{dN}, \mu_1, \dots, \mu_N)} \int_{\mathbb{R}^{dN}} c(x_1, \dots, x_N) \gamma(x_1, \dots, x_N) dx_1, \dots, dx_N, \tag{6.1}$$

where N is the number of marginals μ_i which are probability distributions over $\mathbb{R}^d$, $c(x_1, ..., x_N)$ is the cost function, γ the coupling, is the probability distribution over $\mathbb{R}^{dN}$

Remark 6.1. From now on the marginals μ_i and the coupling γ are densities and when the optimal coupling γ 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 kth marginal as $\{x_{j_k}\}_{j_k = 1, \cdots, M_d}$. If the densities μ_k are approximated by $\mu_k \approx \sum_{j_k} \mu_{j_k} \delta_{x_{j_k}}$, we get

$$\min_{\gamma \in \Diamond\langle} \sum_{j_1, \cdots j_N} c_{j_1, \cdots, j_N} \gamma_{j_1, \cdots, j_N}, \tag{6.2}$$

where Π_k is the discretization of Π, $c_{j_1, \cdots, j_N} = c(x_{j_1}, \cdots, x_{j_N})$ and the coupling support for each coordinate is restricted to the points $\{x_{j_k}\}_{j_k = 1, \cdots, M_d}$ thus becoming a $(M_d)^N$ matrix again

denoted γ with elements $\gamma_{j_1,\cdots,j_N}$. The marginal constraints $\mathcal{C}_i$ (such that $\Pi_k = \bigcap_{k=1}^N \mathcal{C}_k$) becomes

$$\mathcal{C}_k := \Big\{\gamma \in \mathbb{R}_+^{(M_d)^N} : \sum_{j_1,\dots,j_{k-1},j_{k+1},\dots,j_N} \gamma_{j_1,\dots,j_N} = \mu_{j_k},\, \forall j_k = 1,\cdots,M_d\Big\}. \tag{6.3}$$

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

$$\begin{aligned} &\max_{u_{j_k}} \sum_{k=1}^N \sum_{j=1}^M u_{j_k}\mu_{j_k} \\ &s.t. \sum_{k=1}^N u_{j_k} \le c_{j_1\cdots j_N} \quad \forall\ j_k = 1,\cdots,M_d, \end{aligned} \tag{6.4}$$

where $u_{j_k} = u_k(x_{j_k})$ is the kth 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 μ_{j_k} are equal ($\mu_{j_k} = \mu_j$, $\forall\ k \in \{1,\cdots,N\}$). Thus the dual problem can be re-written in a more convenient way (but still computationally unfeasible for many marginals)

$$\begin{aligned} &\max_{u_j} \sum_{j=1}^M N u_j \mu_j \\ &s.t. \sum_{k=1}^N u_{j_k} \le c_{j_1\cdots j_N} \quad \forall j_k = 1,\cdots,M_d, \end{aligned} \tag{6.5}$$

where $u_j = u_{j_k} = u(x_{j_k})$. 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 [81] and, as mentioned above, it has been recently introduced in many applications involving optimal transport [9, 10, 29, 40]. Thus, we consider the following discrete regularized problem

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

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

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

and $\mathcal{C}$ is the intersection of the set associated to the marginal constraints (we remark that the entropy is a penalization of the non-negative contraint on γ). After elementary computations, we can re-write the problem as

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

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

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

As explained in section 2, when the transport plan γ 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 matrix γ 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 γ^ϵ. Moreover the solution γ^ϵ can be obtained through elementary operations. The regularized solutions γ^ϵ then converge to $\gamma^\star$ (see figure 2), the solution of (6.1) with minimal entropy, as $\epsilon \to 0$ (see [26] for a detailed asymptotic analysis and the proof of exponential convergence).

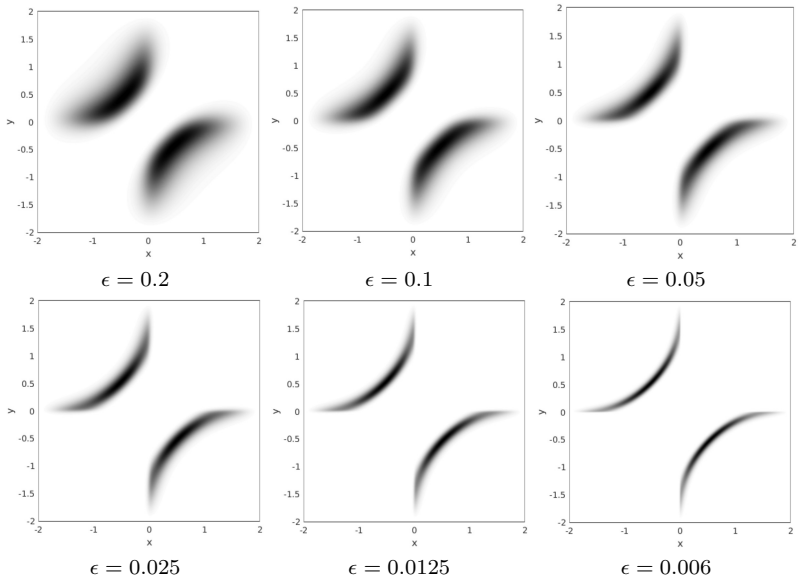


FIGURE 2. *Support of the coupling γ^ϵ for the Coulomb cost and $\mu_1 = \mu_2 = (1 + cos(\pi x/2))/2$. The simulation has been performed on a discretization of $[-2, 2]$ with $M_d = 1000$.*

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

$$\min_{\gamma \in \mathcal{C}} \sum_{i,j} c_{ij}\gamma_{ij}. \tag{6.10}$$

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

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

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

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

Thus, we can now define the following iterative method

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

Remark 6.3. In [77] 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 [78]) proved, in the continous measure 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 cumbersone to write. It leads to a problem set on N M_d-dimensional vectors $a_{j,i_{(\cdot)}}, \quad j = 1, \cdots, N, \quad i_{(\cdot)} = 1, \cdots, M_d$. Each update takes the form

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

Example 6.5 (IPFP and 3 marginals). In order to clarify the extension of the IPFP to the multi-marginal case , we consider 3 marginals and we write down the updates of the algorithm

$$a^{n+1}_{1,i_1} = \frac{\rho_{i_1}}{\sum_{i_2,i_3} \bar{\gamma}_{i_1,i_2,i_3}\, a^n_{2,i_2}\, a^n_{3,i_3}},$$

$$a^{n+1}_{2,i_2} = \frac{\rho_{i_2}}{\sum_{i_1,i_3} \bar{\gamma}_{i_1,i_2,i_3}\, a^{n+1}_{1,i_1}\, a^n_{3,i_3}},$$

$$a^{n+1}_{3,i_3} = \frac{\rho_{i_3}}{\sum_{i_1,i_2} \bar{\gamma}_{i_1,i_2,i_3}\, a^{n+1}_{1,i_1}\, a^{n+1}_{2,i_2}}.$$

In the following sections we present some numerical results obtained by using the IPFP.

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 case where the N marginals are equal to a density μ (we work in a *DFT* framework so the marginals rapresent the electrons which are indistinguishable). We recall that if we split μ into N $\tilde{\mu}_i$ with equal mass ($\int \tilde{\mu}_i(x)dx = \frac{1}{N}\int \mu(x)dx$), then we expect an optimal plan induced by a cyclical map such that $T_\sharp \tilde{\mu}_i = \tilde{\mu}_{i+1}$ $i = 1, \cdots, N-1$ and $T_\sharp \tilde{\mu}_N = \tilde{\mu}_1$.

The simulations in figure 3 are all performed on a discretization of $[-5, 5]$ with $M_d = 200$, with marginals $\mu_i = \mu(x) = \frac{N}{10}(1 + cos(\frac{\pi}{5}x)) \quad i = 1, \cdots, N$ and $\epsilon = 0.02$. If we focus on the support of the coupling $\tilde{\gamma}_{12}(x, y) = (e_1, e_2)\gamma(x, y, z)$ we can notice that the numerical solution correctly reproduces the prescribed behavior: the transport plan is induced by a cyclical optimal map (see section 3). We refer the reader to [10] and [32] for examples in higher dimension.

Remark 6.6. Theorem 3.7 actually works also for other costs functions as $c(x_1, \cdots, x_N) = \sum_{i<j}^N -log(|x_i - x_j|)$. 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 as for 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 as the corresponding case for Coulomb (first row of figure 3). The simulation is performed on a discretizetion $[-5, 5]$ with $M_d = 200$, with marginals $\mu(x) = \frac{N}{10}(1 + cos(\frac{\pi}{5}x)) \quad N = 3$ and $\epsilon = 0.01$.

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 $c(x_1, \cdots, x_N) = |x_1 + \cdots + x_N|^2$, so for the discrete problem (6.2) we take $c_{j_1 \cdots j_N} = |x_{j_1} + \cdots + x_{j_N}|^2$. Let us first 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 $\epsilon = 0.005$.

The multi-marginals case is more delicate to treat: the original OT problem does not admit a unique solution whereas 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 4.13. In this case we have all marginals equal to $\mu = \frac{1}{2}\mathcal{L}|_{[-1,1]}$ and we can find a deterministic coupling given by $\gamma(x, y, z) = \mu(x)\delta(y - T(x))\delta(z - S(x))$ for the maps $T(x)$ and $S(x)$). In order to select this coupling, the idea is to modify the cost function by adding a penalization term $p(x, y) = \frac{\tau}{|x-y|}$ which makes μ_1 and μ_2 be *as far as possible* (if we consider μ_1 and μ_2 as particles). Thus, the discretized cost now reads as

$$c_{j_1,j_2,j_3} = |x_{j_1} + x_{j_2} + x_{j_3}|^2 + \frac{\tau}{|x_{j_1} - x_{j_2}|}. \tag{6.14}$$

Then we expect, as shown in figure 6, that the projection $\tilde{\gamma}_{12}(x, y) = (e_1, e_2)\gamma(x, y, z)$ and $\tilde{\gamma}_{13}(x, z) = (e_1, e_3)\gamma(x, y, z)$ of the computed coupling are induced by map T and S respectively. 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 penalization. In this case we expect a diffuse plan since, if we forget about the marginal constraint, it

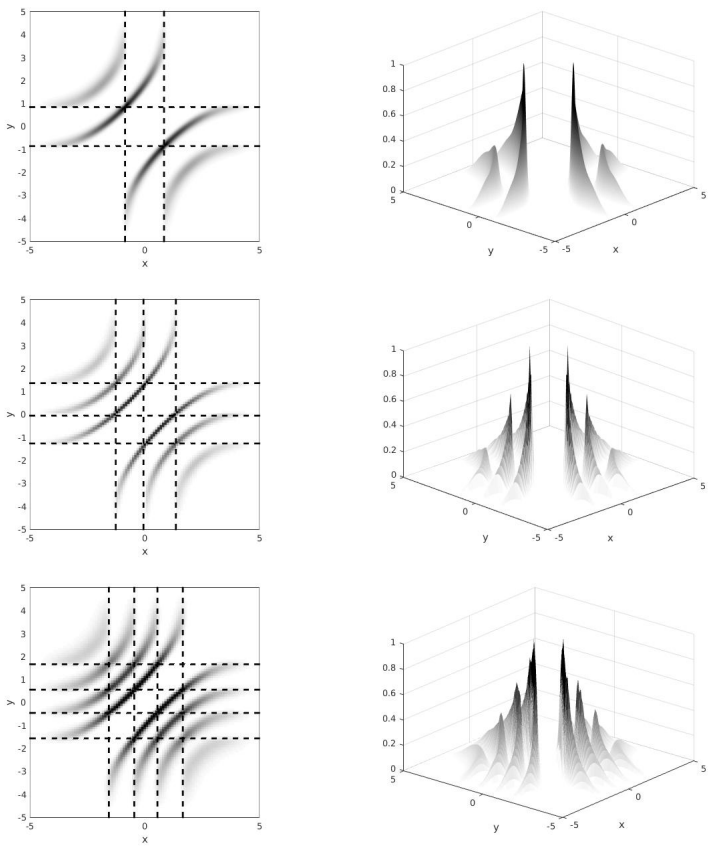


FIGURE 3. *On the Left: support of the optimal coupling $\tilde{\gamma}_{12}(x,y)$, on the Right: graph of the optimal coupling $\tilde{\gamma}_{12}(x,y)$, in the cases $N=3$ (1st Row), $N=4$ (2nd Row), $N=5$ (3rd Row). The dashed lines delimit the supports of $\tilde{\mu}_i$, with $i=1,\cdots,N$.*

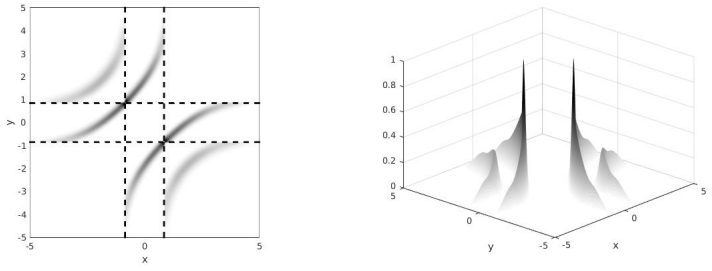


FIGURE 4. *(logarithmic cost) Left: support of the optimal coupling $\tilde{\gamma}_{12}(x,y)$ for $N=3$. Right: support of the optimal coupling $\tilde{\gamma}_{12}(x,y)$ for $N=3$. The dashed lines delimit the intervals where $\tilde{\mu}_i$, with $i=1,\cdots,3$, are defined.*

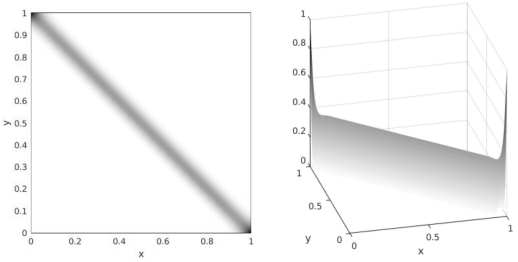


FIGURE 5. *Left: support of the optimal coupling. Right: optimal coupling.*

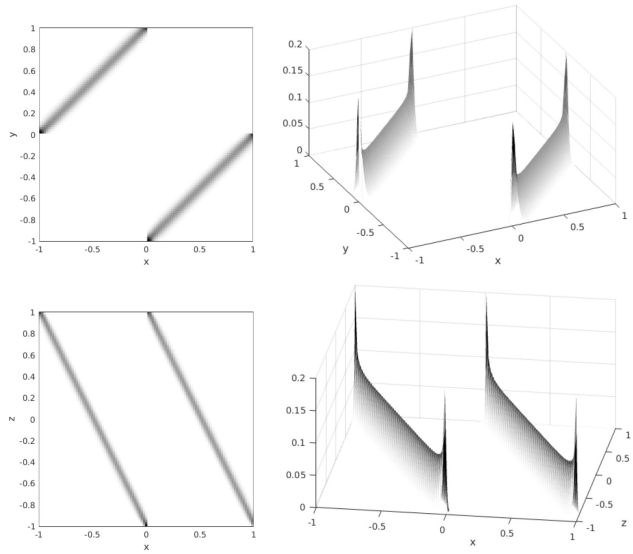


FIGURE 6. *Top-Left: support of the optimal coupling $\tilde{\gamma}_{12}(x,y)$. Top-Right: optimal coupling $\tilde{\gamma}_{12}(x,y)$. Bottom-Left: support of the optimal coupling $\tilde{\gamma}_{13}(x,z)$. Bottom-Right: optimal coupling $\tilde{\gamma}_{13}(x,z)$.*

is simple to show that the second order Γ-limit of the regularized problem as $\epsilon \to 0$ is the relative entropy of γ with respect to $\mathcal{H}^2|_{\{x+y+z=0\}}$; we expect that the Γ-limit is the same also when we add the marginal constraint.

In figure 7 we present the projection $\tilde{\gamma}_{12}(x,y) = (e_1,e_2)\gamma(x,y,z)$ of the computed coupling (it is enough to visualize only this projection because of the symmetries of the problem) and, as expected, we observe a diffuse plan. . 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 $\mathbb{R}^3$ (as the cost function now is $c(x_1,\cdots,x_N) = det(x_1,\cdots,x_N)$ with $x_i \in \mathbb{R}^N$). However if we take radially symmetric densities μ_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 unknowns are the relations between the norms $r_i = \|x_i\|$ of each vector. These relations can be obtained by solving the following problem

$$\min_{\gamma\in\Pi(\lambda_1(r),\cdots,\lambda_N)} -\int (\prod_{i=1}^{N} r_i)\gamma(r_1,\cdots,r_N)dr_1\cdots dr_N, \tag{6.15}$$

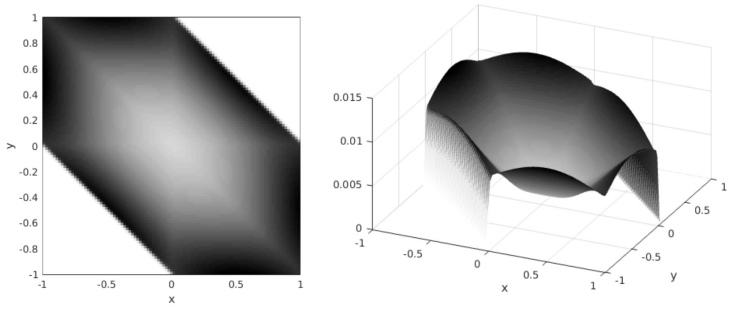


FIGURE 7. *Left: support of the optimal coupling* $\tilde{\gamma}_{12}(x,y)$. *Right: optimal coupling* $\tilde{\gamma}_{12}(x,y)$.

where $\lambda_i(r) = \sigma(r)\mu_i(\mathbf{r})$ (e.g. $N = 3$ 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 [17]). The discretized cost now reads $c_{j_1\cdots j_N} = \prod_{k=1}^{N} r_{j_k}$. Let us consider the 3-marginals case and all densities λ_i equal to the uniform density on the ball $\mathcal{B}_{0.5} = \{x \in \mathbb{R}^3 | \|x\| \leq 0.5\}$. As proved in [17], 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 $\lambda_i(r) = 4\pi r^2 e^{-4r}$

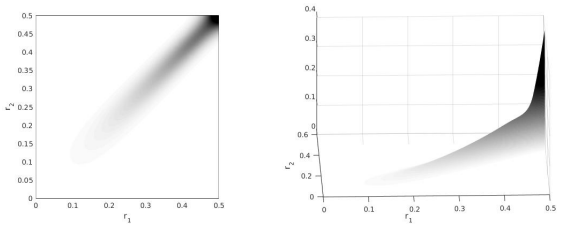


FIGURE 8. *(Uniform density) Left: support of the optimal coupling* $\tilde{\gamma}_{12}(r_1, r_2)$. *Right: optimal coupling* $\tilde{\gamma}_{12}(r_1, r_2)$.

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

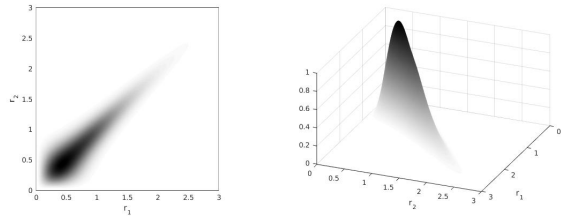
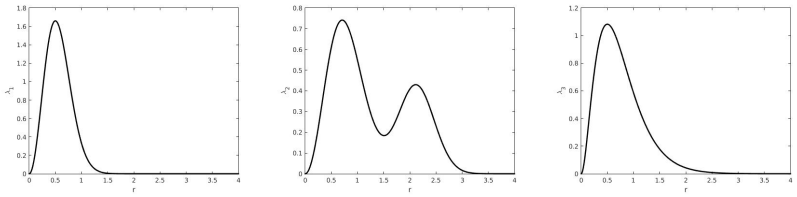


FIGURE 9. *(Exponential density)Left: support of the optimal coupling* $\tilde{\gamma}_{12}(r_1, r_2)$. *Right: optimal coupling* $\tilde{\gamma}_{12}(r_1, r_2)$.

We, finally, present a simulation (see figure 11) with three different marginals λ_i (see figure 10). As one can observe we obtain that the projection of the optimal coupling are concentrated on the graph of a map which rearranges the densities in a monotone way. The simulation has been performed on a discretization of $[0, 4]$ with $M_d = 300$ and $\epsilon = 0.07$.



λ_1 λ_2 λ_3

FIGURE 10. *Densities λ_1, λ_2 and λ_3.*

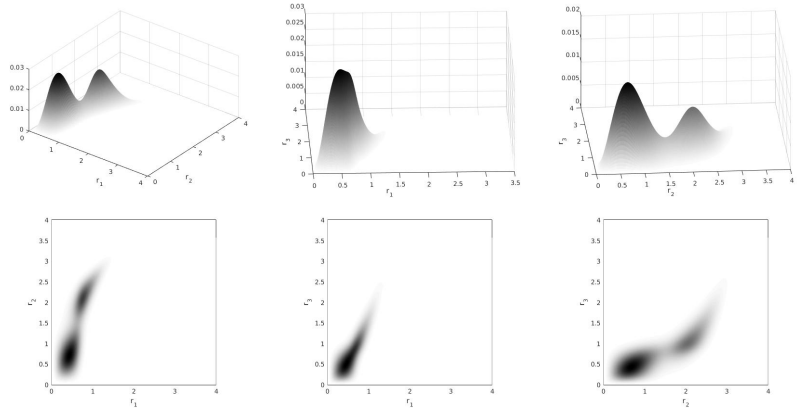


FIGURE 11. *Top-Left: optimal coupling $\tilde{\gamma}_{12}(r_1, r_2)$. Top-Center: optimal coupling $\tilde{\gamma}_{13}(r_1, r_3)$. Top-Right: optimal coupling $\tilde{\gamma}_{23}(r_2, r_3)$. Bottom: supports of those couplings.*

7. Conclusion

In this survey we presented a general famework for the multi-marginal optimal transportation with repulsive cost. We also gave some new results which highlight the importance of the multi-marginal approach. As mentioned there are still a lot of open questions about the solutions for the problems presented and we hope to treat them in future works.

Acknowledgements

We would like to thank Paola Gori-Giorgi for many helpful and stimulating discussions as well as Jean-David Benamou and Guillaume Carlier for a critical reading of the manuscript.

Simone Di Marino and Luca Nenna gratefully acknowledge the support of the ANR, through the project ISOTACE (ANR-12-MONU-0013). Luca Nenna also ackwoledges INRIA through the "action exploratoire" MOKAPLAN.

Augusto Gerolin is member of the *Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA)* of the *Instituto Nazionale di Alta Matematica (INdAM)*. He also acknowledges and thanks CNPq for the financial support of his Ph.D, when this work was done.

References

[1] I Abraham, R Abraham, M Bergounioux, and G Carlier. Tomographic reconstruction from a few views: a multi-marginal optimal transport approach. *Preprint Hal-01065981*, 2014.

[2] M Agueh and G Carlier. Barycenters in the {W}asserstein space. *SIAM J. on Mathematical Analysis*, 43(2):904–924, 2011.

[3] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2006.

[4] Mathias Beiglböck and Claus Griessler. An optimality principle with applications in optimal transport. *arXiv preprint arXiv:1404.7054*, 2014.
[5] Mathias Beiglböck, Pierre Henry-Labordère, and Friedrich Penkner. Model-independent bounds for option prices—a mass transport approach. *Finance and Stochastics*, 17(3):477–501, 2013.
[6] Mathias Beiglböck, Christian Léonard, and Walter Schachermayer. A general duality theorem for the monge–kantorovich transport problem. *arXiv preprint arXiv:0911.4347*, 2009.
[7] Mathias Beiglboeck, Alexander Cox, and Martin Huesmann. Optimal transport and skorokhod embedding. *arXiv preprint arXiv:1307.3656*, 2013.
[8] Mathias Beiglboeck and Nicolas Juillet. On a problem of optimal transport under marginal martingale constraints. *arXiv preprint arXiv:1208.1509*, 2012.
[9] Jean-David Benamou, Guillaume Carlier, Marco Cuturi, Luca Nenna, and Gabriel Peyré. Iterative bregman projections for regularized transportation problems. *SIAM Journal on Scientific Computing*, 37(2):A1111–A1138, 2015.
[10] Jean-David Benamou, Guillaume Carlier, and Luca Nenna. A numerical method to solve optimal transport problems with coulomb cost. *arXiv preprint arXiv:1505.01136, to appear as a chapter in "Splitting Methods in Communication and Imaging, Science and Engineering", Editors R. Glowinski, S. Osher, and W. Yin., Springer*, 2015.
[11] Enrico 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*, 2008.
[12] Giuseppe Buttazzo, Luigi De Pascale, and Paola Gori-Giorgi. Optimal-transport formulation of electronic density-functional theory. *Physical Review A*, 85(6):062502, 2012.
[13] Luis A Caffarelli. A localization property of viscosity solutions to the monge-ampere equation and their strict convexity. *Annals of Mathematics*, pages 129–134, 1990.
[14] Luis A Caffarelli. Some regularity properties of solutions of monge ampere equation. *Communications on pure and applied mathematics*, 44(8-9):965–969, 1991.
[15] G Carlier and I Ekeland. Matching for teams. *Econom. Theory*, 42(2):397–418, 2010.
[16] Guillaume Carlier, Victor Chernozhukov, and Alfred Galichon. Vector quantile regression. *arXiv preprint arXiv:1406.4643*, 2014.
[17] Guillaume Carlier and Bruno Nazaret. Optimal transportation for the determinant. *ESAIM: Control, Optimisation and Calculus of Variations*, 14(04):678–698, 2008.
[18] Guillaume Carlier, Adam Oberman, and Edouard Oudet. Numerical methods for matching for teams and wasserstein barycenters. *arXiv preprint arXiv:1411.3602*, 2014.
[19] Huajie Chen and Gero Friesecke. Pair densities in density functional theory. *arXiv preprint arXiv:1503.07539*, 2015.
[20] Huajie Chen, Gero Friesecke, and Christian B Mendl. Numerical methods for a kohn–sham density functional model based on optimal transport. *Journal of Chemical Theory and Computation*, 10(10):4360–4368, 2014.
[21] Pierre-André Chiappori, Alfred Galichon, and Bernard Salanié. The roommate problem is more stable than you think. 2014.
[22] Pierre-André Chiappori, Robert J McCann, and Lars P Nesheim. Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness. *Economic Theory*, 42(2):317–354, 2010.
[23] Aron J Cohen, Paula Mori-Sánchez, and Weitao Yang. Challenges for density functional theory. *Chemical Reviews*, 112(1):289–320, 2011.
[24] Maria Colombo and Simone Di Marino. Equality between monge and kantorovich multimarginal problems with coulomb cost. *Annali di Matematica Pura ed Applicata*, pages 1–14, 2013.
[25] Maria Colombo and Federico Stra. Counterexamples to multimarginal optimal transport maps with coulomb cost and radial measures. *in preparation.*
[26] Roberto Cominetti and Jaime San Martín. Asymptotic analysis of the exponential penalty trajectory in linear programming. *Mathematical Programming*, 67(1-3):169–187, 1994.
[27] Codina Cotar, Gero Friesecke, and Claudia Klüppelberg. Density functional theory and optimal transportation with coulomb cost. *Communications on Pure and Applied Mathematics*, 66(4):548–599, 2013.
[28] Codina Cotar, Gero Friesecke, and Brendan Pass. Infinite-body optimal transport with coulomb cost. *Calculus of Variations and Partial Differential Equations*, pages 1–26, 2013.

[29] Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems*, pages 2292–2300, 2013.
[30] Luigi De Pascale. Optimal transport with coulomb cost. approximation and duality. *arXiv preprint arXiv:1503.07063*, 2015.
[31] Simone Di Marino, Luigi De Pascale, and Maria Colombo. Multimarginal optimal transport maps for 1-dimensional repulsive costs. *CANADIAN JOURNAL OF MATHEMATICS-JOURNAL CANADIEN DE MATHEMATIQUES*, 67:350–368, 2015.
[32] Simone Di Marino, Augusto Gerolin, Luca Nenna, Michael Seidl, and Paola Gori-Giorgi. The strictly-correlated electron functional for spherically symmetric systems revisited. *in preparation*.
[33] Yan Dolinsky and Halil Mete Soner. Martingale optimal transport and robust hedging in continuous time. *Probability Theory and Related Fields*, 160(1-2):391–427, 2014.
[34] Yan Dolinsky and Halil Mete Soner. Robust hedging with proportional transaction costs. *Finance and Stochastics*, 18(2):327–347, 2014.
[35] Ivar Ekeland. An optimal matching problem. *ESAIM: Control, Optimisation and Calculus of Variations*, 11(01):57–71, 2005.
[36] Lawrence C Evans and Maciej Zworski. Lectures on semiclassical analysis, 2007.
[37] Gero Friesecke. The multiconfiguration equations for atoms and molecules: charge quantization and existence of solutions. *Archive for Rational Mechanics and Analysis*, 169(1):35–71, 2003.
[38] Gero Friesecke, Christian B Mendl, Brendan Pass, Codina Cotar, and Claudia Klüppelberg. N-density representability and the optimal transport limit of the hohenberg-kohn functional. *The Journal of chemical physics*, 139(16):164109, 2013.
[39] Alfred Galichon, Pierre Henry-Labordère, Nizar Touzi, et al. A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *The Annals of Applied Probability*, 24(1):312–336, 2014.
[40] Alfred Galichon and Bernard Salanié. Matching with trade-offs: Revealed preferences over competing characteristics. 2010.
[41] Wilfrid Gangbo and Andrzej Swiech. Optimal maps for the multidimensional monge-kantorovich problem. *Communications on pure and applied mathematics*, 51(1):23–45, 1998.
[42] Nassif Ghoussoub and Bernard Maurey. Remarks on multi-marginal symmetric monge-kantorovich problems. *arXiv preprint arXiv:1212.1680*, 2012.
[43] Nassif Ghoussoub and Abbas Moameni. Symmetric monge–kantorovich problems and polar decompositions of vector fields. *Geometric and Functional Analysis*, 24(4):1129–1166, 2014.
[44] Nassif Ghoussoub and Brendan Pass. Decoupling of degiorgi-type systems via multi-marginal optimal transport. *Communications in Partial Differential Equations*, 39(6):1032–1047, 2014.
[45] Paola Gori-Giorgi, Michael Seidl, and Giovanni Vignale. Density-functional theory for strongly interacting electrons. *Physical review letters*, 103(16):166402, 2009.
[46] Leonid Kantorovich. On a problem of monge. *Journal of Mathematical Sciences*, 133(4):1383–1383, 2006.
[47] Leonid Kantorovitch. On the translocation of masses. *Management Science*, 5(1):1–4, 1958.
[48] Hans G Kellerer. Duality theorems for marginal problems. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 67(4):399–432, 1984.
[49] Young-Heon Kim and Brendan Pass. Multi-marginal optimal transport on riemannian manifolds. *arXiv preprint arXiv:1303.6251*, 2013.
[50] Young-Heon Kim and Brendan Pass. A general condition for monge solutions in the multi-marginal optimal transport problem. *SIAM Journal on Mathematical Analysis*, 46(2):1538–1550, 2014.
[51] Young-Heon Kim and Brendan Pass. Wasserstein barycenters over riemannian manfolds. *arXiv preprint arXiv:1412.7726*, 2014.
[52] Jun Kitagawa and Brendan Pass. The multi-marginal optimal partial transport problem. *arXiv preprint arXiv:1401.7255*, 2014.
[53] Martin Knott and Cyril S. Smith. On a generalization of cyclic monotonicity and distances among random vectors. *Linear algebra and its applications*, 199:363–371, 1994.
[54] Alexander V Kolesnikov and Danila A Zaev. Optimal transportation of processes with infinite kantorovich distance. independence and symmetry. *arXiv preprint arXiv:1303.7255*, 2013.

[55] David C Langreth and John P Perdew. The exchange-correlation energy of a metallic surface. *Solid State Communications*, 17(11):1425–1429, 1975.
[56] Elliott H Lieb. A lower bound for coulomb energies. *Physics Letters A*, 70(5):444–446, 1979.
[57] Elliott H Lieb. Density functionals for coulomb systems. In *Inequalities*, pages 269–303. Springer, 2002.
[58] Elliott H Lieb and Stephen Oxford. Improved lower bound on the indirect coulomb energy. *International Journal of Quantum Chemistry*, 19(3):427–439, 1981.
[59] Elliott H Lieb and Barry Simon. The hartree-fock theory for coulomb systems. *Communications in Mathematical Physics*, 53(3):185–194, 1977.
[60] Pierre-Louis Lions. Solutions of hartree-fock equations for coulomb systems. *Communications in Mathematical Physics*, 109(1):33–97, 1987.
[61] Artur O Lopes and Jairo K Mengue. Duality theorems in ergodic transport. *Journal of Statistical Physics*, 149(5):921–942, 2012.
[62] Artur O Lopes, Elismar R Oliveira, and Philippe Thieullen. The dual potential, the involution kernel and transport in ergodic optimization. *arXiv preprint arXiv:1111.0281*, 2011.
[63] Francesc Malet and Paola Gori-Giorgi. Strong Correlation in Kohn-Sham Density Functional Theory. *Physical Review Letters*, 109(24):246402, 2012.
[64] Christian B. Mendl and Lin Lin. Kantorovich dual solution for strictly correlated electrons in atoms and molecules. *Physical Review B*, 87(12):125106, 2013.
[65] Gaspard Monge. *Mémoire sur la théorie des déblais et des remblais*. De l'Imprimerie Royale, 1781.
[66] Ingram Olkin and Svetlozar T. Rachev. Maximum submatrix traces for positive definite matrices. *SIAM journal on matrix analysis and applications*, 14(2):390–397, 1993.
[67] Brendan Pass. *Structural results on optimal transportation plans*. PhD thesis, University of Toronto, 2011.
[68] Brendan Pass. Uniqueness and monge solutions in the multimarginal optimal transportation problem. *SIAM Journal on Mathematical Analysis*, 43(6):2758–2775, 2011.
[69] Brendan Pass. On the local structure of optimal measures in the multi-marginal optimal transportation problem. *Calculus of Variations and Partial Differential Equations*, 43(3-4):529–536, 2012.
[70] Brendan Pass. Remarks on the semi-classical hohenberg–kohn functional. *Nonlinearity*, 26(9):2731, 2013.
[71] Brendan Pass. Multi-marginal optimal transport and multi-agent matching problems: uniqueness and structure of solutions. *Discrete Contin. Dyn. Syst., 34:1623-1639*, 2014.
[72] Brendan Pass. Multi-marginal optimal transport: theory and applications. *arXiv preprint arXiv:1406.0026*, 2014.
[73] Simone Paziani, Saverio Moroni, Paola Gori-Giorgi, and Giovanni B Bachelet. Local-spin-density functional for multideterminant density functional theory. *Physical Review B*, 73(15):155111, 2006.
[74] Aldo Pratelli. On the equality between monge's infimum and kantorovich's minimum in optimal mass transportation. In *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, volume 43, pages 1–13. Elsevier, 2007.
[75] Aldo Pratelli. On the sufficiency of c-cyclical monotonicity for optimality of transport plans. *Mathematische Zeitschrift*, 2007.
[76] Didier Robert. *Autour de l'approximation semi-classique*, volume 68. Birkhäuser Basel, 1987.
[77] Ludger Ruschendorf. Convergence of the iterative proportional fitting procedure. *The Annals of Statistics*, pages 1160–1174, 1995.
[78] Ludger Rüschendorf and Wolfgang Thomsen. Closedness of sum spaces andthe generalized schrödinger problem. *Theory of Probability & Its Applications*, 42(3):483–494, 1998.
[79] Ludger Rüschendorf and Ludger Uckelmann. *On optimal multivariate couplings*. Springer, 1997.
[80] Walter Schachermayer and Josef Teichmann. Characterization of optimal transport plans for the monge-kantorovich-problem. *proceedings of the A.M.S.*, Vol. 137.(2):519–529, 2009.
[81] Erwin Schrödinger. *Über die umkehrung der naturgesetze*. Verlag Akademie der wissenschaften in kommission bei Walter de Gruyter u. Company, 1931.
[82] Michael Seidl. Strong-interaction limit of density-functional theory. *Physical Review A*, 60(6):4387, 1999.

[83] Michael Seidl, Paola Gori-Giorgi, and Andreas Savin. Strictly correlated electrons in density-functional theory: A general formulation with applications to spherical densities. *Physical Review A*, 75(4):042511, 2007.

[84] Cyril Smith and Martin Knott. On hoeffding-fréchet bounds and cyclic monotone relations. *Journal of multivariate analysis*, 40(2):328–334, 1992.

[85] Julien Toulouse, Francois Colonna, and Andreas Savin. Short-range exchange and correlation energy density functionals: Beyond the local-density approximation. *The Journal of chemical physics*, 122(1):014110, 2005.

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

[87] Gui-Song Xia, Sira Ferradans, Gabriel Peyré, and Jean-François Aujol. Synthesizing and mixing stationary gaussian texture models. *SIAM Journal on Imaging Sciences*, 7(1):476–508, 2014.

[88] Weitao Yang. Generalized adiabatic connection in density functional theory. *The Journal of chemical physics*, 109(23):10107–10110, 1998.

[89] Danila Zaev. On the monge-kantorovich problem with additional linear constraints. *arXiv preprint arXiv:1404.4962*, 2014.

[90] Grigorii M. Zhislin. Discussion of the spectrum of schrödinger operators for systems of many particles. *Trudy Moskovskogo matematiceskogo obscestva*, 9:81–120, 1960.

Laboratoire de Mathématiques d'Orsay,Université Paris-Sud, FRANCE
E-mail address: simone.dimarino@math.u-psud.fr

Dipartimento di Matematica, Università di Pisa, Largo B. Pontecorvo 5, 56126 Pisa, ITALY
E-mail address: gerolin@mail.dm.unipi.it

INRIA, MOKAPLAN, Domaine de Voluceau Le Chesnay, FRANCE
E-mail address: luca.nenna@inria.fr