MOMENT APPROXIMATIONS IN OPTIMAL TRANSPORT: GEOMETRIC ESTIMATES AND A SEMICLASSICAL LIMIT

LUIGI DE PASCALE, LUCA NENNA, AND IGOR PINHEIRO

Abstract. We study finite moment approximations of linear optimization problems over measures. Under suitable assumptions, we establish the existence of discrete minimizers with a controlled number of support points, duality and Γ-convergence as the moment approximation is refined. This framework accommodates several transport problems, in which marginal constraints are replaced by finitely many moment constraints. For moment functions localized at scale r_n, we prove an estimate controlling W_∞ by W_p up to an additive error of order r_n, which also applies to discrete measures. For repulsive costs, we show that minimizers remain separated from the diagonal, uniformly in the number of moments, and we derive consequences for dual potentials. Finally, under a moment-correction assumption, we prove that the fermionic minimum energies converge in the semiclassical limit to the moment-constrained Coulomb transport value for a fixed family of moment functions.

1. Introduction

The numerical approximation of optimal transport problems becomes particularly difficult when several marginals are prescribed. A common approach consists in approximating each marginal by a discrete measure and solving the resulting finite-dimensional problem. If each of M marginals is supported on n points, the transport plan has up to n^M possible configurations. This difficulty motivates approximations that allow the support of the transport plan to be sparse [2, 1].

In [2], the authors study one such approximation, which consists in replacing the marginal constraints by finitely many moment constraints. The prescribed marginals enter the problem through their integrals against a chosen family of test functions, called moment functions. Under suitable assumptions, the relaxed problem admits a discrete minimizer whose number of support points is controlled by the number of constraints. Both the positions of these points and their weights can then be treated as variables in the approximation.

The choice of moment functions is consequently an essential part of the approximation. The relaxed constraints allow measures whose marginals differ from the prescribed ones, and it is necessary to understand what information about those marginals is retained. In this paper, we study finite moment approximations of linear optimization problems over measures within a common framework, and we investigate which geometric information on the marginals is retained by the relaxation. To fix ideas, we first describe the approximation in the Monge–Kantorovich case.

Let X and Y be Polish spaces, and let $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$. Given a lower semicontinuous cost $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$, bounded from below, the Monge–Kantorovich problem consists in minimizing

$$\mathrm{C}(\pi) = \int_{X\times Y} c(x,y)\ \mathrm{d}\pi(x,y)$$

among probability measures π whose first and second marginals are μ and ν, respectively. We denote the set of all such measures by $\Pi(\mu,\nu)$. Equivalently, these marginal constraints

Date: November 18, 2026.

can be expressed using test functions:

$$\int_{X\times Y} \varphi(x)\ \mathrm{d}\pi(x,y) = \int_X \varphi\ \mathrm{d}\mu \quad \text{for every } \varphi \in C_b(X),$$
$$\int_{X\times Y} \psi(y)\ \mathrm{d}\pi(x,y) = \int_Y \psi\ \mathrm{d}\nu \quad \text{for every } \psi \in C_b(Y).$$

The moment relaxation consists in requiring these identities only for selected finite families of test functions. For each $n \in \mathbb{N}$, let $\mathcal{I}^n \subset C_b(X)$ and $\mathcal{J}^n \subset C_b(Y)$ be finite families of moment functions. We denote by $\Pi^n(\mu,\nu)$ the set of probability measures satisfying the preceding identities for every $\varphi \in \mathcal{I}^n$ and $\psi \in \mathcal{J}^n$. The moment-constrained Monge–Kantorovich problem is then

$$\inf_{\pi\in\Pi^n(\mu,\nu)} \mathrm{C}(\pi). \tag{1}$$

Since $\Pi(\mu,\nu) \subseteq \Pi^n(\mu,\nu)$, the infimum of the relaxed problem is no larger than that of the original problem; when the moment functions are refined so as to fully characterize μ and ν, the relaxed values converge to the original one (see Theorem 2.13).

Even when the relaxation preserves the optimal value, it may admit minimizers with very different supports, as the following example shows.

Example 1.1. Let $X = Y = [0,1]$, consider the attractive cost $c(x,y) = |x-y|^2$, and take

$$\mu = \nu = \frac{1}{2}\delta_{1/4} + \frac{1}{2}\delta_{3/4}.$$

The unique minimizer of the original problem is

$$\pi^\star = \frac{1}{2}\delta_{(1/4,1/4)} + \frac{1}{2}\delta_{(3/4,3/4)}.$$

For $\mathcal{I}^3 = \mathcal{J}^3 = \{x \mapsto x^k : k = 1,2,3\}$, a minimizer of the moment-constrained problem is

$$\pi^3 = \frac{1}{8}\delta_{(0,0)} + \frac{3}{4}\delta_{(1/2,1/2)} + \frac{1}{8}\delta_{(1,1)}, \qquad \mathrm{C}(\pi^3) = 0.$$

Both its marginals are

$$\frac{1}{8}\delta_0 + \frac{3}{4}\delta_{1/2} + \frac{1}{8}\delta_1,$$

which have the same first three moments as μ. Nevertheless, their support is disjoint from that of μ, and $\mathrm{supp}(\pi^3) \cap \mathrm{supp}(\pi^\star) = \varnothing$.

Example 1.1 shows that the optimal value alone does not detect how far the relaxed marginals are from the prescribed ones: the moment functions have to be chosen so as to retain geometric information. We now present our main results in a simplified form, together with their consequences; precise statements and assumptions are given in the subsequent sections. Throughout, for a cost c we write $\mathrm{C}(\pi) = \int c\ \mathrm{d}\pi$.

Geometric estimates. We say that the moment functions $\varphi_1,\dots,\varphi_n$ are *localized at scale* r_n on the support of μ if they form a partition of unity on $\mathrm{supp}(\mu)$ made of functions supported in balls of radius r_n. Any probability measure having the same integrals as μ against $\varphi_1,\dots,\varphi_n$ is called a *relaxed marginal* of μ.

Theorem A (Geometric estimates)**.** *Assume that the moment functions are localized at scale r_n. Then every relaxed marginal μ_n of μ satisfies $W_\infty(\mu,\mu_n) \le 2r_n$. Moreover, if μ has a density bounded from below on a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$ and ν is concentrated on $\overline{\Omega}$, then, for all relaxed marginals μ_n of μ and ν_n of ν,*

$$W_\infty(\mu_n,\nu_n) \le 4r_n + C\, W_p(\mu_n,\nu_n)^{\frac{p}{p+d}}.$$

The first estimate rules out the behavior of Example 1.1: with localized moment functions, the relaxation cannot move any portion of mass farther than $2r_n$, so that the relaxed marginals converge to the prescribed ones in the strongest Wasserstein sense as $r_n \to 0$, and not only narrowly. The second estimate compares an averaged distance with a uniform one. If μ_n and ν_n are close on average, namely $W_p(\mu_n, \nu_n)$ is small, then they can be coupled without moving any point farther than $C\, W_p(\mu_n, \nu_n)^{p/(p+d)}$, up to the resolution $4r_n$ of the moment functions. Estimates of this type are classical when one of the measures has a density bounded from below [7, 20]. Here, the constants only depend on the prescribed marginal μ, so that the estimate applies to the relaxed marginals, which are typically discrete. See Lemma 3.2, Theorem 3.4 and Corollary 3.5 for the precise statements.

Repulsive costs. An important motivation for these sparse discretizations comes from multi-marginal optimal transport problems with Coulomb cost, which arise in the strictly correlated electrons limit of density functional theory [9, 12] and for which moment-constrained approximations were introduced in [2, 1, 16]. For such costs, it is natural to ask whether the support of a minimizer of the moment relaxation remains at a positive distance from the diagonal, where the cost is singular.

We consider $M \geq 2$ marginals all equal to $\rho \in \mathcal{P}(X)$ and costs of the form $c(x_1, \dots, x_M) = \sum_{i<j} h(d(x_i, x_j))$, where h is decreasing and blows up at 0, such as the Coulomb cost $h(r) = 1/r$ on $\mathbb{R}^3$. The moment relaxation only prescribes the integrals of the marginals against finitely many moment functions, which need not be localized.

Theorem B (Separation from the diagonal)**.** *Assume that ρ does not concentrate too much mass on small balls. Then there exists $\alpha^\star > 0$, independent of the number n of moment functions, such that every minimizer of the moment-constrained problem gives no mass to configurations in which two points are closer than $\alpha^\star$.*

This has two consequences. First, since minimizers never see the singularity, the cost can be replaced by a bounded continuous truncation without changing the minimizers (Theorem 4.6); this justifies the use of bounded costs in numerical computations. Second, when X is compact and the moment functions are localized, the dual problem admits a symmetric maximizer in the span of the moment functions, with explicit bounds on its coefficients (Theorem 4.8), in the spirit of the bounds on Kantorovich potentials for the Coulomb cost [11, 8]. See Theorem 4.3 and Corollary 4.5 for the precise statements.

A semiclassical limit. Our third main result connects the moment-constrained Coulomb problem with its quantum counterpart. In density functional theory, the multi-marginal problem with Coulomb cost arises as the semiclassical limit of the Lieb functional [12, 5, 24, 13]: when the kinetic energy is multiplied by a small parameter ε, the minimal energy of M fermions with a prescribed density converges to the optimal transport value. We consider the same quantum problem, where the density constraint is replaced by finitely many moment constraints on the M-particle law, as in [16].

Theorem C (Semiclassical limit)**.** *Fix $n \in \mathbb{N}$ and let c be the Coulomb cost. Under a technical assumption ensuring that the moment constraints can be corrected, the minimal energy*

$$\inf_{\Gamma} \left\{ \frac{\varepsilon}{2} \operatorname{Tr}(-\Delta\Gamma) + \mathrm{C}(\pi_\Gamma) \right\},$$

where the infimum is taken over the fermionic states Γ whose M-particle law π_Γ satisfies the moment constraints, converges, as $\varepsilon \to 0$, to the minimum of C under the same moment constraints. Moreover, the laws of almost minimizing states converge, up to subsequences, to minimizers of the moment-constrained transport problem.

In other words, imposing finitely many moments and taking the semiclassical limit are compatible operations: the moment-constrained transport problem studied in the previous results is exactly the strictly correlated limit of the moment-constrained quantum problem. This justifies the use of moment-constrained optimal transport as an approximation of the quantum problem in the strongly correlated regime. The result follows from a Γ-convergence statement (Theorem 5.7). The construction of recovery sequences relies on three steps: we approximate a transport plan by finitely supported plans with finite cost, which are therefore separated from the diagonal; we replace each atom by a smooth fermionic wave packet, which keeps the Coulomb energy under control; finally, we restore the moment constraints exactly through convex combinations with finitely many fixed states, which is where the technical assumption (Assumption 5.1) enters.

Related works. Finite moment relaxations of the marginal constraints were introduced in [2] for multi-marginal optimal transport, where the existence of sparse minimizers and the convergence as the number of moments increases are proved, and numerical methods are proposed; see also [1] for symmetric multi-marginal problems and [16] for the corresponding approximation of the Lieb functional. Other forms of moment constraints in optimal transport have been considered more recently. Additional linear constraints on the transport plan, with exact marginals, are studied in [32]. In [10], weak optimal transport problems with hard and soft moment constraints on the disintegration of the plan, which include martingale transport, are studied, and a Slater-type qualification condition yielding dual attainment is introduced. In [21], a martingale transport problem in which the second marginal is only prescribed through finitely many moments is considered, and a finite-dimensional dual problem without duality gap is derived under an interiority condition on the moment vector, which plays a role similar to our full-support assumption in Theorem 2.7. Finally, the extreme points of moment sets [31, 26] and the Richter–Tchakaloff theorem [3] are classical tools of the theory of moments [29, 15].

Compared with these works, Section 2 gathers in a single framework the existence of minimizers, their discreteness and Γ-convergence for a linear problem over nonnegative measures with finitely many moment inequalities and a confinement constraint, which are then applied to all the problems of Section 6. As concerns the duality result, Theorem 2.7, which holds under a full-support condition and reduces the dual problem to a finite-dimensional maximization over the coefficients of the moment functions (see Remark 2.11), is, to the best of our knowledge, new in this generality.

General framework and organization of the paper. Several variational problems share the linear moment structure used above. In Section 2, we therefore study a general linear optimization problem over nonnegative measures of fixed mass, subject to moment equalities and inequalities. This formulation allows the basic approximation results to be proved once and then applied through the choice of the underlying space, cost and moment functions. Under suitable assumptions, we establish existence of minimizers, obtain discrete minimizers through the Richter–Tchakaloff theorem and extreme-point arguments, and derive duality and optimality conditions. We also prove Γ-convergence as the moment approximation is refined. The section concludes with the Monge–Kantorovich case, including convergence of the optimal values and subsequential convergence of minimizers when the moment functions fully characterize the prescribed marginals.

In Section 3, we study how localized moment functions control the geometry of the relaxed marginals and prove the estimate relating W_p and W_∞. Section 4 establishes the diagonal exclusion estimates and their consequences for cost truncation and dual potentials. Section 5 discusses the semiclassical application through the smoothing of atomic measures and the correction of their moments.

Section 6 shows how the general theory applies to partial optimal transport, JKO steps with linear potential energies, martingale optimal transport on compact sets, capacity-constrained optimal transport, and Wasserstein-on-Wasserstein transport. For each problem, we identify the finite moment relaxation and state the corresponding existence, duality and convergence results. These applications explain the scope of the abstract formulation: moment inequalities describe partial transport and capacity constraints, additional equalities encode the martingale condition, and the underlying space may itself be a space of probability measures.

2. General theory of moment-constrained optimization

In this section, we study the abstract moment-constrained problem underlying all the applications of this paper; the Monge–Kantorovich case is recovered in Section 2.6. For background on the general moment problem, we refer to [29, 15].

2.1. **Formulation.** Let Z be a Polish space, let $\mathcal{M}_+(Z)$ denote the set of finite nonnegative Borel measures on Z, and let $c : Z \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous cost function bounded from below. Consider the functional

$$\mathrm{C} : \mathcal{M}_+(Z) \to \mathbb{R} \cup \{+\infty\}, \qquad \mathrm{C}(\pi) = \int_Z c \, d\pi.$$

Let $w_1, \dots, w_n \in C_b(Z)$ be bounded continuous functions. For a fixed, coercive[1], lower semicontinuous function $\Theta : Z \to [0, +\infty]$, and for $A > 0$ and $m > 0$ we define the following variational problem[2]

$$\inf_{\pi \in \Pi_A^n} \mathrm{C}(\pi). \tag{2}$$

where

$$\begin{aligned} \Pi_A^n \coloneqq \{\, \pi \in \mathcal{M}_+(Z) :\ & \pi(Z) = m, \\ & \int_Z w_i \, d\pi \geq 0 \quad (i = 1, \dots, n), \quad \int_Z \Theta \, d\pi \leq A \,\}. \end{aligned} \tag{3}$$

2.2. **Existence.**

Theorem 2.1. *If the infimum in* (2) *is finite, then there exists at least one minimizer.*

Proof. As (2) is finite we have that $\Pi_A^n \neq \emptyset$. Let $(\pi_k)_{k \in \mathbb{N}} \subset \Pi_A^n$ be a minimizing sequence. Since

$$\int \Theta \, \mathrm{d}\pi_k \leq A \text{ for all } k \in \mathbb{N}$$

and Θ is coercive, the sequence is tight. Passing to a subsequence without changing notation we have $\pi_k \rightharpoonup \pi_\infty$ and $\pi_\infty \in \mathcal{M}_+(Z)$. Let us show that π_∞ is admissible. From weak convergence, $\pi_\infty(Z) = m$ and from the lower semicontinuity of Θ, $\int \Theta d\pi_\infty \leq A$. Moreover, the functional $\pi \mapsto \int w \, \mathrm{d}\pi$ is continuous, as $w \in C_b(Z)$, therefore weak convergence again guarantees that $\int_Z w_i \, \mathrm{d}\pi_\infty \geq 0$ for every $1 \leq i \leq n$ and thus

$$\int c \, \mathrm{d}\pi_\infty \geq \inf_{\pi \in \Pi_A^n} \int_Z c \, \mathrm{d}\pi.$$

The other inequality follows from the lower semicontinuity of the cost. The same argument, applied to an arbitrary sequence in Π_A^n, shows that Π_A^n is narrowly compact. □

[1] Every sublevel set of Θ, $\{\Theta \leq R\}$, is compact for every $R \geq 0$.

[2] If $\Pi_A^n = \emptyset$, we assume the infimum to be $+\infty$.

2.3. **Characterization.**

Theorem 2.2 (Richter–Tchakaloff [3])**.** *Let Z be a Polish space and let π be a nonzero finite nonnegative Borel measure concentrated on a Borel set $E \subset Z$. Let $N \in \mathbb{N}$ and let $\Lambda : Z \to \mathbb{R}^N$ be a Borel measurable map. Assume that*

$$\int_{\mathbb{R}^N} \|u\| \, \mathrm{d}(\Lambda_{\#}\pi)(u) = \int_Z \|\Lambda(z)\| \, \mathrm{d}\pi(z) < +\infty, \tag{4}$$

where $\|\cdot\|$ is the Euclidean norm. Assume moreover that $1 \in \mathrm{span}(\{\Lambda_1, \dots, \Lambda_N\})$. Then, there exist $1 \le K \le N$, distinct points $z_1, \dots, z_K \in E$ and positive weights $t_1, \dots, t_K$ such that

$$\int_Z \Lambda_i(z) \, \mathrm{d}\pi(z) = \sum_{k=1}^K t_k \Lambda_i(z_k), i = 1, \dots, N \quad \text{and} \quad \sum_{k=1}^K t_k = \pi(Z). \tag{5}$$

Theorem 2.3. *If the infimum in (2) is finite, there exist $1 \le K \le n+3$, distinct points $z_1, \dots, z_K$ and positive weights $t_1, \dots, t_K$ such that $\pi = \sum_{i=1}^K t_i \delta_{z_i}$ is a minimizer of (2).*

Proof. By Theorem 2.1, there exists a minimizer $\bar\pi \in \Pi_A^n$ with finite cost. Thus, $\bar\pi$ is concentrated on the set $E = \{z \in Z : c(z) < +\infty, \Theta(z) < +\infty\}$. Define $\tilde c$ and $\tilde\Theta$ to agree with c and Θ on E and to equal zero on E^c. Then define the Borel map $\Lambda : Z \to \mathbb{R}^{n+3}$ by $\Lambda(z) = (w_1(z), \dots, w_n(z), \tilde\Theta(z), \tilde c(z), 1)$. All components Λ_i of Λ are $\bar\pi$-integrable, so

$$\int \|\Lambda(z)\| \, \mathrm{d}\bar\pi(z) < +\infty.$$

Applying Theorem 2.2 yields a discrete measure that is again a minimizer, since we included c as one of the coordinates of Λ. □

The set Π_A^n is convex, as it is determined by linear constraints. We therefore denote by $\mathrm{ex}(\Pi_A^n)$ the set of extreme points: if $\pi \in \mathrm{ex}(\Pi_A^n)$ and $\pi = (1-t)\pi^+ + t(\pi^-)$ for some $\pi^+, \pi^- \in \Pi_A^n$ and $t \in (0,1)$, then $\pi^+ = \pi^-$. We also introduce the following notation:

$$\Phi = (w_1, \dots, w_n, \Theta, 1). \tag{6}$$

Lemma 2.4. *Let $\pi \in \mathrm{ex}(\Pi_A^n)$. If $h \in L^\infty(\pi)$ satisfies*

$$\int h\Phi \, \mathrm{d}\pi = 0$$

then $h = 0$ π-almost everywhere.

Proof. For $0 < \epsilon < \frac{1}{1+\|h\|_\infty}$ define the nonnegative measures $\pi^\pm = (1 \pm \epsilon h)\pi$. From $\int h\Phi \, \mathrm{d}\pi = 0$ we have that $\pi^\pm \in \Pi_A^n$, moreover $\pi = \frac{1}{2}(\pi^+ + \pi^-)$. By extremality we have $\pi^+ = \pi^-$ and thus $h \equiv 0$ π-a.e. □

Theorem 2.5.

$$\mathrm{ex}(\Pi_A^n) \subset \left\{ \begin{array}{c} \pi \in \Pi_A^n \ : \pi = \displaystyle\sum_{i=1}^k t_i \delta_{z_i} \text{ for } \{z_i\}_{i=1}^k \subset Z, \sum_{i=1}^k t_i = m, 1 \le k \le n+2, \\ (w_1(z_i), \dots, w_n(z_i), \Theta(z_i), 1)_{i=1}^k \text{ are linearly independent} \end{array} \right\}$$

Proof. Let $\pi \in \mathrm{ex}(\Pi_A^n)$ and assume that $\#\mathrm{supp}(\pi) > n+2$. We can choose $n+3$ pairwise disjoint Borel sets $B_1, \dots B_{n+3}$, each having positive π-mass, and define the vectors

$$v_j = \int_{B_j} \Phi \, \mathrm{d}\pi \in \mathbb{R}^{n+2}, \ j = 1 \dots, n+3. \tag{7}$$

As we have $n+3$ vectors v_j in $\mathbb{R}^{n+2}$, these vectors are therefore linearly dependent. Thus, there exists $(a_j)_{j=1}^{n+3}$ not all zero such that

$$\sum_{j=1}^{n+3} a_j v_j = 0. \tag{8}$$

Define the function $h = \sum_{j=1}^{n+3} a_j 1_{B_j}$, clearly $h \in L^\infty(\pi)$ and $h \not\equiv 0$. On the other hand, note that

$$\int h\Phi \,\mathrm{d}\pi = \sum_{j=1}^{n+3} a_j \int_{B_j} \Phi \,\mathrm{d}\pi = \sum_{j=1}^{n+3} a_j v_j = 0 \tag{9}$$

and by Lemma 2.4 we have a contradiction.

Therefore, there exist distinct points $z_1, \dots, z_k \in Z$ and positive weights $t_1, \dots, t_k \in \mathbb{R}$ such that $\pi = \sum_{i=1}^k t_i \delta_{z_i}$ for $1 \le k \le n+2$. Let us show now that the vectors $\Phi(z_1), \dots, \Phi(z_k)$ are linearly independent. Assume that there exist $b_1, \dots, b_k$, not all zero, such that $\sum_{i=1}^k b_i \Phi(z_i) = 0$. Define $h(z_i) = b_i/t_i$ and $h = 0$ elsewhere, we have that $h \not\equiv 0$ and $h \in L^\infty(\pi)$. Moreover,

$$\int h\Phi \,\mathrm{d}\pi = \sum_{i=1}^{k} b_i \Phi(z_i) = 0 \tag{10}$$

hence, by Lemma 2.4 we again have a contradiction. □

Since C is linear and Π_A^n is convex and narrowly compact, the minimum of C over Π_A^n is attained at an extreme point, as made precise in the following proposition (see [31, 26] for the extreme-point approach to optimization over moment sets).

Proposition 2.6. *If the infimum in* (2) *is finite, then it is attained at an extreme point of* Π_A^n. *In particular, by Theorem 2.5, there exists a minimizer supported on at most* $n+2$ *points. Moreover, every minimizer is the narrow limit of a sequence of finitely supported minimizers.*

Proof. Define the set

$$F = \{\pi \in \Pi_A^n \,:\, \mathrm{C}(\pi) = \inf_{\pi \in \Pi_A^n} \mathrm{C}(\pi)\}. \tag{11}$$

F is nonempty by Theorem 2.1. The set F is closed as it is the same as $\{\pi \in \Pi_A^n \,:\, \mathrm{C}(\pi) \le \inf_{\pi\in\Pi_A^n} \mathrm{C}(\pi)\}$ and C is lower semicontinuous; moreover, F is narrowly compact as Π_A^n is narrowly compact (Theorem 2.1). Moreover, F is convex since Π_A^n is convex and C is linear. By the Krein–Milman theorem,

$$F = \overline{\mathrm{co}}(\mathrm{ex}(F)) \tag{12}$$

and, as $F \neq \emptyset$, $\mathrm{ex}(F) \neq \emptyset$ and hence

$$\inf_{\pi\in\Pi_A^n} \mathrm{C}(\pi) = \inf_{\pi\in\mathrm{ex}(F)} \mathrm{C}(\pi).$$

Now we show that $\mathrm{ex}(F) \subset \mathrm{ex}(\Pi_A^n)$ by showing that F is a face of Π_A^n. Let $\pi \in F$, $\pi_1, \pi_2 \in \Pi_A^n$ such that $\pi = t\pi_1 + (1-t)\pi_2$ with $t \in (0,1)$, then $\mathrm{C}(\pi_1) \ge \inf_{\pi\in\Pi_A^n} \mathrm{C}(\pi)$ and $\mathrm{C}(\pi_2) \ge \inf_{\pi\in\Pi_A^n} \mathrm{C}(\pi)$ and also

$$\mathrm{C}(\pi) = t\mathrm{C}(\pi_1) + (1-t)\mathrm{C}(\pi_2)$$

guarantees that $\mathrm{C}(\pi_1) = \mathrm{C}(\pi_2) = \inf_{\pi\in\Pi_A^n} \mathrm{C}(\pi)$, thus $\pi_1, \pi_2 \in F$ which shows that F is a face.

Lastly, let $\pi \in F$. The identity (12) gives a sequence $\pi_k \in \mathrm{co}(\mathrm{ex}(F))$ such that $\pi_k \rightharpoonup \pi$ narrowly, π_k is a finite convex combination of extreme points of F, which are finitely supported, thus so is π_k, and finally $\pi_k \in F$, so every π_k is a minimizer. □

2.4. **Duality.** In what follows, we consider $\Theta : Z \to [0, +\infty)$ continuous and the Θ-weighted vector space

$$\mathcal{M}_\Theta(Z) \coloneqq \left\{ \pi \in \mathcal{M}(Z) : \int_Z \Theta \, d|\pi| < +\infty \right\},$$

where $\mathcal{M}(Z)$ denotes the space of finite signed Radon measures on Z, and $|\pi|$ denotes the total variation measure of π. We also set

$$\mathrm{C}_{b,\Theta}(Z) \coloneqq C_b(Z) + \mathrm{span}\{\Theta\}.$$

The pairing between $\mathcal{M}_\Theta(Z)$ and $\mathrm{C}_{b,\Theta}(Z)$ is given by

$$\langle \pi, u \rangle \coloneqq \int_Z u \, d\pi, \qquad \pi \in \mathcal{M}_\Theta(Z), \ u \in \mathrm{C}_{b,\Theta}(Z).$$

We endow $\mathcal{M}_\Theta(Z)$ with the weak topology $\sigma(\mathcal{M}_\Theta(Z), \mathrm{C}_{b,\Theta}(Z))$, i.e. the topology generated by the maps

$$\pi \mapsto \int_Z f \, d\pi, \ f \in C_b(Z) \text{ and } \pi \mapsto \int_Z \Theta \, d\pi.$$

By [28, Theorem 3.10], we have $(\mathcal{M}_\Theta(Z), \sigma(\mathcal{M}_\Theta(Z), \mathrm{C}_{b,\Theta}(Z)))^* \simeq \mathrm{C}_{b,\Theta}(Z)$.

For $m > 0$, $A > 0$ and $W = \{w_1, \dots, w_n\}$ consider the functional $\mathrm{D} : C_b(Z) \times \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R} \cup \{-\infty\}$

$$\mathrm{D}(w, \sigma, \tau) = \begin{cases} \sigma m - \tau A \text{ if } (w, \sigma, \tau) \in \Gamma^n_\Theta \\ -\infty \text{ otherwise} \end{cases} \tag{13}$$

where

$$\Gamma^n_\Theta \coloneqq \{(w, \sigma, \tau) \in C_b(Z) \times \mathbb{R} \times \mathbb{R}_+ \,|\, w \in \mathrm{cone}_+(W), \tau \geq 0, \text{ and } w + \sigma \leq c + \tau\Theta\}.$$

Theorem 2.7. *Let Z, c, Θ and $W = \{w_1, \dots, w_n\}$ be as in (2). Assume that there exists $\eta \in \Pi^n_A$ such that $\int c \, \mathrm{d}\eta < +\infty$ and $\mathrm{supp}(\eta) = \overline{\{c < +\infty\}}$. Then the following duality holds*

$$\max_{(w,\sigma,\tau)} \mathrm{D}(w, \sigma, \tau) = \min_{\pi \in \Pi^n_A} \mathrm{C}(\pi).$$

Moreover, for any minimizer $\pi^\star$ of (2) and dual maximizer $(w^\star, \sigma^\star, \tau^\star)$, the following optimality condition holds:

$$c(z) + \tau^\star \Theta(z) = w^\star(z) + \sigma^\star \qquad \pi^\star\text{-almost everywhere.} \tag{14}$$

Proof. Recall that on $\mathcal{M}_\Theta(Z)$ we consider the Θ-weighted topology. Define $\mathcal{F} : \mathcal{M}_\Theta(Z) \to \mathbb{R} \cup \{+\infty\}$ by

$$\mathcal{F}(\pi) = \begin{cases} \int_Z c \, \mathrm{d}\pi \text{ if } \pi \text{ a positive measure} \\ +\infty \text{ otherwise} \end{cases}$$

note that $\mathcal{F}$ is proper (since it is finite on $\pi = 0$) and lower semicontinuous, as c is lower semicontinuous and bounded below, and $\mathcal{F}$ is convex.

Define $\Xi = \{(a, s, t) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \,|\, a_i \geq 0, \text{ for } 1 \leq i \leq n, \ s = m, \ t \leq A\}$ and let $\mathcal{G} : \mathbb{R}^{n+2} \to \mathbb{R} \cup \{+\infty\}$ be given by

$$\mathcal{G}(a, s, t) = \begin{cases} 0 \text{ if } (a, s, t) \in \Xi \\ +\infty \text{ otherwise} \end{cases}$$

note that Ξ is non-empty, closed and convex, thus $\mathcal{G}$ is proper, lower semicontinuous and convex. Moreover, Ξ is a polyhedron, so $\mathcal{G}$ is polyhedral.

Finally, define the linear transformation $L : \mathcal{M}_\Theta(Z) \to \mathbb{R}^{n+2}$ as

$$L(\pi) = \left(\int_Z w_1 \, \mathrm{d}\pi, \dots, \int_Z w_n \, \mathrm{d}\pi, \int_Z 1 \, \mathrm{d}\pi, \int_Z \Theta \, \mathrm{d}\pi \right)$$

and note that L is continuous for the Θ-weighted topology.

We have

$$\inf_{\pi\in\mathcal{M}_\Theta(Z)}[\mathcal{F}(\pi)+\mathcal{G}(L\pi)] = \inf_{\pi\in\Pi_A^n}\int c\,\mathrm{d}\pi. \tag{15}$$

By assumption, there exists η such that $L(\eta)\in\Xi$. It remains to check that $L(\eta)\in \mathrm{ri}(L(\mathrm{dom}(\mathcal{F})))$, where the relative interior is taken with respect to $H \coloneqq \mathrm{aff}(L(\mathrm{dom}(\mathcal{F})))$. Since $0\in H$, H is a linear space. Thus, we can form a basis of H by choosing points $z_1,\dots,z_q\in\{c<+\infty\}$ and considering the vectors $L(\delta_{z_1}),\dots,L(\delta_{z_q})$. Choose pairwise disjoint neighborhoods U_i of z_i; the vectors $L(\eta|_{U_1}),\dots,L(\eta|_{U_q})$ also form a basis of H, since they are nonzero (here we use the full-support assumption on η) and linear independence is preserved under small perturbations. Consider the measure

$$\eta_t = \eta + \sum_{j=1}^{q} t_j \eta|_{U_j},$$

where each t_j varies in a small neighborhood of zero. For instance, taking $|t_j|<1/2$ ensures that $\eta_t\in\mathrm{dom}(\mathcal{F})$. By linearity,

$$L(\eta_t) = L(\eta) + \sum_{j=1}^{q} t_j L(\eta|_{U_j}).$$

Because the vectors in the sum form a basis, varying each t_j produces a neighborhood of $L(\eta)$ in H. Hence, $L(\eta)\in\mathrm{ri}(L(\mathrm{dom}(\mathcal{F})))$, as required. Since $\mathcal{G}$ is polyhedral, the relative-interior condition above allows us to apply Fenchel–Rockafellar duality (see [6, Theorem 4.2 (ii)]), and we have

$$\inf_{\pi\in\mathcal{M}_\Theta(Z)}[\mathcal{F}(\pi)+\mathcal{G}(L\pi)] = \max_{x^*}[-\mathcal{G}^*(x^*)-\mathcal{F}^*(-L^*x^*)]. \tag{16}$$

Let us compute $\mathcal{F}^*:\mathcal{M}_\Theta(Z)^*\to\mathbb{R}\cup\{+\infty\}$ and $\mathcal{G}^*:\mathbb{R}^{n+2}\to\mathbb{R}\cup\{+\infty\}$. Let $h\in\mathcal{M}_\Theta(Z)^*$, by definition of convex conjugate we have

$$\begin{aligned}\mathcal{F}^*(h) &= \sup_{\pi\in\mathcal{M}_\Theta(Z)\cap\mathcal{M}_+(Z)} \langle\pi,h\rangle - \mathcal{F}(\pi)\\ &= \sup_{\pi\in\mathcal{M}_\Theta(Z)\cap\mathcal{M}_+(Z)} \int_Z h - c\,\mathrm{d}\pi\\ &= \begin{cases} 0 \text{ if } h\le c\\ +\infty \text{ otherwise}\end{cases}\end{aligned} \tag{17}$$

For $\mathcal{G}^*$ let $(\alpha,\sigma,\tau)\in\mathbb{R}^{n+2}$, we have

$$\begin{aligned}\mathcal{G}^*(\alpha,\sigma,\tau) &= \sup_{(a,s,t)\in\mathbb{R}^{n+2}} \langle(a,s,t),(\alpha,\sigma,\tau)\rangle - \mathcal{G}(a,s,t)\\ &= \sup_{\substack{a_i\ge 0,\\ s=m\\ t\le A}} \sum_{i=1}^{n}\alpha_i a_i + \sigma s + \tau t\\ &= \begin{cases} m\sigma + A\tau \text{ if } \alpha_i\le 0 \text{ for } 1\le i\le n, \tau\ge 0\\ +\infty \text{ otherwise}\end{cases}.\end{aligned} \tag{18}$$

Finally, the adjoint of L, L^*, is computed through the identity $\langle(L^*(\alpha,\sigma,\tau),\pi)\rangle = \langle(\alpha,\sigma,\tau),L\pi\rangle$ to be

$$L^*(\alpha,\sigma,\tau) = \sum_{i=1}^{n}\alpha_i w_i + \sigma + \tau\Theta. \tag{19}$$

Substituting (18), (17), and (19) into (16) we obtain, by using that $\mathcal{F}^*(-L^*x^*) = 0$ when $-\sum_{i=1}^n \alpha_i w_i - \sigma - \tau\Theta \leq c$,

$$\begin{aligned} \sup_{x^*} -\mathcal{G}^*(x^*) - \mathcal{F}^*(-L^*x^*) = \\ &= \sup_{\substack{(\alpha,\sigma,\tau)\in\mathbb{R}^{n+2} \\ \tau\geq 0, \alpha_i \leq 0 \\ -\sum_{i=1}^n \alpha_i w_i \leq c+\tau\Theta+\sigma}} -\sigma m - \tau A \\ &= \sup_{\substack{(w,\sigma,\tau)\in \mathrm{cone}_+(W)\times\mathbb{R}\times\mathbb{R}_+ \\ w+\sigma\leq c+\tau\Theta}} \sigma m - \tau A \end{aligned} \tag{20}$$

where we absorbed the negative signs into the coefficients α_i and σ and the conclusion follows from (15).

Let us now show the optimality condition. Take $\pi^\star \in \Pi_A^n$ to be a minimizer of (2) and $(w^\star, \sigma^\star, \tau^\star)$ a maximizer of D. By strong duality we have that

$$0 = \int c \,\mathrm{d}\pi^\star - \sigma^\star m + \tau^\star A = \int (c - \sigma^\star)\,\mathrm{d}\pi^\star + \tau^\star A \tag{21}$$

adding and subtracting the terms $\int w^\star \,\mathrm{d}\pi^\star$ and $\tau^\star \int \Theta \,\mathrm{d}\pi^\star$ we obtain

$$0 = \int (c + \tau^\star\Theta - w^\star - \sigma^\star)\,\mathrm{d}\pi^\star + \int w^\star \,\mathrm{d}\pi^\star + \tau^\star \left(A - \int \Theta \,\mathrm{d}\pi^\star\right) \tag{22}$$

and, as every term on the right-hand side is nonnegative, it follows that

$$c + \tau^\star\Theta - w^\star - \sigma^\star = 0 \,\pi^\star\text{-a.e.}, \ \int w^\star \,\mathrm{d}\pi^\star = 0 \text{ and } \tau^\star \left(A - \int \Theta \,\mathrm{d}\pi^\star\right) = 0. \tag{23}$$

□

2.5. **Γ-convergence.** Thus far we have kept the set of moment functions $\{w_1, \dots, w_n\}$ fixed. In this subsection, we want to let $n \to \infty$ and determine the limit problem of (2). Let us start by fixing some notation. For $n \in \mathbb{N}$ let Π_A^n be the admissible set, as defined in (3) (these sets need not be nested), define $\Pi_A^\infty = \bigcap_{n\in\mathbb{N}} \Pi_A^n$ and the following functionals $\mathrm{C}^n, \mathrm{C}^\infty : \mathcal{M}_+(Z) \to \mathbb{R} \cup \{+\infty\}$ by

$$\mathrm{C}^n(\pi) = \begin{cases} \int_Z c \,\mathrm{d}\pi & \text{if } \pi \in \Pi_A^n, \\ +\infty & \text{otherwise,} \end{cases} \qquad \mathrm{C}^\infty(\pi) = \begin{cases} \int_Z c \,\mathrm{d}\pi & \text{if } \pi \in \Pi_A^\infty, \\ +\infty & \text{otherwise.} \end{cases}$$

Theorem 2.8. *Assume that whenever $n_k \to +\infty$ and $\pi_k \in \Pi_A^{n_k}$ satisfy $\pi_k \rightharpoonup \pi$ narrowly, we have $\pi \in \Pi_A^\infty$. Then C^n Γ-converges to C^∞ with respect to narrow convergence, and the sequence $(\mathrm{C}^n)_{n\in\mathbb{N}}$ is equicoercive.*

In the Monge–Kantorovich case, the assumption of Theorem 2.8 holds when the families of moment functions are increasing and fully characterize the marginals; see Theorem 2.13.

Proof. We start by proving the Γ-lim inf inequality. Let $\pi \in \mathcal{M}_+(Z)$ and consider $\pi^n \rightharpoonup \pi$ narrowly. If $\liminf_{n\to\infty} \mathrm{C}^n(\pi^n) = +\infty$, there is nothing to prove. Assume therefore that $\liminf_{n\to\infty} \mathrm{C}^n(\pi^n) < +\infty$. Then we may extract a subsequence $(n_k)_{k\in\mathbb{N}}$ such that $\pi^{n_k} \in \Pi_A^{n_k}$ for every $k \in \mathbb{N}$ and

$$\liminf_{n\to\infty} \mathrm{C}^n(\pi^n) = \lim_{k\to\infty} \mathrm{C}^{n_k}(\pi^{n_k}).$$

Let us show that this implies $\pi \in \Pi_A^\infty$. Since $\pi^{n_k} \in \Pi_A^{n_k}$ and $\pi^{n_k} \rightharpoonup \pi$ narrowly, this follows directly from the assumption.

Finally, by lower semicontinuity of c we obtain

$$\mathrm{C}^\infty(\pi) = \int_Z c \,\mathrm{d}\pi \leq \liminf_{k\to\infty} \int_Z c \,\mathrm{d}\pi^{n_k} = \liminf_{n\to\infty} \mathrm{C}^n(\pi^n).$$

For the Γ-lim sup inequality, consider now $\pi \in \mathcal{M}_+(Z)$. If $\pi \notin \Pi_A^\infty$, then the limsup inequality is trivially true. Assume $\pi \in \Pi_A^\infty$ and consider the sequence $(\pi^n)_{n\in\mathbb{N}}$ defined by $\pi^n = \pi$ for every $n \in \mathbb{N}$.

By definition of Π_A^∞, $\pi \in \Pi_A^n$ for every n, so $\mathrm{C}^n(\pi) = \mathrm{C}^\infty(\pi)$ for every n. Therefore

$$\limsup_{n\to\infty} \mathrm{C}^n(\pi^n) = \limsup_{n\to\infty} \mathrm{C}^n(\pi) = \mathrm{C}^\infty(\pi).$$

Finally, let us show that $\{\mathrm{C}^n\}_{n\in\mathbb{N}}$ is equicoercive. By Prokhorov's theorem and the fixed mass constraint, it suffices to show that, for every $M > 0$, the sublevel sets $\{\pi : \mathrm{C}^n(\pi) \le M\}$ are uniformly tight with respect to n. Fix $R > 0$. If $\mathrm{C}^n(\pi) \le M < +\infty$, then $\pi \in \Pi_A^n$. Note that, since $\Theta : Z \to [0, +\infty]$ is coercive, the set $K \coloneqq \{z : \Theta(z) \le R\}$ is compact. Hence,

$$\pi(\{\Theta > R\}) = \int_Z \chi_{\{\Theta>R\}} \,\mathrm{d}\pi(z) \le \frac{1}{R}\int_Z \Theta(z)\,\mathrm{d}\pi(z) \le \frac{A}{R},$$

where the last inequality follows from $\pi \in \Pi_A^n$. The result follows by observing that the right-hand side converges to 0 as $R \to +\infty$, uniformly over all these sublevel sets. □

2.6. **The Monge–Kantorovich case.** Let X and Y be Polish spaces, let $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$, and let $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous and bounded below. The Monge–Kantorovich problem is

$$\inf_{\pi\in\Pi(\mu,\nu)} \int_{X\times Y} c(x, y)\,\mathrm{d}\pi(x, y), \tag{24}$$

where

$$\Pi(\mu, \nu) = \big\{\pi \in \mathcal{P}(X \times Y) : (p_X)_\#\pi = \mu,\ (p_Y)_\#\pi = \nu\big\},$$

and $p_X : X \times Y \to X$ and $p_Y : X \times Y \to Y$ denote the canonical projections.

Choose nonnegative, lower semicontinuous functions $\theta_X : X \to [0, +\infty]$, $\theta_Y : Y \to [0, +\infty]$ with compact sublevel sets and finite integrals against μ and ν, respectively. Set

$$\Theta(x, y) = \theta_X(x) + \theta_Y(y)$$

and choose $A > 0$ such that

$$A \ge \int_X \theta_X \,\mathrm{d}\mu + \int_Y \theta_Y \,\mathrm{d}\nu.$$

We set $Z = X \times Y$ and $m = 1$. As moment functions, we choose $\varphi_0, \varphi_1, \dots, \varphi_r \in C_b(X)$ and $\psi_0, \psi_1, \dots, \psi_s \in C_b(Y)$ where $\varphi_0 \equiv \psi_0 \equiv 1$. Let the functions $w_1, \dots, w_n$ of (2) be the functions $f_i, -f_i, g_j, -g_j$, $0 \le i \le r$, $0 \le j \le s$, where

$$f_i = \varphi_i \circ p_X - \int_X \varphi_i \,\mathrm{d}\mu, \qquad g_j = \psi_j \circ p_Y - \int_Y \psi_j \,\mathrm{d}\nu,$$

and we take Θ and A as above. We define $\Pi_A^n(\mu, \nu)$ to be Π_A^n with the choices above and we consider the problem

$$\inf_{\pi\in\Pi_A^n(\mu,\nu)} \int_{X\times Y} c(x, y)\,\mathrm{d}\pi(x, y). \tag{25}$$

Theorem 2.9. *If* (25) *is finite, then it admits a minimizer which is an extreme point of* $\Pi_A^n(\mu, \nu)$ *and it is supported on at most* $r + s + 2$ *points.*

This follows from Theorem 2.5 and Proposition 2.6: up to signs and additive constants, the components of Φ are $\varphi_1 \circ p_X, \dots, \varphi_r \circ p_X$, $\psi_1 \circ p_Y, \dots, \psi_s \circ p_Y$, Θ and 1, so that at most $r + s + 2$ of the vectors $\Phi(z_i)$ can be linearly independent.

For the duality result below, we introduce the following notation: $E_r = \mathrm{span}(\{\varphi_0, \dots, \varphi_r\})$ and $E_s = \mathrm{span}(\{\psi_0, \dots, \psi_s\})$.

Theorem 2.10. *Assume that Θ is finite-valued and continuous. If there exists $\eta \in \Pi_A^n(\mu,\nu)$ such that $\int_{X\times Y} c \,\mathrm{d}\eta < +\infty$ and $\mathrm{supp}(\eta) = \overline{\{c < +\infty\}}$, then*

$$\min_{\pi\in\Pi_A^n(\mu,\nu)} \int_{X\times Y} c \,\mathrm{d}\pi = \max_{\substack{\varphi\in E_r,\ \psi\in E_s,\ \tau\geq 0 \\ \varphi(x)+\psi(y)\leq c(x,y)+\tau\Theta(x,y)}} \left\{ \int_X \varphi \,\mathrm{d}\mu + \int_Y \psi \,\mathrm{d}\nu - \tau A \right\}. \tag{26}$$

The pointwise inequality is required for every $(x,y) \in X \times Y$.

Proof. Since $\mathrm{cone}_+(\{w_1,\dots,w_n\}) = \mathrm{span}(\{f_0,\dots,f_r,g_0,\dots,g_s\})$ we can write any $w \in \mathrm{cone}_+(\{w_1,\dots,w_n\})$ as

$$w = \tilde\varphi \circ p_X - \int_X \tilde\varphi \,\mathrm{d}\mu + \tilde\psi \circ p_Y - \int_Y \tilde\psi \,\mathrm{d}\nu,$$

for some $\tilde\varphi \in E_r$ and $\tilde\psi \in E_s$. Define

$$\varphi = \tilde\varphi - \int \tilde\varphi \,\mathrm{d}\mu + \frac{\sigma}{2} \qquad \text{and} \qquad \psi = \tilde\psi - \int \tilde\psi \,\mathrm{d}\nu + \frac{\sigma}{2}$$

to get

$$\sigma = \int \varphi \,\mathrm{d}\mu + \int \psi \,\mathrm{d}\nu$$

and the conclusion follows from Theorem 2.7. □

Remark 2.11 (A finite dimensional problem). Writing $\varphi = \sum_{i=0}^r a_i\varphi_i$ and $\psi = \sum_{j=0}^s b_j\psi_j$, the dual problem in (26) is a finite-dimensional maximization over the coefficients $(a,b,\tau) \in \mathbb{R}^{r+1} \times \mathbb{R}^{s+1} \times [0,+\infty)$, namely

$$\begin{aligned} &\max_{(a,b,\tau)} \ \sum_{i=0}^r a_i\hat\mu_i + \sum_{j=0}^s b_j\hat\nu_j - \tau A \\ &\text{subject to } \sum_{i=0}^r a_i\varphi_i(x) + \sum_{j=0}^s b_j\psi_j(y) \leq c(x,y) + \tau\Theta(x,y) \quad \text{for every } (x,y) \in X\times Y, \end{aligned}$$

where $\hat\mu_i = \int_X \varphi_i \,\mathrm{d}\mu$ and $\hat\nu_j = \int_Y \psi_j \,\mathrm{d}\nu$. Hence the marginals enter the dual problem only through their moments, and (26) is a semi-infinite linear program, with finitely many unknowns and infinitely many constraints.

The preceding results concern a fixed finite family of moment functions. We now study convergence to the original transport problem as this family is enlarged. To recover the prescribed marginals in the limit, we require that the full collection of moments determines them uniquely. Following [2], we express this requirement through the following notion of full characterization.

Definition 2.12. Let $\mu \in \mathcal{M}_+(X)$. A family of functions $\{\phi^i\}_{i\in\mathbb{N}}$ in $L^1_\mu(X)$ fully characterizes the measure μ if, for any measure $\eta \in \mathcal{M}_+(X)$ satisfying

$$\int_X \phi^i \,\mathrm{d}\eta = \int_X \phi^i \,\mathrm{d}\mu, \ \text{for all } i \in \mathbb{N},$$

it follows that $\eta = \mu$.

Theorem 2.13. *Let $(\varphi_i)_{i\in\mathbb{N}} \subset C_b(X)$ and $(\psi_j)_{j\in\mathbb{N}} \subset C_b(Y)$ fully characterize μ and ν, respectively. Let I^n be the minimum in (25) with moment functions $(\varphi_i)_{i=1}^n$ and $(\psi_j)_{j=1}^n$, and let I be the minimum in (24). If I is finite, then $I^n \nearrow I$ as $n \to +\infty$. Moreover, if π^n minimizes (25) for every $n \in \mathbb{N}$, then, up to a subsequence, π^n converges narrowly to a minimizer π of (24).*

Proof. By Theorem 2.8 and the fundamental theorem of Γ-convergence, it suffices to show that $\Pi_A^\infty = \Pi(\mu,\nu)$ and that the assumption of Theorem 2.8 holds. The families of moment functions being increasing, the sets $\Pi_A^n(\mu,\nu)$ are nested, which also gives the monotonicity of $(I^n)_{n\in\mathbb{N}}$. Moreover, these sets are narrowly closed: if $\pi_k \in \Pi_A^{n_k}(\mu,\nu)$ with $n_k \to +\infty$ converge narrowly to π, then $\pi \in \Pi_A^N(\mu,\nu)$ for every N, that is, $\pi \in \Pi_A^\infty$. The choice of Θ and A guarantees that $\Pi(\mu,\nu) \subseteq \Pi_A^n(\mu,\nu)$ for every $n \in \mathbb{N}$. Conversely, let $\pi \in \bigcap_{n\in\mathbb{N}} \Pi_A^n(\mu,\nu)$. Then, for every $i,j \in \mathbb{N}$,

$$\int_{X\times Y} \varphi_i(x)\ \mathrm{d}\pi = \int_X \varphi_i\ \mathrm{d}\mu, \qquad \int_{X\times Y} \psi_j(y)\ \mathrm{d}\pi = \int_Y \psi_j\ \mathrm{d}\nu.$$

Since these families fully characterize μ and ν, we obtain $(p_X)_\#\pi = \mu$ and $(p_Y)_\#\pi = \nu$. Therefore,

$$\Pi_A^\infty = \bigcap_{n\in\mathbb{N}} \Pi_A^n(\mu,\nu) = \Pi(\mu,\nu).$$

□

Increasing the number of moment constraints does not by itself ensure that the prescribed marginals are recovered in the limit. As the following example shows, two distinct probability measures may have the same moments against every function in an infinite family.

Example 2.14. Consider $X = [-1,1]$ and the moment functions $\phi_0(x) = 1$ and $\phi_k(x) = x^{2k-1}$ for $k \in \mathbb{N}$. We have that

$$\int \phi_k\,\mathrm{d}(\delta_0) = \int \phi_k\,\mathrm{d}\big(\frac{1}{2}(\delta_{-1}+\delta_1)\big) \text{ for every } k \in \mathbb{N}$$

and the measures are not equal.

The following lemma provides a sufficient condition for full characterization.

Lemma 2.15. *Let X be a locally compact Polish space and let $(\varphi_i)_{i\in\mathbb{N}} \subset C_b(X)$. If*

$$C_0(X) \subseteq \overline{\operatorname{span}\{\varphi_i : i \in \mathbb{N}\}}^{\|\cdot\|_\infty},$$

then this family fully characterizes every finite nonnegative Borel measure on X.

Proof. Let $\mu,\eta \in \mathcal{M}_+(X)$ satisfy $\int_X \varphi_i\ \mathrm{d}\mu = \int_X \varphi_i\ \mathrm{d}\eta$ for every $i \in \mathbb{N}$. Given $f \in C_0(X)$, choose $g_k \in \operatorname{span}\{\varphi_i : i \in \mathbb{N}\}$ such that $\|f - g_k\|_\infty \to 0$. Since the integrals of g_k against μ and η coincide, we have

$$\left|\int_X f\ \mathrm{d}\mu - \int_X f\ \mathrm{d}\eta\right| \le \big(\mu(X)+\eta(X)\big)\|f-g_k\|_\infty \longrightarrow 0.$$

Thus μ and η agree on every function in $C_0(X)$ and thus $\mu = \eta$. □

Let us conclude this discussion of Definition 2.12 by noting that, even when this condition fails, one still obtains an interesting problem, which we call **optimal transport between equivalence classes**.

Definition 2.16 (Equivalence class)**.** For a probability measure $\mu \in \mathcal{P}(X)$, we define its equivalence class with respect to a family of moment functions $\mathcal{I} = \operatorname{span}(\{\varphi_i\}_i)$ as

$$[\mu]_{\mathcal{I}} \coloneqq \left\{\eta \in \mathcal{M}_+(X) \;\middle|\; \int_X \varphi\,\mathrm{d}\eta = \int_X \varphi\,\mathrm{d}\mu, \text{ for all } \varphi \in \mathcal{I}\right\}.$$

Definition 2.17 (Optimal transport between equivalence classes)**.** The optimal transport between equivalence classes with confinement bound A is the following minimization problem

$$\inf_{\substack{\pi\in\Pi([\mu]_{\mathcal{I}},[\nu]_{\mathcal{J}})\\ \int_{X\times Y}\Theta\,\mathrm{d}\pi\le A}}\int_{X\times Y}c(x,y)\,\mathrm{d}\pi(x,y), \tag{27}$$

where $\Pi([\mu]_{\mathcal{I}},[\nu]_{\mathcal{J}})$ is the set of all couplings between $[\mu]_{\mathcal{I}}$ and $[\nu]_{\mathcal{J}}$ for $\mathcal{I}=\operatorname{span}(\{\varphi_i\}_i)$ and $\mathcal{J}=\operatorname{span}(\{\psi_j\}_j)$. In particular, we say that a measure $\pi\in\mathcal{P}(X\times Y)$ is a coupling between $[\mu]_{\mathcal{I}}$ and $[\nu]_{\mathcal{J}}$ if there exist representatives $\tilde{\mu}\in[\mu]_{\mathcal{I}}$ and $\tilde{\nu}\in[\nu]_{\mathcal{J}}$ such that $\pi\in\Pi(\tilde{\mu},\tilde{\nu})$.

Notice that (25) is a special case of (27) when $\mathcal{I}$ and $\mathcal{J}$ are finite-dimensional. Moreover, (27) is the limit problem for (25) when the moment functions do not fully characterize the marginals.

Corollary 2.18. *Under the assumptions of Theorem 2.13, assume that* (24) *is finite and that Theorem 2.10 applies for every $n\in\mathbb{N}$. Then*

$$\min_{\pi\in\Pi(\mu,\nu)}\int_{X\times Y}c\,\mathrm{d}\pi=\sup_{\substack{\varphi\in C_b(X),\ \psi\in C_b(Y)\\ \varphi(x)+\psi(y)\le c(x,y)}}\left\{\int_X\varphi\,\mathrm{d}\mu+\int_Y\psi\,\mathrm{d}\nu\right\}.$$

Proof. Denote the dual supremum by D. Weak duality gives $D\le I$. Let $(\varphi_n,\psi_n,\tau_n)$ be a maximizer in (26) and define

$$f_n=\varphi_n-\tau_n\theta_X,\qquad g_n=\psi_n-\tau_n\theta_Y.$$

Then $f_n(x)+g_n(y)\le c(x,y)$ and

$$\int_X f_n\,\mathrm{d}\mu+\int_Y g_n\,\mathrm{d}\nu=I^n+\tau_n\left(A-\int_X\theta_X\,\mathrm{d}\mu-\int_Y\theta_Y\,\mathrm{d}\nu\right)\ge I^n.$$

The functions f_n and g_n are continuous and bounded above. Choose $B_n\ge 0$ bounding both functions above, a lower bound b for c, and $M_n\ge 0$ such that $B_n-M_n\le b$. The bounded continuous functions

$$\widehat{f}_n=\max\{f_n,-M_n\},\qquad \widehat{g}_n=\max\{g_n,-M_n\}$$

remain dual admissible, and, since truncation from below can only increase the integrals, $D\ge I^n$. Letting $n\to\infty$ gives $D\ge I$, which concludes the proof. □

3. Estimates for the W_∞ distance

As Example 1.1 shows, matching finitely many moments can preserve the optimal value while allowing very different marginal supports. We therefore study which geometric information is retained by localized moment constraints, defined below. Our first result is Lemma 3.2, which shows that, under this choice of moment functions, the support of the moment-relaxed marginals stays quantitatively close to the support of the original marginal. We also establish a W_p–W_∞ estimate, Theorem 3.4, in the same spirit as [20, 7] with an additive error $4r_n$, showing how average transport accuracy controls the maximal displacement needed to couple the relaxed marginals.

Assumption 3.1 (Localized moment functions)**.** *The measure $\mu\in\mathcal{P}(X)$ has compact and connected support and $\nu\in\mathcal{P}(\operatorname{supp}(\mu))$. For fixed $n\in\mathbb{N}$, the nodes $x_1,\dots,x_n\in\operatorname{supp}(\mu)$ and $r_n>0$ are such that the balls $B(x_i,r_n)$ cover $K=\operatorname{supp}(\mu)$, and the moment functions $\varphi_1,\dots,\varphi_n\in C_b(X)$ are nonnegative and form a partition of unity on K, subordinate to this cover, that is,*

$$\sum_{i=1}^n\varphi_i=1\quad\text{on } K,\qquad \sum_{i=1}^n\varphi_i\le 1\quad\text{on } X,\qquad \operatorname{supp}(\varphi_i)\subset B(x_i,r_n),\quad i=1,\dots,n. \tag{28}$$

Throughout this section, we work under Assumption 3.1. The parameter r_n gives the localization scale of the moment functions. For $\lambda \in \mathcal{P}(\mathrm{supp}(\mu))$, we denote by $[\lambda]_n$ the set of probability measures on X having the same integrals as λ against $\varphi_1, \dots, \varphi_n$, and we set

$$m_\mu(s) \coloneqq \inf_{z \in \mathrm{supp}(\mu)} \mu(B(z,s)), \qquad s > 0.$$

We also write $A^r = \{x \in X : d(x,A) \le r\}$ for the closed r-neighborhood of a set A.

Lemma 3.2. *Under Assumption 3.1, for every $\lambda \in \mathcal{P}(\mathrm{supp}(\mu))$ and $\lambda_n \in [\lambda]_n$,*

$$W_\infty(\lambda, \lambda_n) \le 2r_n, \qquad \mathrm{supp}(\lambda_n) \subset \mathrm{supp}(\mu)^{r_n}.$$

Proof. Set $S = \sum_{i=1}^n \varphi_i$. Since $S = 1$ on $\mathrm{supp}(\mu)$, the moment constraints give $\int_X (1 - S)\, \mathrm{d}\lambda_n = 0$. The function $1 - S$ is continuous and nonnegative, so it vanishes on $\mathrm{supp}(\lambda_n)$. Thus, for every $y \in \mathrm{supp}(\lambda_n)$, there exists i with $\varphi_i(y) > 0$. The support condition on φ_i then gives $d(y, x_i) < r_n$, proving the support inclusion.

To estimate the displacement, set $a_i = \int_X \varphi_i \, \mathrm{d}\lambda$ and define

$$\gamma(\mathrm{d}x, \mathrm{d}y) = \sum_{i : a_i > 0} \frac{\varphi_i(x)\varphi_i(y)}{a_i} \lambda(\mathrm{d}x)\lambda_n(\mathrm{d}y).$$

The moment identities and the equality $S = 1$ almost everywhere for both measures show that $\gamma \in \Pi(\lambda, \lambda_n)$. Each summand is supported on $\mathrm{supp}(\varphi_i) \times \mathrm{supp}(\varphi_i)$, where $d(x,y) \le 2r_n$. Hence $W_\infty(\lambda, \lambda_n) \le 2r_n$. □

Lemma 3.3. *Assume Assumption 3.1. Let $A \subset X$ be closed and let $R > 0$. Suppose that there exist $z_0, z_1 \in \mathrm{supp}(\mu)$ such that*

$$d(z_0, A) \le r_n, \qquad d(z_1, A) > R - r_n.$$

Then, for every $s > 0$ satisfying $R > 4r_n + 3s$ and every $\mu_n \in [\mu]_n$,

$$\mu_n(A^{R-s} \setminus A) \ge m_\mu(s).$$

Proof. Choose $0 < \eta < R - 4r_n - 3s$. The number $2r_n + s + \eta$ lies between r_n and $R - r_n$. Since $\mathrm{supp}(\mu)$ is connected and the distance to A is continuous, there exists $z \in \mathrm{supp}(\mu)$ such that $d(z, A) = 2r_n + s + \eta$.

By the construction in Lemma 3.2, there exists $\gamma \in \Pi(\mu, \mu_n)$ with displacement at most $2r_n$. For γ-almost every pair (x, y) with $x \in B(z, s)$, we have

$$\begin{aligned} d(y, z) &\le d(y, x) + d(x, z) < 2r_n + s, \\ |d(y, A) - d(z, A)| &\le d(y, z), \end{aligned}$$

where the second inequality follows because the distance to A is 1-Lipschitz. Consequently, $\eta < d(y, A) < 4r_n + 2s + \eta < R - s$, so $y \in A^{R-s} \setminus A$.

Thus all the mass of μ in $B(z, s)$ is transported into $A^{R-s} \setminus A$, and

$$\begin{aligned} \mu_n(A^{R-s} \setminus A) &\ge \gamma\big(B(z,s) \times (A^{R-s} \setminus A)\big) \\ &= \mu(B(z,s)) \ge m_\mu(s). \end{aligned}$$

□

Theorem 3.4. *Under Assumption 3.1, for every $1 \le p < \infty$, $s > 0$ and $(\mu_n, \nu_n) \in [\mu]_n \times [\nu]_n$,*

$$W_p^p(\mu_n, \nu_n) < m_\mu(s) s^p \quad \Longrightarrow \quad W_\infty(\mu_n, \nu_n) \le 4r_n + 3s.$$

Proof. Fix $s > 0$ such that $W_p^p(\mu_n, \nu_n) < m_\mu(s)s^p$, and let $\pi^n \in \Pi(\mu_n, \nu_n)$ be optimal for W_p. The transport cost bounds the mass displaced by at least s:

$$\begin{aligned} \pi^n(\{d(x,y) \geq s\}) &\leq \frac{1}{s^p} \int_{X\times X} d(x,y)^p \, \mathrm{d}\pi^n \\ &= \frac{W_p^p(\mu_n, \nu_n)}{s^p} < m_\mu(s). \end{aligned} \tag{29}$$

We argue by contradiction. Suppose that $W_\infty(\mu_n, \nu_n) > 4r_n + 3s$, and choose $4r_n + 3s < R < W_\infty(\mu_n, \nu_n)$. By Strassen's theorem (see [30, Theorem 1.27 and Corollary 1.28]), there exists a closed set $A \subset X$ such that

$$\mu_n(A) > \nu_n(A^R). \tag{30}$$

In particular, $\mu_n(A) > 0$ and $\nu_n((A^R)^c) > 0$. By Lemma 3.2, both measures are supported in $\mathrm{supp}(\mu)^{r_n}$. We can therefore find $z_0, z_1 \in \mathrm{supp}(\mu)$ with $d(z_0, A) \leq r_n$ and $d(z_1, A) > R - r_n$. Applying Lemma 3.3 gives

$$\mu_n(A^{R-s} \setminus A) \geq m_\mu(s). \tag{31}$$

On the other hand, if $x \in A^{R-s}$ and $y \notin A^R$, then $d(x,y) > s$. Using the marginals of π^n and (31), we obtain

$$\begin{aligned} &\mu_n(A^{R-s}) \leq \nu_n(A^R) + \pi^n(\{d(x,y) \geq s\}), \\ &\mu_n(A^{R-s}) = \mu_n(A) + \mu_n(A^{R-s} \setminus A) \geq \mu_n(A) + m_\mu(s). \end{aligned}$$

Combining these inequalities with (29) yields $\mu_n(A) < \nu_n(A^R)$, contradicting (30). □

The next corollary is similar to [7, Theorem 1.2] and to the Euclidean case of [20, Theorem 1.1]. Under the assumptions below, $m_\mu(t) \geq ct^d$ for sufficiently small $t > 0$. Let us highlight that both μ_n and ν_n are allowed to be discrete measures, while the lower mass bound is imposed on the reference measure μ. Thus, this corollary provides an analogue of these L^∞ estimates for discrete measures satisfying localized moment constraints, with an additive error $4r_n$.

Corollary 3.5. *Let $X = \mathbb{R}^d$. Under Assumption 3.1, let $\Omega \subset \mathbb{R}^d$ be a bounded, connected domain with Lipschitz boundary and $\mu = f\mathbf{1}_\Omega \, \mathrm{d}x$, where $f \geq a > 0$ almost everywhere on Ω. Then there exists $c_{p,d,a,\Omega} > 0$ such that*

$$W_\infty(\mu_n, \nu_n) \leq 4r_n + c_{p,d,a,\Omega} W_p^{\frac{p}{p+d}}(\mu_n, \nu_n)$$

for every $(\mu_n, \nu_n) \in [\mu]_n \times [\nu]_n$.

Proof. Here $\mathrm{supp}(\mu) = \overline{\Omega}$. Set $D = \mathrm{diam}(\mathrm{supp}(\mu))$. The Lipschitz boundary condition gives a constant $c_\Omega > 0$ such that $|\Omega \cap B(z,s)| \geq c_\Omega s^d$ for every $z \in \mathrm{supp}(\mu)$ and $0 < s \leq D$. Since $f \geq a$, it follows that $m_\mu(s) \geq ac_\Omega s^d$ on this range.

Set

$$t = (ac_\Omega)^{-\frac{1}{p+d}} W_p^{\frac{p}{p+d}}(\mu_n, \nu_n).$$

If $t < D$, then every $s \in (t, D]$ satisfies $W_p^p(\mu_n, \nu_n) < ac_\Omega s^{p+d} \leq m_\mu(s)s^p$. By Theorem 3.4, $W_\infty(\mu_n, \nu_n) \leq 4r_n + 3s$. Letting $s \downarrow t$ gives the desired estimate.

If $t \geq D$, both measures are supported in $\mathrm{supp}(\mu)^{r_n}$, whose diameter is at most $D + 2r_n$. Hence $W_\infty(\mu_n, \nu_n) \leq D + 2r_n \leq 4r_n + 3t$. The result follows in both cases with $c_{p,d,a,\Omega} = 3(ac_\Omega)^{-1/(p+d)}$. □

Remark 3.6. In particular, if $r_n \to 0$ as $n \to +\infty$, since $c_{p,d,a,\Omega}$ does not depend on n, one recovers the inequality originally proven in [7, 20]. Note that our strategy is similar to Lemma 2.2 in [20], but we obtain an improvement on the underlying constants.

4. Estimates for repulsive costs

In Section 3, the moment functions were localized and the estimates concerned the geometry of the relaxed marginals. We now turn to multi-marginal problems with repulsive costs, such as the Coulomb cost, for which the relevant question is whether minimizers of the relaxed problem stay away from the diagonal, where the cost is singular. Here the moment functions need not be localized. We prove a diagonal exclusion estimate which is uniform in n (Theorem 4.3 and Corollary 4.5), and we derive two consequences: the cost can be truncated without changing the minimizers (Theorem 4.6) and, for localized moment functions, the dual problem admits potentials with explicit bounds (Theorem 4.8).

Fix $M \geq 2$ and let X be a Polish space endowed with a distance d. We consider costs of the form

$$c(x_1, \dots, x_M) = \sum_{1 \leq i < j \leq M} h(d(x_i, x_j)), \tag{32}$$

where $h : (0, +\infty) \to [0, +\infty)$ is strictly decreasing and continuous, with $\lim_{r \to 0^+} h(r) = +\infty$, and we set $h(0) = +\infty$. Let $\rho \in \mathcal{P}(X)$. For each $n \in \mathbb{N}$, we take the same family $\mathcal{I}^n = \{\varphi_i\}_{i=1}^n \subset C_b(X)$ of moment functions for all M marginals; these moment functions need not form a partition of unity. We denote by $\Pi(\rho)$ the set of probability measures on X^M whose marginals are all equal to ρ, and by $\Pi^n(\rho)$ the set of probability measures on X^M whose marginals have the same integrals as ρ against every function in $\mathcal{I}^n$. To guarantee the existence of minimizers, we also impose a confinement constraint. Let $\theta : X \to [0, +\infty)$ be continuous with compact sublevel sets, assume $\int_X \theta \, \mathrm{d}\rho < +\infty$, and choose $A > 0$ such that $A \geq M \int_X \theta \, \mathrm{d}\rho$. Set $\Theta(x_1, \dots, x_M) = \theta(x_1) + \cdots + \theta(x_M)$ and define for every $n \in \mathbb{N}$

$$\Pi^n_A(\rho) \coloneqq \Big\{ \pi \in \Pi^n(\rho) : \int_{X^M} \Theta \, \mathrm{d}\pi \leq A \Big\}, \tag{33}$$

$$I^n(\rho) \coloneqq \inf_{\pi \in \Pi^n_A(\rho)} \mathrm{C}(\pi), \tag{34}$$

and

$$I(\rho) \coloneqq \inf_{\pi \in \Pi(\rho)} \mathrm{C}(\pi), \tag{35}$$

where, as before, $\mathrm{C}(\pi) = \int_{X^M} c \, \mathrm{d}\pi$. Problem (34) is (2) with $Z = X^M$, $m = 1$ and moment functions $\pm\big(\varphi_i \circ p_j - \int_X \varphi_i \, \mathrm{d}\rho\big)$, $1 \leq i \leq n$, $1 \leq j \leq M$, where p_j denotes the projection onto the j-th factor. For $\alpha > 0$, we set

$$D_\alpha \coloneqq \big\{ (x_1, \dots, x_M) \in X^M : d(x_i, x_j) < \alpha, \text{ for some } i \neq j \big\}, \qquad \kappa_\rho(\alpha) \coloneqq \sup_{x \in X} \rho(B(x, \alpha)).$$

Proposition 4.1. *Let $\pi \in \Pi^n_A(\rho)$ satisfy $\int_{X^M} c \, \mathrm{d}\pi < +\infty$. Then there exists a finitely supported sequence $\pi_k \in \Pi^n_A(\rho)$ converging narrowly to π such that, for every $k \in \mathbb{N}$,*

$$\mathrm{supp}(\pi_k) \subset \mathrm{supp}(\pi), \ \#\mathrm{supp}(\pi_k) \leq Mn + k + 3,$$

$$\int_{X^M} c \, \mathrm{d}\pi_k = \int_{X^M} c \, \mathrm{d}\pi, \ \int_{X^M} \Theta \, \mathrm{d}\pi_k = \int_{X^M} \Theta \, \mathrm{d}\pi$$

and $\pi_k(D_{\delta_k}) = 0$, where $\delta_k > 0$ and $\delta_k \to 0$ as $k \to +\infty$.

Proof. Choose a countable family $\{w_j\}_{j \in \mathbb{N}} \subset C_b(X^M)$ fully characterizing π in the sense of Definition 2.12. The finiteness of the cost implies that π is concentrated on the Borel set $E = \mathrm{supp}(\pi) \cap \{c < +\infty\}$. Define $\widetilde{c}$ to agree with c on $\{c < +\infty\}$ and to equal zero elsewhere. For each $k \in \mathbb{N}$, consider the following set of moment functions

$$\bigcup_{j=1}^M (\{\varphi_i \circ p_j\}_{i=1}^n) \cup \{1, \Theta, \widetilde{c}, w_1, \dots, w_k\}. \tag{36}$$

By Theorem 2.2, there exist $1 \le L_k \le Mn + k + 3$, distinct points $x_{k,\ell} \in E$ and positive weights $a_{k,\ell}$ such that $\pi_k = \sum_{\ell=1}^{L_k} a_{k,\ell}\delta_{x_{k,\ell}}$ and π_k has the same integral as π with respect to all functions in (36), in particular $\pi_k \in \Pi_A^n(\rho)$. Since the points $x_{k,\ell} = (x_{k,\ell}^1, \dots, x_{k,\ell}^M) \in E$ it follows that $x_{k,\ell}^i \neq x_{k,\ell}^j$ for $i \neq j$ and every $1 \le \ell \le L_k$.

As the set Π_A^n is narrowly compact, the sequence $(\pi_k)_{k\in\mathbb{N}}$ is tight and, by full characterization, the limit of any converging subsequence, and thus of $(\pi_k)_{k\in\mathbb{N}}$, is π.

Finally, since π_k has finite support and every point in its support has pairwise distinct coordinates, we may choose $0 < \delta_k \le 1/k$ such that $\pi_k(D_{\delta_k}) = 0$. □

Remark 4.2. Proposition 4.1 shows that finite cost alone yields finitely supported approximations which give no mass to D_{δ_k}, but $\delta_k \to 0$, so that no positive exclusion radius is obtained for π itself. For minimizers, under a condition on κ_ρ, Theorem 4.3 gives a positive exclusion radius, which by Corollary 4.5 is independent of n. We also notice that the proof of Proposition 4.1 only relies on Theorem 2.2, and therefore applies to any finite family of moment functions in $C_b(X^M)$; this will be used in Section 5.

Assume henceforth that $I(\rho) < +\infty$.

Theorem 4.3. *If $\kappa_\rho(\alpha) < 1/(2(M-1))$ and $\pi^n \in \Pi_A^n(\rho)$ is optimal in (34) then,*

$$\pi^n(D_{\alpha_n}) = 0$$

where $\alpha_n > 0$ is such that

$$h(\alpha_n) = \frac{I(\rho) - I^n(\rho)}{1 - 2(M-1)\kappa_\rho(\alpha)} + 2(M-1)h(\alpha).$$

Proof. The definition of α_n and the strict decrease of h give $0 < \alpha_n < \alpha$. Suppose initially that π^n is finitely supported. Assume that $\pi^n(D_{\alpha_n}) > 0$. We will show that this allows us to construct a better competitor for (34), contradicting the optimality of π^n. Let $\bar{x} = (\bar{x}_1, \dots, \bar{x}_M) \in \operatorname{supp}(\pi^n) \cap D_{\alpha_n}$ and let $a = \pi^n(\{\bar{x}\})$. After relabeling the coordinates, we may assume that $d(\bar{x}_1, \bar{x}_2) < \alpha_n$. Let $\pi^\star$ be optimal for (35) and define the "good" set

$$\mathcal{G}_{\bar{x}} = \Big\{(z_1, \dots, z_M) \in X^M : d(z_1, \bar{x}_j) \ge \alpha, \quad d(\bar{x}_1, z_j) \ge \alpha \text{ for } j = 2, \dots, M\Big\}.$$

We have the following estimate for the mass of $\mathcal{G}_{\bar{x}}$:

$$\begin{aligned}
\pi^\star\big(X^M \setminus \mathcal{G}_{\bar{x}}\big) &= \pi^\star\left(\bigcup_{j=2}^{M}\Big[\big(B(\bar{x}_j, \alpha) \times X^{M-1}\big) \cup \big(X^{j-1} \times B(\bar{x}_1, \alpha) \times X^{M-j}\big)\Big]\right) \\
&\le \sum_{j=2}^{M}\Big[\rho\big(B(\bar{x}_j, \alpha)\big) + \rho\big(B(\bar{x}_1, \alpha)\big)\Big] \\
&\le 2(M-1)\kappa_\rho(\alpha).
\end{aligned}$$

Hence, from $\kappa_\rho(\alpha) < 1/(2(M-1))$,

$$\pi^\star(\mathcal{G}_{\bar{x}}) \ge 1 - 2(M-1)\kappa_\rho(\alpha) > 0. \tag{37}$$

Define the following transport plan:

$$\pi_\epsilon = (1-\epsilon)\pi^n + \epsilon\pi^\star, \qquad \epsilon \in (0,1). \tag{38}$$

Clearly, $\pi_\epsilon \in \Pi_A^n(\rho)$. Moreover, we can explicitly estimate its cost:

$$\int_{X^M} c \, \mathrm{d}\pi_\epsilon = (1-\epsilon)I^n(\rho) + \epsilon I(\rho) = I^n(\rho) + \epsilon(I(\rho) - I^n(\rho)). \tag{39}$$

Since $d(\bar{x}_1, \bar{x}_2) < \alpha_n < \alpha$, we have $(\bar{x}_1, \dots, \bar{x}_M) \notin \mathcal{G}_{\bar{x}}$. Choose $\epsilon \in (0,1)$ such that $\epsilon\pi^\star(\mathcal{G}_{\bar{x}}) \le (1-\epsilon)a$ and define

$$\tilde{\pi}_\epsilon = \pi_\epsilon - \epsilon\pi^\star(\mathcal{G}_{\bar{x}})\delta_{(\bar{x}_1,\dots,\bar{x}_M)} - \epsilon\pi^\star|_{\mathcal{G}_{\bar{x}}} + \epsilon(T_1)_\#\big(\pi^\star|_{\mathcal{G}_{\bar{x}}}\big) + \epsilon(T_2)_\#\big(\pi^\star|_{\mathcal{G}_{\bar{x}}}\big),$$

where $T_1, T_2 : \mathcal{G}_{\bar{x}} \to X^M$ are given by $T_1(z_1, \dots, z_M) = (\bar{x}_1, z_2, \dots, z_M)$ and $T_2(z_1, \dots, z_M) = (z_1, \bar{x}_2, \dots, \bar{x}_M)$.

Observe that $\tilde{\pi}_\epsilon$ is a nonnegative measure, since $\epsilon\pi^\star(\mathcal{G}_{\bar{x}})\delta_{(\bar{x}_1,\dots,\bar{x}_M)} + \epsilon\big(\pi^\star|_{\mathcal{G}_{\bar{x}}}\big)$ is a submeasure of π_ϵ by our choice of ϵ. Moreover, $\tilde{\pi}_\epsilon$ is a probability measure. We also have

$$\begin{aligned}(p_1)_\#\tilde{\pi}_\epsilon &= (p_1)_\#\pi_\epsilon - \epsilon\pi^\star(\mathcal{G}_{\bar{x}})\delta_{\bar{x}_1} - \epsilon(p_1)_\#\big(\pi^\star|_{\mathcal{G}_{\bar{x}}}\big)\\ &\quad + \epsilon(p_1)_\#(T_1)_\#\big(\pi^\star|_{\mathcal{G}_{\bar{x}}}\big) + \epsilon(p_1)_\#(T_2)_\#\big(\pi^\star|_{\mathcal{G}_{\bar{x}}}\big)\\ &= (p_1)_\#\pi_\epsilon - \epsilon\pi^\star(\mathcal{G}_{\bar{x}})\delta_{\bar{x}_1} - \epsilon(p_1)_\#\big(\pi^\star|_{\mathcal{G}_{\bar{x}}}\big)\\ &\quad + \epsilon\pi^\star(\mathcal{G}_{\bar{x}})\delta_{\bar{x}_1} + \epsilon(p_1)_\#\big(\pi^\star|_{\mathcal{G}_{\bar{x}}}\big)\\ &= (p_1)_\#\pi_\epsilon.\end{aligned} \tag{40}$$

and similarly $(p_j)_\#\tilde{\pi}_\epsilon = (p_j)_\#\pi_\epsilon$ for $j = 2, \dots, M$. Since $\tilde{\pi}_\epsilon$ and π_ϵ have the same marginals, they have the same moments against $\varphi_i \circ p_j$ for every $i = 1, \dots, n$ and $j = 1, \dots, M$. Moreover, since $\Theta(x_1, \dots, x_M) = \theta(x_1) + \cdots + \theta(x_M)$, we have

$$\int_{X^M} \Theta \,\mathrm{d}\tilde{\pi}_\epsilon = \int_{X^M} \Theta \,\mathrm{d}\pi_\epsilon \le A.$$

Therefore $\tilde{\pi}_\epsilon \in \Pi_A^n(\rho)$. Integrating the cost c against the measure $\tilde{\pi}_\epsilon$ we obtain

$$\begin{aligned}&\int_{X^M} c \,\mathrm{d}\tilde{\pi}_\epsilon\\ &\quad= \int_{X^M} c \,\mathrm{d}\pi_\epsilon - \epsilon\pi^\star(\mathcal{G}_{\bar{x}}) \int_{X^M} c \,\mathrm{d}\delta_{(\bar{x}_1,\dots,\bar{x}_M)} - \epsilon \int_{\mathcal{G}_{\bar{x}}} c \,\mathrm{d}\pi^\star\\ &\qquad + \epsilon \int_{\mathcal{G}_{\bar{x}}} c \circ T_1 \,\mathrm{d}\pi^\star + \epsilon \int_{\mathcal{G}_{\bar{x}}} c \circ T_2 \,\mathrm{d}\pi^\star\\ &\quad= \int_{X^M} c \,\mathrm{d}\pi_\epsilon - \epsilon \int_{\mathcal{G}_{\bar{x}}} c(\bar{x}_1, \dots, \bar{x}_M) \,\mathrm{d}\pi^\star - \epsilon \int_{\mathcal{G}_{\bar{x}}} c(z_1, \dots, z_M) \,\mathrm{d}\pi^\star\\ &\qquad + \epsilon \int_{\mathcal{G}_{\bar{x}}} c(\bar{x}_1, z_2, \dots, z_M) \,\mathrm{d}\pi^\star + \epsilon \int_{\mathcal{G}_{\bar{x}}} c(z_1, \bar{x}_2, \dots, \bar{x}_M) \,\mathrm{d}\pi^\star\\ &\quad= \int_{X^M} c \,\mathrm{d}\pi_\epsilon + \epsilon \sum_{j=2}^{M} \int_{\mathcal{G}_{\bar{x}}} [-h(d(\bar{x}_1, \bar{x}_j)) - h(d(z_1, z_j))\\ &\qquad\qquad + h(d(\bar{x}_1, z_j)) + h(d(z_1, \bar{x}_j))] \,\mathrm{d}\pi^\star.\end{aligned} \tag{41}$$

Since $d(\bar{x}_1, \bar{x}_2) < \alpha_n$ and for $(z_1, \dots, z_M) \in \mathcal{G}_{\bar{x}}$ we have $d(z_1, \bar{x}_j) \ge \alpha$ and $d(\bar{x}_1, z_j) \ge \alpha$ for every $j = 2, \dots, M$, we can estimate the sum of the integrands as

$$\begin{aligned}\sum_{j=2}^{M} \Big[-h(d(\bar{x}_1, \bar{x}_j)) - h(d(z_1, z_j)) + h(d(\bar{x}_1, z_j)) + h(d(z_1, \bar{x}_j))\Big]&\\ \le -h(d(\bar{x}_1, \bar{x}_2)) + 2(M-1)h(\alpha)&\\ < -h(\alpha_n) + 2(M-1)h(\alpha).&\end{aligned} \tag{42}$$

Substituting (42) and (39) into (41) we obtain

$$\begin{aligned}\int_{X^M} c \,\mathrm{d}\tilde{\pi}_\epsilon &< \int_{X^M} c \,\mathrm{d}\pi_\epsilon + \epsilon\pi^\star(\mathcal{G}_{\bar{x}}) \left(-h(\alpha_n) + 2(M-1)h(\alpha)\right)\\ &\le I^n(\rho) + \epsilon \left[I(\rho) - I^n(\rho) + \pi^\star(\mathcal{G}_{\bar{x}}) \left(-h(\alpha_n) + 2(M-1)h(\alpha)\right)\right].\end{aligned} \tag{43}$$

Since, by assumption, $h(\alpha_n) = \dfrac{I(\rho) - I^n(\rho)}{1 - 2(M-1)\kappa_\rho(\alpha)} + 2(M-1)h(\alpha)$, combining this with (37) yields

$$\begin{aligned}\int_{X^M} c\,\mathrm{d}\tilde{\pi}_\epsilon &< I^n(\rho) + \epsilon\left[I(\rho) - I^n(\rho) - \pi^\star(\mathcal{G}_{\bar{x}})\left(\frac{I(\rho)-I^n(\rho)}{1-2(M-1)\kappa_\rho(\alpha)}\right)\right]\\ &\leq I^n(\rho) + \epsilon\left[I(\rho) - I^n(\rho) - (1-2(M-1)\kappa_\rho(\alpha))\left(\frac{I(\rho)-I^n(\rho)}{1-2(M-1)\kappa_\rho(\alpha)}\right)\right]\\ &= I^n(\rho)\end{aligned} \tag{44}$$

This is a contradiction, since $\tilde{\pi}_\epsilon \in \Pi_A^n(\rho)$.

Let now π^n be an arbitrary minimizer. By Proposition 2.6, there exists a sequence of finitely supported minimizers $(\pi^{n;k})_{k\in\mathbb{N}}$ converging narrowly to π^n. The preceding argument gives us that $\pi^{n;k}(D_{\alpha_n}) = 0$ for every $k \in \mathbb{N}$. Since D_{α_n} is open and does not depend on k, we have

$$0 \leq \pi^n(D_{\alpha_n}) \leq \liminf_{k\to\infty} \pi^{n;k}(D_{\alpha_n}) = 0.$$

This concludes the proof. □

Remark 4.4. If $\bigcup_n \mathcal{I}^n$ fully characterizes ρ and the admissible sets Π_A^n satisfy the limit condition in Theorem 2.8, then $I^n(\rho) \to I(\rho)$. Consequently, the exclusion radii in Theorem 4.3 satisfy $h(\alpha_n) \to 2(M-1)h(\alpha)$. Thus, in the limit one recovers the diagonal exclusion radius (see [11, Theorem 1.3 (i)]).

Since $0 \leq I(\rho) - I^n(\rho) \leq I(\rho)$ and h is decreasing, the diagonal exclusion is uniform in n:

Corollary 4.5. *If $\kappa_\rho(\alpha) < 1/(2(M-1))$ for some $\alpha > 0$, set*

$$\alpha^\star = h^{-1}\left(\frac{I(\rho)}{1-2(M-1)\kappa_\rho(\alpha)} + 2(M-1)h(\alpha)\right).$$

Then, for every $n \in \mathbb{N}$, every minimizer π^n of (34) satisfies $\pi^n(D_{\alpha^\star}) = 0$.

For $0 < \bar{\alpha} < \alpha^\star$, define $\bar{h}(r) = \min(h(r), h(\bar{\alpha}))$, the truncated cost function $\bar{c}$ defined as in (32) with respect to the function $\bar{h}$, and $\bar{I}^n(\rho)$ and $\bar{I}(\rho)$ defined as in (34) and (35), respectively, with respect to $\bar{c}$.

Theorem 4.6. *$\pi^n \in \Pi_A^n(\rho)$ is optimal for $I^n(\rho)$ if, and only if, it is optimal for $\bar{I}^n(\rho)$.*

Proof. We first show that every minimizer $\bar{\pi}^n$ of $\bar{I}^n(\rho)$ satisfies $\bar{\pi}^n(D_{\alpha^\star}) = 0$. As in the proof of Theorem 4.3, one may initially assume that $\bar{\pi}^n$ is finitely supported and then pass to the limit. The only change concerns (42): Since $d(\bar{x}_1, \bar{x}_2) < \alpha^\star$ and $\bar{\alpha} < \alpha^\star$ we have $\bar{h}(d(\bar{x}_1, \bar{x}_2)) > h(\alpha^\star)$, while on $\mathcal{G}_{\bar{x}}$, both $\bar{h}(d(\bar{x}_1, z_j))$ and $\bar{h}(d(z_1, \bar{x}_j))$ are at most $h(\alpha)$ for every $j = 2, \ldots, M$. Since $0 \leq \bar{I}(\rho) - \bar{I}^n(\rho) \leq I(\rho)$, the definition of $\alpha^\star$ then yields the same contradiction.

Let π^n and $\bar{\pi}^n$ be optimal for $I^n(\rho)$ and $\bar{I}^n(\rho)$, respectively. Since $\bar{c} \leq c$ pointwise and $\bar{\pi}^n(D_{\alpha^\star}) = 0$, we obtain

$$\int_{X^M} \bar{c}\,\mathrm{d}\bar{\pi}^n \leq \int_{X^M} \bar{c}\,\mathrm{d}\pi^n \leq \int_{X^M} c\,\mathrm{d}\pi^n \leq \int_{X^M} c\,\mathrm{d}\bar{\pi}^n = \int_{X^M} \bar{c}\,\mathrm{d}\bar{\pi}^n,$$

hence all the inequalities are equalities, and the result follows. □

The equivalence obtained in Theorem 4.6 has been used in the literature to obtain bounds on the Kantorovich potential of the corresponding dual problem (see [11, 8]). For the moment-constrained problems considered in this section, a similar result is possible but requires stronger assumptions:

Assumption 4.7. *The space X is compact and $\rho \in \mathcal{P}(X)$ has full support on X. For fixed n, the moment functions $\mathcal{I}^n = \{\varphi_1, \dots, \varphi_n\} \subset C(X)$ satisfy (28), with the compact set taken to be X and nodes $z_1, \dots, z_n$, and $\varphi_i(z_j) = \delta_{ij}$ for every $i, j = 1, \dots, n$.*

The compactness of X permits us to choose $\Theta \equiv 0$. Set $E_n = \operatorname{span}(\mathcal{I}^n)$ and define $U_i \coloneqq \{\varphi_i > 0\}$, $J_i \coloneqq \{j \in \{1, \dots, n\} : U_i \cap U_j \neq \emptyset\}$,

$$R_i \coloneqq \frac{\displaystyle\sum_{j \in J_i} \int_X \varphi_j \, \mathrm{d}\rho}{\displaystyle\int_X \varphi_i \, \mathrm{d}\rho},$$

which measures the total ρ-mass assigned to the moment functions whose positivity set intersects the positivity set of φ_i relative to the mass assigned to φ_i itself. The denominator is positive because ρ has full support and $\varphi_i(z_i) = 1$. Since $\bar{c}$ is continuous on the compact space X^M, the function

$$\omega(t) \coloneqq \sup_{\substack{x_1, x_1', x_2, \dots, x_M \in X \\ d(x_1, x_1') \le t}} |\bar{c}(x_1, x_2, \dots, x_M) - \bar{c}(x_1', x_2, \dots, x_M)|$$

satisfies $\omega(t) \to 0$ as $t \to 0$.

Theorem 4.8. *Under Assumption 4.7, the dual problem to $I^n(\rho)$, namely*

$$\sup_{\substack{u_1, \dots, u_M \in E_n \\ \sum_{j=1}^M u_j(x_j) \le c(x_1, \dots, x_M)}} \left\{ \sum_{j=1}^M \int_X u_j \, \mathrm{d}\rho \right\} \tag{45}$$

admits a maximizer of the form $(u, \dots, u)$ where $u = \sum_{i=1}^n a_i \varphi_i$ and

$$-\frac{\|\bar{c}\|_\infty}{M} - \omega(r_n) R_i \le a_i \le \frac{\|\bar{c}\|_\infty}{M}, \qquad i = 1, \dots, n.$$

Proof. We only treat the case $M = 2$; the case $M > 2$ is analogous. Since $\bar{c}$ is bounded and continuous and $\rho \otimes \rho$ has full support, Theorem 2.10 gives attainment in the dual problem with cost $\bar{c}$. By symmetry of the cost, if (u_1, u_2) is a maximizer, then (u, u) with $u = (u_1 + u_2)/2$ is also a maximizer.

Let (u, u) be such a maximizer, with $u = \sum_{i=1}^n a_i \varphi_i$. Taking $x = y$ in the dual constraint gives $2u(x) \le \bar{c}(x, x) = \|\bar{c}\|_\infty$. Since $u(z_i) = a_i$, we obtain $a_i \le \|\bar{c}\|_\infty / 2$ for every i.

To prove the lower bound, fix $i \in \{1, \dots, n\}$. If $a_i \ge -\|\bar{c}\|_\infty / 2$, there is nothing to prove. Otherwise, define

$$\widetilde{u} = u + \left(-\frac{\|\bar{c}\|_\infty}{2} - a_i \right) \varphi_i.$$

This replaces the coefficient a_i by $-\|\bar{c}\|_\infty / 2$ and leaves the other coefficients unchanged. In particular, $\widetilde{u}(z_i) = -\|\bar{c}\|_\infty / 2$, whereas $\widetilde{u}(z_j) = u(z_j)$ for $j \neq i$.

Let us estimate the error in the admissibility of $(\widetilde{u}, u)$. For $j \neq i$, feasibility of (u, u) gives $\widetilde{u}(z_j) + u(y) \le \bar{c}(z_j, y)$. The same inequality holds for $j = i$, since $\widetilde{u}(z_i) + u(y) = -\|\bar{c}\|_\infty / 2 + u(y) \le 0$ and $\bar{c}$ is nonnegative. Multiplying these inequalities by $\varphi_j(x)$ and summing over j, we obtain

$$\widetilde{u}(x) + u(y) \le \sum_{j=1}^n \bar{c}(z_j, y) \varphi_j(x) \le \bar{c}(x, y) + \omega(r_n),$$

where the last inequality follows from $\operatorname{supp}(\varphi_j) \subset B(z_j, r_n)$. Moreover, $\widetilde{u} = u$ outside U_i, so the error is confined to U_i:

$$\widetilde{u}(x) + u(y) \le \bar{c}(x, y) + \omega(r_n) \chi_{U_i}(x). \tag{46}$$

We correct this error by setting $\zeta_i = \sum_{j\in J_i}\varphi_j$. The partition of unity property gives $0 \le \zeta_i \le 1$ and $\zeta_i = 1$ on U_i. Indeed, whenever $x \in U_i$ and $\varphi_j(x) > 0$, we have $j \in J_i$. Thus $\chi_{U_i} \le \zeta_i$, and (46) shows that $(\widetilde{u} - \omega(r_n)\zeta_i, u)$ is admissible for the dual problem with cost $\bar{c}$.

Comparing its value with that of the maximizer (u, u) gives

$$\begin{aligned} 0 &\ge \int_X \left(\widetilde{u} - \omega(r_n)\zeta_i - u\right) \mathrm{d}\rho \\ &= \left(-\frac{\|\bar{c}\|_\infty}{2} - a_i\right)\int_X \varphi_i \,\mathrm{d}\rho - \omega(r_n)\sum_{j\in J_i}\int_X \varphi_j \,\mathrm{d}\rho. \end{aligned}$$

Dividing by $\int_X \varphi_i \,\mathrm{d}\rho > 0$ yields $a_i \ge -\|\bar{c}\|_\infty/2 - \omega(r_n)R_i$, as required.

Finally, since $\bar{c} \le c$, the pair (u, u) is also admissible for the original dual problem. By Theorem 4.6, its value is $2\int_X u \,\mathrm{d}\rho = \bar{I}^n(\rho) = I^n(\rho)$. Weak duality therefore shows that (u, u) is optimal for the original cost c. □

Remark 4.9. Since the functions φ_i are nonnegative and sum to one, the coefficient bounds imply

$$\|u\|_\infty \le \frac{\|\bar{c}\|_\infty}{M} + \omega(r_n)\max_{1\le i\le n} R_i.$$

For the Coulomb cost, $\|\bar{c}\|_\infty = \binom{M}{2}/\bar{\alpha}$ and $\bar{c}$ is $(M-1)\bar{\alpha}^{-2}$-Lipschitz in each variable. Consequently,

$$\|u\|_\infty \le \frac{M-1}{2\bar{\alpha}} + \frac{(M-1)r_n}{\bar{\alpha}^2}\max_{1\le i\le n} R_i.$$

In particular, for fixed M and $\bar{\alpha}$, if $r_n \to 0$ and the ratios R_i are bounded uniformly in n and i, the error term tends to zero.

5. A semiclassical Γ-convergence result

We now explain how the moment-constrained Coulomb problem of Section 4 and, more generally, transport problems with finitely many moment equalities arise as the zero-order limit of a fermionic Lieb functional. We restrict the discussion to the finite-dimensional moment problem, that is, to one fixed family W^n of moment functions encoding the equality constraints, as specified below. We consider measures of mass one satisfying the confinement bound. In the following, we present the semiclassical analogue of the regularization results for the Lieb and Hohenberg–Kohn functionals in [5, 24, 13], combined with the finite-moment transport framework of [2].

Let $M \ge 2$, $Z = (\mathbb{R}^3)^M$, and let $\mathcal{P}_{\mathrm{sym}}(Z)$ be the narrowly closed set of probability measures invariant under permutations of the particle coordinates. We write $\operatorname{Sym}\pi = (M!)^{-1}\sum_{\sigma\in S_M}\sigma_\#\pi$. The cost c of the preceding sections is now the Coulomb cost

$$c(\mathbf{x}) = \sum_{1\le i<j\le M}\frac{1}{|x_i - x_j|}, \qquad \mathbf{x} = (x_1, \dots, x_M),$$

with the convention $c = +\infty$ whenever two particle coordinates coincide. We assume that $\Theta : Z \to [0, +\infty)$ is continuous, coercive and permutation invariant, and fix $A > 0$.

For the equality constraints, let $g_1, \dots, g_q \in C_b(Z)$ be permutation-invariant moment functions and let $b \in \mathbb{R}^q$ be the prescribed vector. We write

$$\Psi_n(\pi) := \left(\int_Z g_1 \,\mathrm{d}\pi, \dots, \int_Z g_q \,\mathrm{d}\pi\right), \qquad W^n := \{g_i - b_i,\ b_i - g_i :\ i = 1, \dots, q\}.$$

Here n labels the moment family and $q = q_n$ counts the moment equalities. With the moment functions W^n and mass $m = 1$, the inequalities in (3) encode precisely $\Psi_n(\pi) = b$.

In this section the admissible set includes symmetry:

$$\Pi^n_{A,\mathrm{sym}} := \Pi^n_A \cap \mathcal{P}_{\mathrm{sym}}(Z) = \left\{ \pi \in \mathcal{P}_{\mathrm{sym}}(Z) : \Psi_n(\pi) = b,\ \int_Z \Theta \,\mathrm{d}\pi \le A \right\}.$$

We define, on $\mathcal{P}(Z)$,

$$\mathrm{C}^n_{\mathrm{sym}}(\pi) := \begin{cases} \displaystyle\int_Z c \,\mathrm{d}\pi, & \pi \in \Pi^n_{A,\mathrm{sym}}, \\ +\infty, & \pi \notin \Pi^n_{A,\mathrm{sym}}. \end{cases}$$

Thus $\mathrm{C}^n_{\mathrm{sym}}$ agrees with the extended functional C^n of Section 2.5 on $\mathcal{P}_{\mathrm{sym}}(Z)$, and is infinite outside this set. When the moment equalities are prescribed by a common marginal ρ, that is, $g_i = M^{-1}\sum_{j=1}^M \varphi_i \circ p_j$ and $b_i = \int_{\mathbb{R}^3} \varphi_i \,\mathrm{d}\rho$, and Θ is as in Section 4, symmetrization shows that

$$\inf_{\pi \in \Pi^n_{A,\mathrm{sym}}} \mathrm{C}(\pi) = I^n(\rho).$$

This is the symmetric M-particle formulation of the M-marginal value considered in Section 4.

Let $\mathfrak{S}_a$ denote the set of mixed fermionic states on $L^2_a((\mathbb{R}^3)^M)$, let π_Γ be the associated symmetric M-particle law, and let

$$K(\Gamma) := \frac{1}{2}\operatorname{Tr}(-\Delta\Gamma) \in [0, +\infty]$$

be the kinetic energy. For a probability measure π on Z, define the partially minimized kinetic energy

$$T(\pi) := \inf\left\{ K(\Gamma) : \Gamma \in \mathfrak{S}_a,\ \pi_\Gamma = \pi \right\}, \qquad \inf \varnothing := +\infty.$$

The lifted semiclassical functionals are

$$F^\varepsilon_n(\pi) := \varepsilon T(\pi) + \mathrm{C}^n_{\mathrm{sym}}(\pi), \qquad \varepsilon > 0, \tag{47}$$

and their zero-order candidate is $F^0_n := \mathrm{C}^n_{\mathrm{sym}}$. Partial minimization gives the identity

$$\inf_{\pi \in \mathcal{P}(Z)} F^\varepsilon_n(\pi) = \inf_{\substack{\Gamma \in \mathfrak{S}_a \\ \pi_\Gamma \in \Pi^n_{A,\mathrm{sym}}}} \left\{ \varepsilon K(\Gamma) + \int_Z c \,\mathrm{d}\pi_\Gamma \right\} =: \mathcal{L}^\varepsilon_n.$$

Set

$$H_n := \operatorname{aff}\{\Psi_n(\pi) : \pi \in \mathcal{P}_{\mathrm{sym}}(Z)\}.$$

The recovery construction below replaces the atoms of a transport plan by fermionic wave packets, which slightly perturbs the moments. To recover the constraint $\Psi_n(\pi) = b$ exactly, we mix the resulting state with finitely many fixed states whose moment vectors have a convex hull containing a neighborhood of b in H_n (these moment vectors need not be close to b); since only a small mixing weight is needed, the energies are almost unchanged. The following assumption guarantees that such states exist. It is the analogue of Slater's condition in convex optimization: b must lie in the relative interior, within the affine hull H_n, of the moment vectors of admissible finite-energy states since, on the boundary, small errors in the moments could not be corrected in every direction.

Assumption 5.1. *There are finitely many states $\Gamma^0, \dots, \Gamma^J \in \mathfrak{S}_a$ such that*

$$K(\Gamma^j) < +\infty, \qquad \int_Z c \,\mathrm{d}\pi_{\Gamma^j} < +\infty, \qquad \int_Z \Theta \,\mathrm{d}\pi_{\Gamma^j} < A \quad (j = 0, \dots, J).$$

Moreover, $b \in H_n$ and, for some $r > 0$,

$$B(b, r) \cap H_n \subset \operatorname{conv}\left\{ \Psi_n(\pi_{\Gamma^0}), \dots, \Psi_n(\pi_{\Gamma^J}) \right\}.$$

Remark 5.2. Assumption 5.1 requires b to lie in the relative interior, within H_n, of the moment vectors of strictly confined fermionic states with finite energies, so that small errors in the moments can be corrected by convex combinations. It plays the same role as the full-support assumption of Theorem 2.7. Similarly, it holds as soon as there is a state $\Gamma^\circ \in \mathfrak{S}_a$ with finite kinetic and Coulomb energies such that

$$\int_Z \Theta \, \mathrm{d}\pi_{\Gamma^\circ} < A, \qquad \Psi_n(\pi_{\Gamma^\circ}) = b, \qquad \mathrm{supp}(\pi_{\Gamma^\circ}) = Z :$$

localizing Γ° with smooth, permutation-invariant cutoff functions produces finite-energy fermionic states whose moment vectors surround b.

In particular, there are weights $\lambda_j^* \geq 0$, summing to one, such that

$$\Gamma^{\mathrm{ref}} := \sum_{j=0}^{J} \lambda_j^* \Gamma^j, \qquad \sigma := \pi_{\Gamma^{\mathrm{ref}}}, \qquad \Psi_n(\sigma) = b, \quad \int_Z \Theta \, \mathrm{d}\sigma < A.$$

Both energies of Γ^{ref} are finite. We use this reference state below.

We next give the recovery construction for the Γ-convergence result. We use

$$D := \{\mathbf{x} \in Z : \min_{i \neq j} |x_i - x_j| = 0\}, \qquad D_\delta := \{\mathbf{x} \in Z : \min_{i \neq j} |x_i - x_j| < \delta\},$$

and the bounded-Lipschitz distance d_{BL},

$$d_{\mathrm{BL}}(\pi, \pi') := \sup \left\{ \int_Z f \, \mathrm{d}\pi - \int_Z f \, \mathrm{d}\pi' : \ \|f\|_\infty \leq 1, \ \mathrm{Lip}(f) \leq 1 \right\}, \qquad \pi, \pi' \in \mathcal{P}(Z),$$

which metrizes narrow convergence on $\mathcal{P}(Z)$.

Lemma 5.3 (Sparse approximation in energy)**.** *Let $\pi \in \Pi^n_{A,\mathrm{sym}}$ satisfy $\int_Z c \, \mathrm{d}\pi < +\infty$ and $\int_Z \Theta \, \mathrm{d}\pi < A$. Then there are finite symmetric atomic laws*

$$\pi^k = \sum_{\ell=1}^{L_k} \alpha_{k,\ell} \, \mathrm{Sym} \, \delta_{\mathbf{x}_{k,\ell}}$$

with all $\mathbf{x}_{k,\ell} \notin D$, such that

$$d_{\mathrm{BL}}(\pi^k, \pi) \to 0, \qquad \int_Z c \, \mathrm{d}\pi^k = \int_Z c \, \mathrm{d}\pi, \qquad \int_Z \Theta \, \mathrm{d}\pi^k = \int_Z \Theta \, \mathrm{d}\pi, \qquad \Psi_n(\pi^k) = b.$$

Proof. By Remark 4.2, the proof of Proposition 4.1 applies with the moment functions $g_1, \ldots, g_q$ in place of $\varphi_i \circ p_j$. It gives finitely supported measures $\widehat{\pi}^k$, converging narrowly to π, with atoms in $\mathrm{supp}(\pi) \setminus D$ and with the same integrals as π against $g_1, \ldots, g_q$, c and Θ. Since π is symmetric and $g_1, \ldots, g_q$, c and Θ are permutation invariant, the symmetrized measures $\pi^k := \mathrm{Sym} \, \widehat{\pi}^k$ have the same properties, converge narrowly to $\mathrm{Sym} \, \pi = \pi$, and are of the stated form. Finally, d_{BL} metrizes narrow convergence.

□

Lemma 5.4 (Packet representation of a sparse law)**.** *Let $\pi^k = \sum_{\ell=1}^{L_k} \alpha_{k,\ell} \, \mathrm{Sym} \, \delta_{\mathbf{x}_{k,\ell}}$ be as in Lemma 5.3 and set*

$$d_k := \min_{\ell} \min_{i \neq j} |x_{k,\ell,i} - x_{k,\ell,j}| > 0.$$

Fix a real, even $\chi \in C_c^\infty(B(0,1))$ with $\|\chi\|_2 = 1$ and write $\chi_{\eta,a}(x) = \eta^{-3/2} \chi((x-a)/\eta)$. For $0 < \eta < d_k/3$ there is a mixed fermionic state $\Gamma_{k,\eta}$ such that

$$K(\Gamma_{k,\eta}) = \frac{M}{2\eta^2} \int_{\mathbb{R}^3} |\nabla \chi|^2,$$

$$\pi_{\Gamma_{k,\eta}} = \sum_{\ell=1}^{L_k} \alpha_{k,\ell} \, \mathrm{Sym} \bigotimes_{i=1}^{M} \chi^2_{\eta, x_{k,\ell,i}}.$$

Consequently, as $\eta \downarrow 0$,

$$\pi_{\Gamma_{k,\eta}} \rightharpoonup \pi^k, \qquad \int_Z c \,\mathrm{d}\pi_{\Gamma_{k,\eta}} \to \int_Z c \,\mathrm{d}\pi^k, \qquad \Psi_n(\pi_{\Gamma_{k,\eta}}) \to \Psi_n(\pi^k).$$

If $\int_Z \Theta \,\mathrm{d}\pi^k < A$, *then the confinement inequality also holds for* η *small.*

Proof. For each ℓ, the supports of the orbitals $\chi_{\eta,x_{k,\ell,i}}$ are disjoint, hence they are orthonormal. Let $\mathcal{S}_{k,\ell,\eta}$ be the associated Slater determinant and set

$$\Gamma_{k,\eta} := \sum_{\ell=1}^{L_k} \alpha_{k,\ell} |\mathcal{S}_{k,\ell,\eta}\rangle\langle\mathcal{S}_{k,\ell,\eta}|.$$

Disjointness removes all exchange terms and gives the displayed formula for the law. The kinetic identity follows by scaling (see [24] for a similar construction). The remaining convergences follow from continuity of c away from D, bounded continuity of the moments, and continuity of Θ on the compact support. □

Lemma 5.5 (Exact finite-dimensional correction)**.** *Let* (Γ_η) *be states with* $K(\Gamma_\eta) \le A_0\eta^{-2}$, $\int_Z c \,\mathrm{d}\pi_{\Gamma_\eta} \le B_0$, $\int_Z \Theta \,\mathrm{d}\pi_{\Gamma_\eta} < A$, *and*

$$\beta_\eta := |\Psi_n(\pi_{\Gamma_\eta}) - b| \longrightarrow 0.$$

Under Assumption 5.1, there are exactly feasible states $\widetilde{\Gamma}_\eta$ *satisfying*

$$\Psi_n(\pi_{\widetilde{\Gamma}_\eta}) = b, \qquad K(\widetilde{\Gamma}_\eta) \le A_0\eta^{-2} + C\beta_\eta, \qquad \left| \int_Z c \,\mathrm{d}\pi_{\widetilde{\Gamma}_\eta} - \int_Z c \,\mathrm{d}\pi_{\Gamma_\eta} \right| \le C\beta_\eta,$$

and $d_{\mathrm{BL}}(\pi_{\widetilde{\Gamma}_\eta}, \pi_{\Gamma_\eta}) \to 0$.

Proof. Let $v_j := \Psi_n(\pi_{\Gamma^j})$. For $s \in (0,1]$ and a probability vector $\lambda = (\lambda_0, \dots, \lambda_J)$, set

$$\widetilde{\Gamma}_\eta(s,\lambda) := (1-s)\Gamma_\eta + s\sum_{j=0}^{J} \lambda_j \Gamma^j.$$

If $\delta_\eta := b - \Psi_n(\pi_{\Gamma_\eta})$, exact feasibility is equivalent to

$$\sum_{j=0}^{J} \lambda_j v_j = b + \frac{1-s}{s}\delta_\eta.$$

If $\beta_\eta = 0$, leave the state unchanged. Otherwise, choose $s_\eta := \min\{1, 2\beta_\eta/r\}$. For small η the right-hand side belongs to $B(b,r) \cap H_n$, so the assumed convex-hull condition gives a suitable λ_η. Convexity preserves the confinement inequality. Since kinetic and Coulomb energies are affine in the state, and $s_\eta = O(\beta_\eta)$, the two stated estimates follow. Finally,

$$d_{\mathrm{BL}}(\pi_{\widetilde{\Gamma}_\eta}, \pi_{\Gamma_\eta}) \le 2s_\eta \to 0.$$

□

Proposition 5.6 (Density of exactly feasible finite-energy laws)**.** *Under Assumption 5.1, for every* $\pi \in \Pi^n_{A,\mathrm{sym}}$ *with* $\int_Z c \,\mathrm{d}\pi < +\infty$ *there are exactly feasible states* $\widetilde{\Gamma}_j$ *such that, with* $\widetilde{\pi}_j := \pi_{\widetilde{\Gamma}_j}$,

$$d_{\mathrm{BL}}(\widetilde{\pi}_j, \pi) \to 0, \qquad \int_Z c \,\mathrm{d}\widetilde{\pi}_j \to \int_Z c \,\mathrm{d}\pi, \qquad K(\widetilde{\Gamma}_j) < +\infty.$$

Proof. Let $\sigma := \pi_{\Gamma^{\mathrm{ref}}}$ and, for $t \in (0,1]$, set

$$\pi_t := (1-t)\pi + t\sigma.$$

Then $\Psi_n(\pi_t) = b$, $\pi_t \rightharpoonup \pi$ and $\int c \,\mathrm{d}\pi_t \to \int c \,\mathrm{d}\pi$, while $A - \int \Theta \,\mathrm{d}\pi_t > 0$. Put $t_j := 1/j$ and $\sigma_j := A - \int_Z \Theta \,\mathrm{d}\pi_{t_j}$. For each fixed j, apply Lemma 5.3 to π_{t_j}. Then choose one of

its atomic laws, say $\pi_j^{k(j)}$, sufficiently close in distance, energy, moments and confinement, and write $d_{j,k(j)}$ for its separation. Choose a packet radius

$$0 < \eta_j < \min\{d_{j,k(j)}/3, 1/j\},$$

so small that the packet state Γ_{j,η_j} satisfies

$$\begin{aligned} d_{\mathrm{BL}}(\pi_{\Gamma_{j,\eta_j}}, \pi_{t_j}) &\leq 1/j, \\ \left| \int_Z c \, \mathrm{d}\pi_{\Gamma_{j,\eta_j}} - \int_Z c \, \mathrm{d}\pi_{t_j} \right| &\leq 1/j, \\ |\Psi_n(\pi_{\Gamma_{j,\eta_j}}) - b| &\leq 1/j, \\ \int_Z \Theta \, \mathrm{d}\pi_{\Gamma_{j,\eta_j}} &\leq A - \sigma_j/2. \end{aligned}$$

This is possible because the sparse approximations converge in energy and the packet construction converges for each fixed atomic law. Apply Lemma 5.5 to Γ_{j,η_j} and denote the corrected state by $\widetilde{\Gamma}_j$. Convexity preserves the confinement inequality, and the correction changes the bounded-Lipschitz distance and Coulomb energy by $o(1)$. Therefore

$$d_{\mathrm{BL}}(\pi_{\widetilde{\Gamma}_j}, \pi_{t_j}) \to 0, \qquad \int_Z c \, \mathrm{d}\pi_{\widetilde{\Gamma}_j} - \int_Z c \, \mathrm{d}\pi_{t_j} \to 0.$$

Finally, $d_{\mathrm{BL}}(\pi_{t_j}, \pi) \leq 2/j$ and

$$\left| \int_Z c \, \mathrm{d}\pi_{t_j} - \int_Z c \, \mathrm{d}\pi \right| \leq \frac{1}{j} \left| \int_Z c \, \mathrm{d}\sigma - \int_Z c \, \mathrm{d}\pi \right|.$$

The triangle inequality gives the two asserted convergences, while $K(\widetilde{\Gamma}_j) < +\infty$ follows from the packet and correction estimates. □

Theorem 5.7 (Semiclassical Γ-convergence)**.** *Under the standing hypotheses of this section, including continuity of Θ and Assumption 5.1, fix n. Then, as $\varepsilon \downarrow 0$, the functionals F_n^ε Γ-converge to $F_n^0 = \mathrm{C}_{\mathrm{sym}}^n$ on $\mathcal{P}(Z)$ endowed with the narrow topology, where $\mathrm{C}_{\mathrm{sym}}^n = +\infty$ outside the symmetric admissible set $\Pi_{A,\mathrm{sym}}^n$. Moreover, the family $(F_n^\varepsilon)_{0<\varepsilon\leq 1}$ is equicoercive.*

Consequently,

$$\mathcal{L}_n^\varepsilon \longrightarrow \min_{\pi \in \mathcal{P}(Z)} \mathrm{C}_{\mathrm{sym}}^n(\pi),$$

which equals $I^n(\rho)$ in the marginal-moment setting. If Γ_ε are minimizers (or $o(1)$-minimizers) of the semiclassical problem, then the family $(\pi_{\Gamma_\varepsilon})_\varepsilon$ is narrowly precompact and every cluster point is a minimizer of $\mathrm{C}_{\mathrm{sym}}^n$.

Proof. We first prove the Γ-liminf inequality. Let $\pi_\varepsilon \rightharpoonup \pi$ and suppose that

$$\liminf_{\varepsilon \to 0} F_n^\varepsilon(\pi_\varepsilon) < +\infty.$$

Up to a subsequence, $\pi_\varepsilon \in \Pi_{A,\mathrm{sym}}^n$. Symmetry passes to the narrow limit. Since all functions in W^n are bounded and continuous, the moment equalities pass to the limit. The mass constraint is preserved by narrow convergence, and lower semicontinuity of Θ gives

$$\int_Z \Theta \, \mathrm{d}\pi \leq \liminf_{\varepsilon \to 0} \int_Z \Theta \, \mathrm{d}\pi_\varepsilon \leq A.$$

Consequently $\pi \in \Pi_{A,\mathrm{sym}}^n$. Since $T \geq 0$ and c is lower semicontinuous and bounded from below, the Portmanteau theorem yields

$$\liminf_{\varepsilon \to 0} F_n^\varepsilon(\pi_\varepsilon) \geq \liminf_{\varepsilon \to 0} \int_Z c \, \mathrm{d}\pi_\varepsilon \geq \int_Z c \, \mathrm{d}\pi = F_n^0(\pi).$$

For the Γ-limsup, let

$$\mathcal{D}_n := \left\{ \pi_\Gamma : \Gamma \in \mathfrak{S}_a, K(\Gamma) < +\infty,\ \pi_\Gamma \in \Pi^n_{A,\mathrm{sym}},\ \int_Z c \,\mathrm{d}\pi_\Gamma < +\infty \right\}.$$

If $Q = \pi_\Gamma \in \mathcal{D}_n$, the constant sequence gives

$$\lim_{\varepsilon \downarrow 0} F^\varepsilon_n(Q) = \int_Z c \,\mathrm{d}Q = F^0_n(Q).$$

By Proposition 5.6, $\mathcal{D}_n$ is dense in the domain of F^0_n with convergence of the Coulomb energy. Let F' denote the Γ-upper limit. Since F' is lower semicontinuous, for every π with $F^0_n(\pi) < +\infty$ one obtains

$$F'(\pi) \le \liminf_{j\to\infty} F'(Q_j) \le \liminf_{j\to\infty} F^0_n(Q_j) = F^0_n(\pi), \qquad Q_j \in \mathcal{D}_n, Q_j \rightharpoonup \pi.$$

This is the Γ-limsup inequality.

It remains to check equicoercivity. Every sublevel set of F^ε_n, uniformly for $0 < \varepsilon \le 1$, is contained in $\Pi^n_{A,\mathrm{sym}}$. Since Θ is coercive, $\int_Z \Theta \,\mathrm{d}\pi \le A$ implies uniform tightness. Prokhorov's theorem therefore gives equicoercivity in the narrow topology.

The convergence of minima and the assertion on minimizers follow from the fundamental theorem of Γ-convergence (see, for example, [14, Chapters 7–9]). In the present setting one can also see the value convergence directly: the liminf inequality gives the lower bound, while a recovery sequence at any minimizer of $\mathrm{C}^n_{\mathrm{sym}}$ gives the upper bound. □

Remark 5.8 (Recovery from the diagonal estimate). For the classical problem, in which the density is prescribed, Lewin [24] proved the convergence of the minimal values in the semiclassical limit of the Levy–Lieb functional by constructing a recovery sequence at a minimizer only: an optimal transport plan is supported away from the diagonal, it is regularized at fixed marginals as in [5], and a fermionic trial state is built from Slater determinants of well-separated orbitals. One may wonder whether the same strategy applies to the moment-constrained problem. In this setting, Theorem 4.3 plays the role of the diagonal exclusion for optimal plans.

We explain how Theorem 4.3 supplies the recovery sequence needed for the value convergence. In the marginal-moment setting of Section 4, with $X = \mathbb{R}^3$ and the Coulomb cost, if $\pi^\star_n$ is a minimizer of the n-moment problem and $\kappa_\rho(\alpha) < 1/(2(M-1))$, Theorem 4.3 gives

$$\pi^\star_n(D_{\alpha_n}) = 0, \qquad h(\alpha_n) = \frac{I(\rho) - I^n(\rho)}{1 - 2(M-1)\kappa_\rho(\alpha)} + 2(M-1)h(\alpha).$$

Thus the minimizer has a positive distance from the repulsive diagonal, which, by Corollary 4.5, is bounded from below independently of the number of moments. After approximating $\pi^\star_n$ by the finitely supported measures of Proposition 4.1, whose atoms lie in $\mathrm{supp}(\pi^\star_n)$, one may smooth the atoms by packets of radius $\eta < \alpha_n/3$. The packets then remain away from the diagonal, and the Coulomb cost is continuous on their common support. The usual moment correction restores the constraints exactly.

More precisely, if this construction gives states $\Gamma_{n,\eta}$ with

$$K(\Gamma_{n,\eta}) \le A_n \eta^{-2}, \qquad \pi_{\Gamma_{n,\eta}} \rightharpoonup \pi^\star_n, \qquad \int_Z c \,\mathrm{d}\pi_{\Gamma_{n,\eta}} \longrightarrow I^n(\rho),$$

then for fixed n one may take $\eta = \varepsilon^{1/4}$ for all sufficiently small ε. This gives $\varepsilon K(\Gamma_{n,\eta}) = O(\sqrt{\varepsilon})$ and hence

$$\limsup_{\varepsilon \downarrow 0} \mathcal{L}^\varepsilon_n \le I^n(\rho).$$

Together with the lower bound, this proves convergence of the minimum values and of minimizers. Notice that this argument uses recovery at minimizers only. It is therefore

enough for the variational conclusion, but it does not replace the density construction above when the full Γ-limsup is asserted.

Finally, since $\alpha_n \geq \alpha^\star$ for every n by Corollary 4.5, the separation part of the construction is uniform in n. A joint limit $n = n(\varepsilon) \to \infty$ would additionally require uniform bounds for the moment-correction constants; the diagonal estimate alone does not provide such bounds.

6. Other applications

We now show how the general framework of Section 2 applies to other transport problems: moment inequalities describe partial transport and capacity constraints, additional equalities encode the martingale condition, and the underlying space may itself be a space of probability measures. We state the theorems without proof because their proofs are analogous to those in Section 2.6. We highlight any differences in the arguments when necessary.

6.1. **Partial optimal transport.** Let X, Y and c be as above, let $\mu \in \mathcal{M}_+(X)$ and $\nu \in \mathcal{M}_+(Y)$, and set $m = \min\{\mu(X), \nu(Y)\}$. The partial optimal transport problem with transported mass m is

$$\inf_{\pi \in \Pi_\leq(\mu,\nu)} \int_{X\times Y} c(x,y)\ \mathrm{d}\pi(x,y), \tag{48}$$

where

$$\Pi_\leq(\mu,\nu) \coloneqq \{\, \pi \in \mathcal{M}_+(X\times Y) :\ \pi(X\times Y) = m, \\ (p_X)_\#\pi \leq \mu, \quad (p_Y)_\#\pi \leq \nu \,\}.$$

If $m = 0$, the only admissible measure is the zero measure, and the optimal value is zero. Henceforth assume $m > 0$.

To obtain a finite moment relaxation, take $Z = X \times Y$, choose nonnegative functions $\varphi_1, \dots, \varphi_r \in C_b(X)$ and $\psi_1, \dots, \psi_s \in C_b(Y)$, and set $n = r + s$ and

$$w_i = \frac{1}{m}\int_X \varphi_i\ \mathrm{d}\mu - \varphi_i \circ p_X, \quad i = 1, \dots, r,$$
$$w_{r+j} = \frac{1}{m}\int_Y \psi_j\ \mathrm{d}\nu - \psi_j \circ p_Y, \quad j = 1, \dots, s.$$

Retain a coercive moment bound to ensure narrow compactness of the finite moment class. Choose θ_X and θ_Y with the properties described above and set

$$\Theta(x,y) = \theta_X(x) + \theta_Y(y), \qquad A > 0, \quad A \geq \int_X \theta_X\ \mathrm{d}\mu + \int_Y \theta_Y\ \mathrm{d}\nu.$$

Marginal domination then gives

$$\int_{X\times Y} \Theta\ \mathrm{d}\pi \leq \int_X \theta_X\ \mathrm{d}\mu + \int_Y \theta_Y\ \mathrm{d}\nu \leq A \qquad \text{for every } \pi \in \Pi_\leq(\mu,\nu).$$

Thus $\Pi_\leq(\mu,\nu) \subseteq \Pi^n_A$, and (2) provides a relaxation of (48).

With these choices, set $\Pi^n_{A,\leq}(\mu,\nu) \coloneqq \Pi^n_A$ and consider the finite moment approximation

$$\inf_{\pi \in \Pi^n_{A,\leq}(\mu,\nu)} \int_{X\times Y} c(x,y)\ \mathrm{d}\pi(x,y). \tag{49}$$

Theorem 6.1. *If* (49) *is finite, then it admits a minimizer which is an extreme point of* $\Pi^n_{A,\leq}(\mu,\nu)$ *and it is supported on at most* $r + s + 2$ *points.*

For the duality result below we introduce the following notation $E_r = \mathrm{cone}_+(\{\varphi_1, \dots, \varphi_r\})$, and $E_s = \mathrm{cone}_+(\{\psi_1, \dots, \psi_s\})$. In particular, every function in $-E_r$ or $-E_s$ is nonpositive.

Theorem 6.2. *Assume that Θ is finite-valued and continuous. If there exists $\eta \in \Pi^n_{A,\leq}(\mu,\nu)$ such that $\int_{X\times Y} c \, \mathrm{d}\eta < +\infty$ and $\operatorname{supp}(\eta) = \overline{\{c < +\infty\}}$, then*

$$\min_{\pi \in \Pi^n_{A,\leq}(\mu,\nu)} \int_{X\times Y} c \, \mathrm{d}\pi = \max_{\substack{\varphi\in -E_r,\ \psi \in -E_s \\ \lambda\in\mathbb{R},\ \tau\geq 0 \\ \varphi(x)+\psi(y)\leq c(x,y)-\lambda+\tau\Theta(x,y)}} \left\{ \int_X \varphi \, \mathrm{d}\mu + \int_Y \psi \, \mathrm{d}\nu + m\lambda - \tau A \right\}. \tag{50}$$

Definition 6.3. Let $\mu \in \mathcal{M}_+(X)$ and $m > 0$. A family of nonnegative functions $(\varphi_i)_{i\in\mathbb{N}} \subset C_b(X)$ characterizes domination by μ among measures of mass m if, for every $\eta \in \mathcal{M}_+(X)$ with $\eta(X) = m$,

$$\left(\int_X \varphi_i \, \mathrm{d}\eta \leq \int_X \varphi_i \, \mathrm{d}\mu \text{ for every } i \in \mathbb{N} \right) \Longrightarrow \eta \leq \mu.$$

Theorem 6.4. *Let $(\varphi_i)_{i\in\mathbb{N}} \subset C_b(X)$ and $(\psi_j)_{j\in\mathbb{N}} \subset C_b(Y)$ characterize domination by μ and ν, respectively, among measures of mass m. With Θ and A chosen as above, let I^n be the minimum in (49) using the first n tests of each family, and let I be the minimum in (48). If $I < +\infty$, then $I^n \nearrow I$ as $n \to +\infty$. Moreover, every sequence of minimizers of (49) admits a narrowly convergent subsequence, and every accumulation point is a minimizer of (48).*

Corollary 6.5. *Under the assumptions of Theorem 6.4, assume that (48) is finite and that Theorem 6.2 applies for every $n \in \mathbb{N}$. Then*

$$\min_{\pi\in\Pi_\leq(\mu,\nu)} \int_{X\times Y} c \, \mathrm{d}\pi = \sup_{\substack{\varphi\in C_b(X),\ \psi\in C_b(Y) \\ \varphi\leq 0,\ \psi\leq 0,\ \lambda\in\mathbb{R} \\ \varphi(x)+\psi(y)\leq c(x,y)-\lambda}} \left\{ \int_X \varphi \, \mathrm{d}\mu + \int_Y \psi \, \mathrm{d}\nu + m\lambda \right\}.$$

The dual problem in Corollary 6.5 has the same structure as that in [18, Theorem 2.2].

6.2. **The JKO scheme.** Let $F : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R} \cup \{+\infty\}$ be a functional, $\tau > 0$, and $\nu \in \mathcal{P}_2(\mathbb{R}^d)$. The associated JKO step is

$$\inf_{\rho\in\mathcal{P}_2(\mathbb{R}^d)} \left\{ F(\rho) + \frac{1}{2\tau} W_2^2(\rho,\nu) \right\}. \tag{51}$$

Assume that F is linear and of the form

$$F(\rho) = \int_{\mathbb{R}^d} V(x) \, \mathrm{d}\rho(x),$$

where $V : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous, bounded below, and not identically $+\infty$. Define $c_\tau(x,y) = V(x) + \frac{1}{2\tau}|x-y|^2$. Then (51) is equivalent to

$$\inf_{\substack{\pi\in\mathcal{P}_2(\mathbb{R}^d\times\mathbb{R}^d) \\ (p_Y)_\#\pi=\nu}} \int_{\mathbb{R}^d\times\mathbb{R}^d} c_\tau(x,y) \, \mathrm{d}\pi(x,y),$$

where $p_Y(x,y) = y$. Thus, the first marginal is free, while the second marginal is prescribed. To obtain a finite moment relaxation, take $Z = \mathbb{R}^d \times \mathbb{R}^d$ and $m = 1$, choose $\psi_1, \ldots, \psi_s \in C_b(\mathbb{R}^d)$, and define

$$g_j = \psi_j \circ p_Y - \int_{\mathbb{R}^d} \psi_j \, \mathrm{d}\nu, \qquad j = 1, \ldots, s.$$

Using $g_j, -g_j$ as the constraint functions gives $n = 2s$ inequalities of the form appearing in (2).

A compatible coercive moment bound can be obtained by taking

$$\Theta(x,y) = |x-y|^2 + |y|^2.$$

Indeed, let $b \in \mathbb{R}$ satisfy $V \geq b$, choose $x_0 \in \mathbb{R}^d$ with $V(x_0) < +\infty$, and set

$$B = V(x_0) + \frac{1}{2\tau}\int_{\mathbb{R}^d} |x_0 - y|^2 \, \mathrm{d}\nu(y).$$

The plan $\delta_{x_0} \otimes \nu$ has cost B. Every plan with second marginal ν and cost at most B satisfies

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \Theta \, \mathrm{d}\pi \leq 2\tau(B-b) + \int_{\mathbb{R}^d} |y|^2 \, \mathrm{d}\nu(y).$$

Consequently, choosing $A > 0$ at least as large as the right-hand side preserves the optimal value of (51). With this choice, (2) is a finite moment relaxation of the JKO step, with a narrowly compact admissible set.

Define

$$\begin{aligned}\Pi_A^n(\nu) \coloneqq \big\{\pi \in \mathcal{P}(\mathbb{R}^d\times\mathbb{R}^d) : \int_{\mathbb{R}^d\times\mathbb{R}^d} \Theta \, \mathrm{d}\pi \leq A,\\ \int_{\mathbb{R}^d\times\mathbb{R}^d} \psi_j(y) \, \mathrm{d}\pi(x,y) = \int_{\mathbb{R}^d} \psi_j \, \mathrm{d}\nu, \quad j = 1,\dots,s\big\}.\end{aligned} \tag{52}$$

We consider the finite moment problem

$$\inf_{\pi\in\Pi_A^n(\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} c_\tau(x,y) \, \mathrm{d}\pi(x,y). \tag{53}$$

Theorem 6.6. *If* (53) *is finite, then it admits a minimizer which is an extreme point of* $\Pi_A^n(\nu)$ *and it is supported on at most* $s+2$ *points.*

For the duality result below we use β for the multiplier associated with the coercive moment bound, since τ is classically used to denote the time step. Define $E_s = \mathrm{span}(\{1,\psi_1,\dots,\psi_s\})$.

Theorem 6.7. *Assume that* Θ *is finite-valued and continuous. If there exists* $\eta \in \Pi_A^n(\nu)$ *such that* $\int_{\mathbb{R}^d\times\mathbb{R}^d} c_\tau \, \mathrm{d}\eta < +\infty$ *and* $\mathrm{supp}(\eta) = \overline{\{c_\tau < +\infty\}}$, *then*

$$\min_{\pi\in\Pi_A^n(\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} c_\tau \, \mathrm{d}\pi = \max_{\substack{\psi\in E_s,\ \beta\geq 0\\ \psi(y)\leq c_\tau(x,y)+\beta\Theta(x,y)}} \left\{\int_{\mathbb{R}^d} \psi \, \mathrm{d}\nu - \beta A\right\}. \tag{54}$$

Theorem 6.8. *Let* $(\psi_j)_{j\in\mathbb{N}} \subset C_b(\mathbb{R}^d)$ *fully characterize* ν *in the sense of Definition 2.12. Let* I^n *be the minimum in* (53) *using the first* n *tests, and let* I *be the minimum in* (51). *Assume that* $I < +\infty$ *and that some optimal plan for the lifted formulation of* (51) *satisfies* $\int \Theta \, \mathrm{d}\pi \leq A$. *Then* $I^n \nearrow I$ *as* $n \to +\infty$. *Moreover, every sequence of minimizers of* (53) *admits a narrowly convergent subsequence, and every accumulation point* π *satisfies* $(p_Y)_\#\pi = \nu$ *and has first marginal* $(p_X)_\#\pi$ *minimizing* (51).

Corollary 6.9. *Under the assumptions of Theorem 6.8, assume that Theorem 6.7 applies for every* $n \in \mathbb{N}$. *Then*

$$\min_{\rho\in\mathcal{P}_2(\mathbb{R}^d)} \left\{\int_{\mathbb{R}^d} V \, \mathrm{d}\rho + \frac{1}{2\tau}W_2^2(\rho,\nu)\right\} = \sup_{\substack{\psi\in C_b(\mathbb{R}^d),\beta\geq 0\\ \psi(y)\leq c_\tau(x,y)+\beta\Theta(x,y)}} \int_{\mathbb{R}^d} \psi \, \mathrm{d}\nu - \beta A.$$

Our dual formula differs in appearance from [19, equation (1.4)], where the maximization involves two potentials and the convex conjugate of the energy. Here, only one potential appears, and the energy enters directly through the pointwise constraint

$$\psi(y) \leq V(x) + \frac{1}{2\tau}|x-y|^2.$$

This simplification results from the fact that the contribution of a linear energy can be incorporated into the transport cost, allowing the auxiliary potential in their formulation to be eliminated.

6.3. **Martingale optimal transport.** Let $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^d)$, where $\mathcal{P}_1(\mathbb{R}^d)$ denotes the probability measures with finite first moment, and let $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous and bounded below. For $\pi \in \Pi(\mu, \nu)$, write $\pi(\mathrm{d}x, \mathrm{d}y) = \mu(\mathrm{d}x)\pi_x(\mathrm{d}y)$ for its disintegration with respect to the first marginal. The martingale optimal transport problem is

$$\inf_{\substack{\pi \in \Pi(\mu,\nu) \\ \int_{\mathbb{R}^d} y \, \mathrm{d}\pi_x(y) = x \ \mu\text{-a.e.}}} \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x, y) \, \mathrm{d}\pi(x, y). \tag{55}$$

Its admissible set is nonempty if and only if $\mu \preceq_{\mathrm{cx}} \nu$, where $\preceq_{\mathrm{cx}}$ denotes convex order (see [2] for details).

Take $Z = \mathbb{R}^d \times \mathbb{R}^d$ and $m = 1$. The martingale condition is equivalent to

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} \phi(x) \cdot (y - x) \, \mathrm{d}\pi(x, y) = 0 \qquad \text{for every } \phi \in C_b(\mathbb{R}^d; \mathbb{R}^d).$$

Thus, in addition to the marginal moment constraints, one may select finitely many functions $\phi_1, \ldots, \phi_q$ and impose the inequalities associated with h_ℓ and $-h_\ell$, where

$$h_\ell(x, y) = \phi_\ell(x) \cdot (y - x), \qquad \ell = 1, \ldots, q.$$

These functions are continuous but generally unbounded. Consequently, this falls outside the assumption $w_i \in C_b(Z)$ and requires an extension to continuous unbounded tests, together with suitable uniform integrability conditions.

For exact marginal constraints, such an extension is considered in [32]. For Polish spaces X and Y, the author introduces the space $C_L(\mu, \nu)$ of functions $h \in C(X \times Y)$ satisfying $|h(x, y)| \le f(x) + g(y)$ for some $f \in C(X) \cap L^1(\mu)$ and $g \in C(Y) \cap L^1(\nu)$, equipped with the seminorm

$$\|h\|_L \coloneqq \sup_{\pi \in \Pi(\mu,\nu)} \int_{X \times Y} |h| \, \mathrm{d}\pi.$$

The space $C_b(X \times Y)$ is dense in $C_L(\mu, \nu)$ for this seminorm, and $\pi \mapsto \int h \, \mathrm{d}\pi$ is narrowly continuous on $\Pi(\mu, \nu)$ for every $h \in C_L(\mu, \nu)$. In particular, the martingale tests belong to this space, since

$$|\phi(x) \cdot (y - x)| \le \|\phi\|_\infty (|x| + |y|).$$

This continuity statement relies on the exact marginal constraints and does not automatically extend to their finite moment relaxation.

For the latter, one may impose growth conditions of the type used in [2]. Specifically, for the selected marginal tests φ_i, ψ_j and martingale tests h_ℓ, a sufficient condition is

$$|\varphi_i(x)| + |\psi_j(y)| + |h_\ell(x, y)| \le C\big(1 + \Theta(x, y)\big)^s, \qquad 0 < s < 1,$$

for some $C > 0$, where

$$\Theta(x, y) = \theta_\mu(|x|) + \theta_\nu(|y|).$$

Here $\theta_\mu, \theta_\nu : \mathbb{R}_+ \to \mathbb{R}_+$ are continuous, nondecreasing functions tending to infinity, with

$$\int_{\mathbb{R}^d} \theta_\mu(|x|) \, \mathrm{d}\mu(x) + \int_{\mathbb{R}^d} \theta_\nu(|y|) \, \mathrm{d}\nu(y) < \infty.$$

These are additional assumptions on the selected tests and the marginals. Choose $A > 0$ at least as large as this sum. The bound $\int \Theta \, \mathrm{d}\pi \le A$, together with the sublinear growth condition, ensures uniform integrability of the constraint functions. Their moment constraints therefore pass to narrow limits, allowing the finite moment formulation to accommodate these unbounded tests.

For the next results, assume that $X, Y \subset \mathbb{R}^d$ are compact, so that we may take $\Theta \equiv 0$. In addition to the finite marginal constraints, choose $H_1, \dots, H_q \in C(X; \mathbb{R}^d)$ and impose

$$\int_{X\times Y} H_\ell(x) \cdot (y - x) \, \mathrm{d}\pi(x, y) = 0, \qquad \ell = 1, \dots, q.$$

The functions $h_\ell(x, y) = H_\ell(x) \cdot (y - x)$ belong to $C_b(X \times Y)$, so these equalities are encoded by adding $h_\ell, -h_\ell$ to the moment functions. We denote the resulting admissible set by $\Pi^n_{\mathrm{mart}}(\mu, \nu)$, with $n = 2(r + s + q)$, and consider

$$\inf_{\pi\in\Pi^n_{\mathrm{mart}}(\mu,\nu)} \int_{X\times Y} c(x, y) \, \mathrm{d}\pi(x, y). \tag{56}$$

Theorem 6.10. *If* (56) *is finite, then it admits a minimizer which is an extreme point of* $\Pi^n_{\mathrm{mart}}(\mu, \nu)$ *and is supported on at most* $r + s + q + 1$ *points.*

For the duality result below we introduce the following notation

$$E_r = \mathrm{span}\{1, \varphi_1, \dots, \varphi_r\}, \qquad E_s = \mathrm{span}\{1, \psi_1, \dots, \psi_s\}, \qquad \mathcal{H}_q = \mathrm{span}\{H_1, \dots, H_q\}.$$

Theorem 6.11. *If there exists* $\eta \in \Pi^n_{\mathrm{mart}}(\mu, \nu)$ *such that* $\int_{X\times Y} c \, \mathrm{d}\eta < +\infty$ *and* $\mathrm{supp}(\eta) = \overline{\{c < +\infty\}}$, *then*

$$\min_{\pi\in\Pi^n_{\mathrm{mart}}(\mu,\nu)} \int_{X\times Y} c \, \mathrm{d}\pi \quad = \max_{\substack{\varphi\in E_r,\ \psi\in E_s,\ H\in\mathcal{H}_q \\ \varphi(x)+\psi(y)+H(x)\cdot(y-x)\le c(x,y)}} \left\{ \int_X \varphi \, \mathrm{d}\mu + \int_Y \psi \, \mathrm{d}\nu \right\}. \tag{57}$$

Theorem 6.12. *Let* $(\varphi_i)_{i\in\mathbb{N}} \subset C(X)$ *and* $(\psi_j)_{j\in\mathbb{N}} \subset C(Y)$ *fully characterize* μ *and* ν, *respectively, and let* $(H_\ell)_{\ell\in\mathbb{N}} \subset C(X; \mathbb{R}^d)$ *have a uniformly dense linear span in* $C(X; \mathbb{R}^d)$. *Let* I^k *be the minimum in* (56) *using the first* k *tests of each family, so that* $r = s = q = k$ *and* $n = 6k$, *and let* I *be the minimum in* (55). *If* $I < +\infty$, *then* $I^k \nearrow I$ *as* $k \to +\infty$. *Moreover, every sequence of minimizers of* (56) *admits a narrowly convergent subsequence, and every accumulation point is a minimizer of* (55).

Corollary 6.13. *Under the assumptions of Theorem 6.12, assume that* (55) *is finite and that Theorem 6.11 applies at every approximation level. Then*

$$\min_{\pi\in\Pi_{\mathrm{mart}}(\mu,\nu)} \int_{X\times Y} c \, \mathrm{d}\pi = \sup_{\substack{\varphi\in C(X),\ \psi\in C(Y) \\ H\in C(X;\mathbb{R}^d) \\ \varphi(x)+\psi(y)+H(x)\cdot(y-x)\le c(x,y)}} \left\{ \int_X \varphi \, \mathrm{d}\mu + \int_Y \psi \, \mathrm{d}\nu \right\},$$

where $\Pi_{\mathrm{mart}}(\mu, \nu)$ *denotes the admissible set in* (55).

The dual formulation above has the same structure as in [4, 17].

6.4. **Capacity constrained optimal transport.** Let X, Y and c be as in the Monge–Kantorovich problem, let $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$, and let $\sigma \in \mathcal{M}_+(X \times Y)$ be a prescribed capacity measure. Consider

$$\inf_{\substack{\pi\in\Pi(\mu,\nu) \\ \pi\le\sigma}} \int_{X\times Y} c(x, y) \, \mathrm{d}\pi(x, y). \tag{58}$$

This includes a prescribed upper bound on the transport density by taking $\sigma = \bar{h}\,\lambda$ for a reference measure λ and a nonnegative $\bar{h} \in L^1(\lambda)$. The Euclidean density formulation is studied in [22].

Take $Z = X \times Y$ and $m = 1$. The capacity constraint is equivalent to

$$\int_{X\times Y} u \, \mathrm{d}\pi \le \int_{X\times Y} u \, \mathrm{d}\sigma \qquad \text{for every } u \in C_b(X \times Y), \quad u \ge 0.$$

Thus, choose finitely many nonnegative functions $u_1, \dots, u_q \in C_b(X \times Y)$ and define

$$a_k = \int_{X\times Y} u_k \, \mathrm{d}\sigma - u_k, \qquad k = 1, \dots, q.$$

Adding the inequalities $\int a_k \, \mathrm{d}\pi \geq 0$ to the finite marginal constraints gives a problem of the form (2). Retaining the compatible coercive moment bound constructed from μ and ν ensures narrow compactness and preserves every admissible plan of (58).

Define, for $n = 2(r+s) + q$,

$$\Pi^n_{A,\sigma}(\mu,\nu) \coloneqq \big\{\pi \in \Pi^{2(r+s)}_A(\mu,\nu) : \quad \int_{X\times Y} a_k \, \mathrm{d}\pi \geq 0, \quad k = 1, \dots, q\big\}, \tag{59}$$

where $\Pi^{2(r+s)}_A(\mu,\nu)$ is the marginal moment class introduced in the Monge–Kantorovich case. We consider the finite moment approximation of (58)

$$\inf_{\pi\in\Pi^n_{A,\sigma}(\mu,\nu)} \int_{X\times Y} c(x,y) \, \mathrm{d}\pi(x,y). \tag{60}$$

Theorem 6.14. *If* (60) *is finite, then it admits a minimizer which is an extreme point of* $\Pi^n_{A,\sigma}(\mu,\nu)$ *and is supported on at most* $r+s+q+2$ *points.*

For the duality result below we consider E_r and E_s as in the Monge–Kantorovich case and we introduce also $U_q = \mathrm{cone}_+(\{u_1, \dots, u_q\})$.

Theorem 6.15. *Assume that* Θ *is finite-valued and continuous. If there exists* $\eta \in \Pi^n_{A,\sigma}(\mu,\nu)$ *such that* $\int_{X\times Y} c \, \mathrm{d}\eta < +\infty$ *and* $\mathrm{supp}(\eta) = \overline{\{c < +\infty\}}$, *then*

$$\min_{\pi\in\Pi^n_{A,\sigma}(\mu,\nu)} \int_{X\times Y} c \, \mathrm{d}\pi = \max_{\substack{\varphi\in E_r,\ \psi\in E_s \\ u\in U_q,\ \tau\geq 0 \\ \varphi+\psi\leq c+u+\tau\Theta}} \left\{\int_X \varphi \, \mathrm{d}\mu + \int_Y \psi \, \mathrm{d}\nu - \int_{X\times Y} u \, \mathrm{d}\sigma - \tau A\right\}. \tag{61}$$

Remark 6.16. The capacity constraint is imposed through finitely many moment inequalities, so measures in $\Pi^n_{A,\sigma}(\mu,\nu)$ need not satisfy $\pi \leq \sigma$. In particular, they may be atomic even if σ has no atoms, whereas the constraint $\pi \leq \sigma$ would rule out atoms in this case.

Theorem 6.17. *Let* $(\varphi_i)_{i\in\mathbb{N}} \subset C_b(X)$ *and* $(\psi_j)_{j\in\mathbb{N}} \subset C_b(Y)$ *fully characterize* μ *and* ν, *respectively, in the sense of Definition 2.12. Let* $(u_k)_{k\in\mathbb{N}} \subset C_b(X\times Y)$ *be nonnegative and characterize domination by* σ *among measures of mass* 1, *in the sense of Definition 6.3. With* Θ *and* A *chosen as above, let* I^n *be the minimum in* (60) *using the first* n *tests of each family, and let* I *be the minimum in* (58). *If* $I < +\infty$, *then* $I^n \nearrow I$ *as* $n \to +\infty$. *Moreover, every sequence of minimizers of* (60) *admits a narrowly convergent subsequence, and every accumulation point is a minimizer of* (58).

Corollary 6.18. *Under the assumptions of Theorem 6.17, assume that Theorem 6.15 applies for every* $n \in \mathbb{N}$. *Then*

$$I = \sup_{\substack{\varphi\in C_b(X) \\ \psi\in C_b(Y)}} \left\{\int_X \varphi \, \mathrm{d}\mu + \int_Y \psi \, \mathrm{d}\nu - \int_{X\times Y} \big(\varphi(x) + \psi(y) - c(x,y)\big)_+ \, \mathrm{d}\sigma\right\}.$$

Here $a_+ = \max\{a, 0\}$, *with the convention* $(-\infty)_+ = 0$.

Our dual formula is similar to [23, eqs. (4)–(6)] with $\sigma = \bar{h} \, \mathrm{d}x \, \mathrm{d}y$ and surplus $s = -c$. Therefore, our expression is the negative of theirs.

6.5. **Wasserstein-on-Wasserstein transport.** Let (S,d) be a complete separable metric space, $1 \leq p < \infty$, and set $E = (\mathcal{P}_p(S), W_p)$, which is itself a Polish space. For two laws $\mathbb{M}, \mathbb{N} \in \mathcal{P}_p(E)$, the Wasserstein-on-Wasserstein distance is defined by

$$\mathbb{W}_p^p(\mathbb{M},\mathbb{N}) = \inf_{\Gamma \in \Pi(\mathbb{M},\mathbb{N})} \int_{E\times E} W_p^p(\rho,\sigma)\, \mathrm{d}\Gamma(\rho,\sigma). \tag{62}$$

See [27]. Here the transported objects are probability measures, and the objective remains linear in the coupling Γ.

This problem fits our framework with $Z = E \times E$, $m = 1$ and $c(\rho,\sigma) = W_p^p(\rho,\sigma)$. Choose $F_1,\dots,F_r,G_1,\dots,G_s \in C_b(E)$ and define

$$f_i = F_i \circ p_1 - \int_E F_i \,\mathrm{d}\mathbb{M}, \quad i = 1,\dots,r,$$
$$g_j = G_j \circ p_2 - \int_E G_j \,\mathrm{d}\mathbb{N}, \quad j = 1,\dots,s,$$

where $p_1, p_2 : E \times E \to E$ are the canonical projections. Using $f_i, -f_i, g_j, -g_j$ as the constraint functions gives the finite moment relaxation.

As before, choose nonnegative lower semicontinuous functions $\theta_{\mathbb{M}}, \theta_{\mathbb{N}} : E \to [0,+\infty]$ with compact sublevel sets in the W_p topology and finite integrals against the respective laws. Set

$$\Theta(\rho,\sigma) = \theta_{\mathbb{M}}(\rho) + \theta_{\mathbb{N}}(\sigma),$$
$$A \geq \int_E \theta_{\mathbb{M}} \,\mathrm{d}\mathbb{M} + \int_E \theta_{\mathbb{N}} \,\mathrm{d}\mathbb{N}, \qquad A > 0.$$

This bound preserves every exact coupling and ensures narrow compactness of the finite moment class.

The tests F_i and G_j are functionals on the space of measures. Restricting them to linear evaluations $\rho \mapsto \int_S \varphi \,\mathrm{d}\rho$ would constrain only the barycenters of the marginal laws, rather than determine the laws themselves.

We denote by $\Pi_A^n(\mathbb{M},\mathbb{N})$ the admissible set corresponding to the chosen marginal tests, with $n = 2(r+s)$, and consider

$$\inf_{\Gamma \in \Pi_A^n(\mathbb{M},\mathbb{N})} \int_{E\times E} W_p^p(\rho,\sigma)\, \mathrm{d}\Gamma(\rho,\sigma). \tag{63}$$

Theorem 6.19. *Problem* (63) *admits a minimizer which is an extreme point of* $\Pi_A^n(\mathbb{M},\mathbb{N})$ *and is supported on at most* $r+s+2$ *points of* $E \times E$.

For the duality result let $E_r = \mathrm{span}(\{F_0,\dots,F_r\})$ and $E_s = \mathrm{span}(\{G_0,\dots,G_s\})$ where $F_0 \equiv G_0 \equiv 1$.

Theorem 6.20. *Assume that* S *is compact and choose* $\Theta \equiv 0$. *If there exists* $\eta \in \Pi_A^n(\mathbb{M},\mathbb{N})$ *such that* $\mathrm{supp}(\eta) = E \times E$, *then*

$$\min_{\Gamma \in \Pi_A^n(\mathbb{M},\mathbb{N})} \int_{E\times E} W_p^p(\rho,\sigma)\, \mathrm{d}\Gamma(\rho,\sigma) = \max_{\substack{F\in E_r,\ G\in E_s \\ F(\rho)+G(\sigma)\leq W_p^p(\rho,\sigma)}} \left\{ \int_E F \,\mathrm{d}\mathbb{M} + \int_E G \,\mathrm{d}\mathbb{N} \right\}. \tag{64}$$

Remark 6.21. The discrete characterization concerns the coupling Γ on $E \times E$: the measures ρ and σ appearing in its support need not themselves be discrete.

Remark 6.22. Since S is compact, E is compact and the cost W_p^p is bounded, so finiteness of the cost of η is automatic. If $\mathbb{M}$ and $\mathbb{N}$ have full support on E, one may choose $\eta = \mathbb{M}\otimes\mathbb{N}$. When S is noncompact, a finite-valued continuous coercive function Θ on $E \times E$ need not exist ([25, p. 978]). Indeed, the existence of such a function would imply that $E \times E$, and hence E, is locally compact.

Theorem 6.23. *Let $(F_i)_{i\in\mathbb{N}}, (G_j)_{j\in\mathbb{N}} \subset C_b(E)$ fully characterize $\mathbb{M}$ and $\mathbb{N}$, respectively, in the sense of Definition 2.12. With Θ and A chosen as above, let I^n be the minimum in* (63) *using the first n tests of each family. Then $I^n \nearrow \mathbb{W}_p^p(\mathbb{M},\mathbb{N})$ as $n \to +\infty$. Moreover, every sequence of minimizers of* (63) *admits a narrowly convergent subsequence in $\mathcal{P}(E\times E)$, and every accumulation point is a minimizer of* (62).

Corollary 6.24. *Under the assumptions of Theorem 6.23, assume that S is compact and take $\Theta \equiv 0$. If Theorem 6.20 applies for every $n \in \mathbb{N}$, then*

$$\mathbb{W}_p^p(\mathbb{M},\mathbb{N}) = \sup_{\substack{F,G\in C_b(E)\\ F(\rho)+G(\sigma)\le W_p^p(\rho,\sigma) \text{ for every } \rho,\sigma\in E}} \left\{ \int_E F \,\mathrm{d}\mathbb{M} + \int_E G \,\mathrm{d}\mathbb{N} \right\}.$$

The dual formula above coincides with Kantorovich duality applied to the metric space $(\mathcal{P}_p(S), W_p)$; compare [27, equation (2.7)].

A.I. Disclosure

In this paper both Claude and ChatGPT were used for grammar and spelling corrections, as well as a first check for the correctness of the proofs. Theorem 3.4 was originally obtained with a worst constant for h_n. Discussions with ChatGPT helped to improve the constant, hence obtaining the current argument. The authors take full responsibility for correctness and the results of this paper.

Acknowledgements.

L.N. benefited from the ANR project GOTA (ANR-23-CE46-0001) and also supported by the French National Research Agency (ANR) under the France 2030 grant reference number ANR-24-RRII-0002 operated by the Inria Quadrant Program (project Pi-QuanTO).

References

[1] Aurélien Alfonsi, Rafaël Coyaud, and Virginie Ehrlacher. Constrained overdamped Langevin dynamics for symmetric multimarginal optimal transportation. *Mathematical Models and Methods in Applied Sciences*, 32(3):403–455, 2022.

[2] Aurélien Alfonsi, Rafaël Coyaud, Virginie Ehrlacher, and Damiano Lombardi. Approximation of optimal transport problems with marginal moments constraints. *Mathematics of Computation*, 90(328):689–737, 2021.

[3] Christian Bayer and Josef Teichmann. The proof of Tchakaloff's theorem. *Proceedings of the American Mathematical Society*, 134(10):3035–3040, 2006.

[4] 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.

[5] Ugo Bindini and Luigi De Pascale. Optimal transport with Coulomb cost and the semiclassical limit of density functional theory. *Journal de l'École polytechnique — Mathématiques*, 4:909–934, 2017.

[6] Jonathan M. Borwein and Adrian S. Lewis. Partially finite convex programming, part I: Quasi relative interiors and duality theory. *Mathematical Programming*, 57(1–3):15–48, 1992.

[7] Guy Bouchitté, Chloé Jimenez, and Mahadevan Rajesh. A new L^∞ estimate in optimal mass transport. *Proceedings of the American Mathematical Society*, 135(11):3525–3536, 2007.

[8] Giuseppe Buttazzo, Thierry Champion, and Luigi De Pascale. Continuity and estimates for multimarginal optimal transportation problems with singular costs. *Applied Mathematics & Optimization*, 78(1):185–200, 2018.

[9] Giuseppe Buttazzo, Luigi De Pascale, and Paola Gori-Giorgi. Optimal-transport formulation of electronic density-functional theory. *Physical Review A*, 85(6):062502, 2012.

[10] Guillaume Carlier, Hugo Malamut, and Maxime Sylvestre. Weak optimal transport with moment constraints: constraint qualification, dual attainment and entropic regularization. arXiv preprint arXiv:2511.16211, 2025.

[11] Maria Colombo, Simone Di Marino, and Federico Stra. Continuity of multimarginal optimal transport with repulsive cost. *SIAM Journal on Mathematical Analysis*, 51(4):2903–2926, 2019.

[12] 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.
[13] Codina Cotar, Gero Friesecke, and Claudia Klüppelberg. Smoothing of transport plans with fixed marginals and rigorous semiclassical limit of the Hohenberg–Kohn functional. *Archive for Rational Mechanics and Analysis*, 228(3):891–922, 2018.
[14] Gianni Dal Maso. *An Introduction to Γ-Convergence*, volume 8 of *Progress in Nonlinear Differential Equations and Their Applications*. Birkhäuser, Boston, 1993.
[15] Philipp J. di Dio and Konrad Schmüdgen. The multidimensional truncated moment problem: The moment cone. *Journal of Mathematical Analysis and Applications*, 511(1):126066, 2022.
[16] Virginie Ehrlacher and Luca Nenna. Sparse approximation of the Lieb functional with moment constraints. *SIAM Journal on Mathematical Analysis*, 58(3):2865–2893, 2026.
[17] Gaoyue Guo and Jan Obłój. Computational methods for martingale optimal transport problems. *The Annals of Applied Probability*, 29(6):3311–3347, 2019.
[18] Noureddine Igbida and Van Thanh Nguyen. Optimal partial transport problem with Lagrangian costs. *ESAIM: Mathematical Modelling and Numerical Analysis*, 52(5):2109–2132, 2018.
[19] Matt Jacobs, Wonjun Lee, and Flavien Léger. The back-and-forth method for Wasserstein gradient flows. *ESAIM: Control, Optimisation and Calculus of Variations*, 27:28, 2021.
[20] Heikki Jylhä and Tapio Rajala. L^∞ estimates in optimal mass transportation. *Journal of Functional Analysis*, 270(11):4297–4321, 2016.
[21] Alexander V. Kolesnikov, Aleksandra Novikova, and Kirill Sokolov. Duality for a martingale transport problem with moment constraints. arXiv preprint arXiv:2609.03187, 2026.
[22] Jonathan Korman and Robert J. McCann. Optimal transportation with capacity constraints. *Transactions of the American Mathematical Society*, 367(3):1501–1521, 2015.
[23] Jonathan Korman, Robert J. McCann, and Christian Seis. Dual potentials for capacity constrained optimal transport. *Calculus of Variations and Partial Differential Equations*, 54(1):573–584, 2015.
[24] Mathieu Lewin. Semi-classical limit of the Levy–Lieb functional in density functional theory. *Comptes Rendus. Mathématique*, 356(4):449–455, 2018.
[25] John Lott and Cédric Villani. Ricci curvature for metric-measure spaces via optimal transport. *Annals of Mathematics*, 169(3):903–991, 2009.
[26] Iosif Pinelis. On the extreme points of moments sets. *Mathematical Methods of Operations Research*, 83(3):325–349, 2016.
[27] Alessandro Pinzi and Giuseppe Savaré. Totally convex functions, L^2-optimal transport for laws of random measures, and solution to the Monge problem. arXiv preprint arXiv:2509.01768, 2025.
[28] Walter Rudin. *Functional Analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, NY, second edition, 1991.
[29] Konrad Schmüdgen. *The moment problem*, volume 277 of *Graduate Texts in Mathematics*. Springer, Cham, 2017.
[30] Cédric Villani. *Topics in Optimal Transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
[31] Gerhard Winkler. Extreme points of moment sets. *Mathematics of Operations Research*, 13(4):581–587, 1988.
[32] Danila A. Zaev. On the Monge–Kantorovich problem with additional linear constraints. *Mathematical Notes*, 98(5–6):725–741, 2015.

Dipartimento di Matematica e Informatica, Università di Firenze, Viale Morgagni 67/a, 50134 Firenze, Italy
Email address: luigi.depascale@unifi.it

Université Paris-Saclay, CNRS, Laboratoire de mathématiques d'Orsay, ParMA, Inria Saclay, 91405, Orsay, France. Institut Universitaire de France (IUF).
Email address: luca.nenna@universite-paris-saclay.fr

Dipartimento di Matematica e Informatica "Ulisse Dini", Università degli Studi di Firenze
Email address: igorvinicius.pereirapinheiro@unifi.it