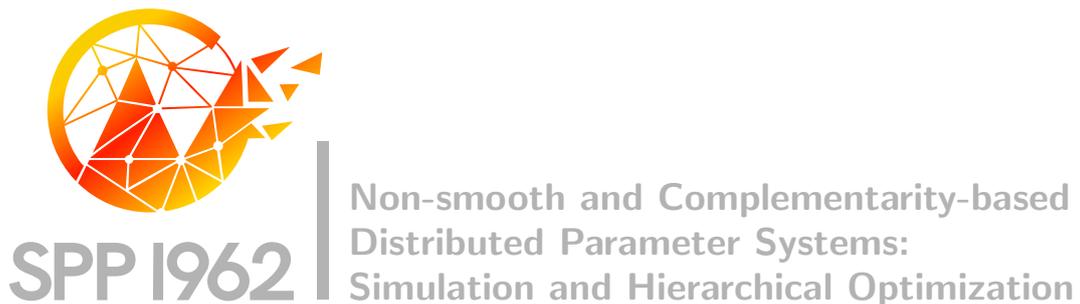


DFG Deutsche
Forschungsgemeinschaft
Priority Programme 1962

Quadratically Regularized Optimal Transport

Dirk A. Lorenz, Paul Manns, Christian Meyer



Preprint Number SPP1962-120

received on September 16, 2019

Edited by
SPP1962 at Weierstrass Institute for Applied Analysis and Stochastics (WIAS)
Leibniz Institute in the Forschungsverbund Berlin e.V.
Mohrenstraße 39, 10117 Berlin, Germany
E-Mail: spp1962@wias-berlin.de

World Wide Web: <http://spp1962.wias-berlin.de/>

Quadratically regularized optimal transport

Dirk A. Lorenz · Paul Manns · Christian Meyer

the date of receipt and acceptance should be inserted later

Abstract We investigate the problem of optimal transport in the so-called Kantorovich form, i.e. given two Radon measures on two compact sets, we seek an optimal transport plan which is another Radon measure on the product of the sets that has these two measures as marginals and minimizes a certain cost function.

We consider quadratic regularization of the problem, which forces the optimal transport plan to be a square integrable function rather than a Radon measure. We derive the dual problem and show strong duality and existence of primal and dual solutions to the regularized problem. Then we derive two algorithms to solve the dual problem of the regularized problem: A Gauss-Seidel method and a semismooth quasi-Newton method and investigate both methods numerically. Our experiments show that the methods perform well even for small regularization parameters. Quadratic regularization is of interest since the resulting optimal transport plans are sparse, i.e. they have a small support (which is not the case for the often used entropic regularization where the optimal transport plan always has full measure).

Keywords optimal transport, regularization, semismooth Newton method, Gauss-Seidel method, duality

Mathematics Subject Classification (2010) 49Q20, 65D99, 90C25

1 Introduction

In this paper we will investigate a regularized version of the optimal transport problem. Optimal transport dates back to the work of Monge in 1781 but the

Dirk A. Lorenz, E-mail: d.lorenz@tu-braunschweig.de
TU Braunschweig, Institute of Analysis and Algebra, Germany, ORCID iD: 0000-0002-7419-769X

Paul Manns, E-mail: paul.manns@tu-braunschweig.de
TU Braunschweig, Institute of Mathematical Optimization, Germany

Christian Meyer, E-mail: christian.meyer@math.tu-dortmund.de
TU Dortmund, Fakultät für Mathematik, Germany, ORCID iD: 0000-0003-3340-2316

problem formulation we use here is the one of Kantorovich [15]. Let us fix some notation and formulate the problem: Let $\Omega_1 \subset \mathbb{R}^{d_1}$, $\Omega_2 \subset \mathbb{R}^{d_2}$ be two compact domains, denote $\Omega = \Omega_1 \times \Omega_2$, and assume we are given two positive regular Radon measures μ_1 and μ_2 on Ω_1 and Ω_2 , respectively. Further we assume that a cost function $c : \Omega_1 \times \Omega_2 \rightarrow \mathbb{R}$ is given that models the cost of transporting a unit of mass from $x_1 \in \Omega_1$ to $x_2 \in \Omega_2$. The optimal transport problem asks to find a transport plan π , which is a Radon measure on Ω , such that it has minimal overall transport cost $\int_{\Omega} c(x_1, x_2) d\pi(x_1, x_2)$ among all measures π which have μ_1 and μ_2 as first and second marginals, respectively, i.e. for all Borel sets $A \in \Omega_1$ it holds that $\pi(A \times \Omega_2) = \mu_1(A)$ and for all Borel sets $B \in \Omega_2$ it holds that $\pi(\Omega_1 \times B) = \mu_2(B)$. This problem has been studied extensively and we refer to the books [18, 19, 23, 24, 21]. One particular result is, that an optimal plan π^* exists and that the support of optimal plans is contained in the so-called c -superdifferential of a c -concave function [1, Theorem 1.13]. For many cost functions c , this means that optimal transport plans are supported on small sets and that they are in fact singular with respect to the Lebesgue measure on Ω . This makes the numerical treatment of optimal transport problems difficult and one can employ regularization to obtain approximately optimal plans π that are functions on Ω . The regularization method that has got the most attention recently is regularization with the negative entropy of π and we refer to [16, 10, 4]. Entropic regularization has gotten popular in machine learning applications due to the fact that it allows for the very simple Sinkhorn algorithm (in the discrete case), see [9, 13] and also [17] for a recent and thorough review of the computational aspects of optimal transport.

Regularizations different from entropic regularization has been much less studied. We are only aware of works in the discrete case, e.g. [3, 11]. In this work we will investigate the case where we regularize the problem in $L^2(\Omega)$. The paper is organized as follows: In Section 2 we state the problem and analyze existence and duality. It will turn out that existence of solutions of the dual problem will be quite tricky to show, but we will show that dual solutions exist in respective L^2 spaces and that a straightforward optimality system characterizes primal-dual optimality. In Section 3 we derive two different algorithms for the discrete version of the quadratically regularized optimal transport problem, and in Section 4 we comment on a simple discretization scheme and report numerical examples.

Notation. We will abbreviate $x_+ = \max(x, 0)$ (and will apply this also to functions and to measures where $_+$ will mean the positive part from the Hahn-Jordan decomposition). By $C(\Omega)$ we denote that space of continuous functions on Ω (and we will always work on compact sets) equipped with the supremum norm $\|\cdot\|_{\infty}$ and by $\mathfrak{M}(\Omega)$ we denote the space of Radon measures on a compact domain and we use the norm $\|\mu\|_{\mathfrak{M}} = \sup\{\int f d\mu \mid f \in C(\Omega), |f| \leq 1\}$. The Lebesgue measure will be λ (and we also use λ_1 and λ_2 to specify the Lebesgue measure on sets Ω_1 and Ω_2 , respectively). For convenience, we use $|\Omega|$ for the Lebesgue measure of the set Ω . Furthermore, for a Radon measure $w \in \mathfrak{M}$, we denote the absolutely and singular part arising from the Lebesgue decomposition with respect to the Lebesgue measure by w_{ac} and w_s , i.e. they satisfy $w_{ac} \ll \lambda$ and $w_s \perp \lambda$. Duality pairings are denoted by $\langle \cdot, \cdot \rangle$. If both arguments of the duality pairing are positive and the duality pairing does not necessarily exist, e.g. for $\psi \in \mathfrak{M}(\Omega)$ and $x \in L^2(\Omega)$, we set $\langle \psi, x \rangle := +\infty$.

2 Quadratic regularization in the continuous case

For the quadratically regularized optimal transport problem we seek a transport plan $\pi \in L^2(\Omega_1 \times \Omega_2)$ which for a given cost function $c \in L^2(\Omega_1 \times \Omega_2)$, a regularization parameter $\gamma > 0$, and given functions $\mu_i \in L^2(\Omega_i)$, $i = 1, 2$ solves

$$\begin{aligned} \min_{\pi} \langle c, \pi \rangle_{L^2} + \frac{\gamma}{2} \|\pi\|_{L^2}^2 \quad \text{subject to} \quad & \int_{\Omega_2} \pi(x_1, x_2) \, d\lambda_2 = \mu_1(x_1), \\ & \int_{\Omega_1} \pi(x_1, x_2) \, d\lambda_1 = \mu_2(x_2), \\ & \pi(x_1, x_2) \geq 0 \end{aligned} \quad (1)$$

where the constraints are understood pointwise almost everywhere.

2.1 Solutions of the primal problem

It is straight forward to show, that optimal transport plans exist:

Lemma 2.1 *Problem (1) has an optimal solution if and only if $\mu_1 \in L^2(\Omega_1)$, $\mu_2 \in L^2(\Omega_2)$, $\mu_1, \mu_2 \geq 0$ almost everywhere, and $\int_{\Omega_1} \mu_1(x_1) \, d\lambda_1 = \int_{\Omega_2} \mu_2(x_2) \, d\lambda_2$.*

Proof Assume that there is an optimal solution $\pi^* \in L^2(\Omega_1 \times \Omega_2)$. By Jensen's inequality we get

$$\begin{aligned} \int_{\Omega_1} \mu_1^2(x_1) \, d\lambda_1 &= \int_{\Omega_1} \left(\int_{\Omega_2} \pi^*(x_1, x_2) \, d\lambda_2 \right)^2 \, d\lambda_1 \\ &\leq |\Omega_2| \iint_{\Omega_1 \times \Omega_2} \pi^*(x_1, x_2)^2 \, d\lambda_1 \, d\lambda_2 < \infty \end{aligned}$$

which shows $\mu_1 \in L^2(\Omega_1)$. The argument for μ_2 is similar. Non-negativity of μ_1 and μ_2 follows from non-negativity of π^* . Finally, by Fubini's theorem

$$\begin{aligned} \int_{\Omega_1} \mu_1(x_1) \, d\lambda_1 &= \iint_{\Omega_1 \times \Omega_2} \pi^*(x_1, x_2) \, d\lambda_1 \, d\lambda_2 \\ &= \int_{\Omega_2} \mu_2(x_2) \, d\lambda_2 \end{aligned}$$

Conversely, if $\mu_1 \in L^2(\Omega_1)$ and $\mu_2 \in L^2(\Omega_2)$ and $\mu_1, \mu_2 \geq 0$ we set $C := \int_{\Omega_1} \mu_1(x_1) \, d\lambda_1 = \int_{\Omega_2} \mu_2(x_2) \, d\lambda_2$. Then $\pi(x_1, x_2) = \frac{1}{C} \mu_1(x_1) \mu_2(x_2)$ is feasible for (1) and since the objective is continuous, coercive, and strongly convex a (unique) minimizer exists. \square

2.2 Dual problem and existence of dual solutions

In the following section, we apply the classical Lagrange duality to the linear-quadratic program (1). To this end, let us define the Lagrangian associated with (1). In order to shorten the notation, we set

$$\mu := \gamma \mu_1 \otimes \mu_2.$$

Furthermore, we define

$$P_1 : L^2(\Omega) \ni \pi \mapsto \int_{\Omega_2} \pi \, d\lambda_2 \in L^2(\Omega_1), \quad P_2 : L^2(\Omega) \ni \pi \mapsto \int_{\Omega_1} \pi \, d\lambda_1 \in L^2(\Omega_2), \quad (2)$$

and denote the the primal objective by

$$E_\gamma : L^2(\Omega) \rightarrow \mathbb{R}, \quad E_\gamma(\pi) := \int_{\Omega} c \pi \, d\lambda + \frac{\gamma}{2} \|\pi\|_{L^2(\Omega)}^2. \quad (3)$$

Then, the Lagrangian associated with (1) is given by

$$\begin{aligned} \mathcal{L} : L^2(\Omega) \times L^2(\Omega_1) \times L^2(\Omega_2) \times L^2(\Omega) &\rightarrow \mathbb{R}, \\ \mathcal{L}(\pi, \alpha_1, \alpha_2, \varrho) &:= E_\gamma(\pi) - \langle \varrho, \pi \rangle_{L^2(\Omega)} \\ &\quad + \langle \alpha_1, P_1 \pi - \mu_1 \rangle_{L^2(\Omega_1)} + \langle \alpha_2, P_2 \pi - \mu_2 \rangle_{L^2(\Omega_2)}. \end{aligned}$$

Then, by standard arguments, the primal problem in (1) is equivalent to

$$\inf_{\pi \in L^2(\Omega)} \sup_{\substack{\alpha_1 \in L^2(\Omega_1), \alpha_2 \in L^2(\Omega_2) \\ \varrho \in L^2(\Omega), \varrho \geq 0}} \mathcal{L}(\pi, \alpha_1, \alpha_2, \varrho), \quad (\text{PP})$$

while its (Lagrangian) dual is given by

$$\sup_{\substack{\alpha_1 \in L^2(\Omega_1), \alpha_2 \in L^2(\Omega_2) \\ \varrho \in L^2(\Omega), \varrho \geq 0}} \inf_{\pi \in L^2(\Omega)} \mathcal{L}(\pi, \alpha_1, \alpha_2, \varrho). \quad (\text{DP})$$

The main part of the upcoming analysis is devoted to the existence of solutions to (DP). Once this is established, the necessary and sufficient optimality condition associated with (1) in form of the variational inequality will allow us to derive an optimality system that is also amenable for numerical computations.

To show existence for (DP), we first reformulate the dual problem. Since \mathcal{L} is quadratic w.r.t. π , the inner inf-problem is solved by

$$\pi = \frac{1}{\gamma} (\rho + \alpha_1 \oplus \alpha_2 - c), \quad (4)$$

where the mapping $\oplus : L^2(\Omega_1) \times L^2(\Omega_2) \rightarrow L^2(\Omega)$ is defined via

$$(v_1 \oplus v_2)(x_1, x_2) := v_1(x_1) + v_2(x_2) \quad (5)$$

for almost all $(x_1, x_2) \in \Omega$ and all $v_i \in L^2(\Omega_i)$, $i = 1, 2$.

Remark 2.2 The map \oplus is related to the adjoints of the projections P_1 and P_2 from (2) by $\alpha_1 \oplus \alpha_2 = P_1^* \alpha_1 + P_2^* \alpha_2$.

Inserting (4) into (DP) yields

$$\sup_{\alpha_1 \in L^2(\Omega_1), \alpha_2 \in L^2(\Omega_2)} \sup_{\rho \geq 0} \left(-\frac{1}{2\gamma} \int_{\Omega} (\rho + \alpha_1 \oplus \alpha_2 - c)^2 \, d\lambda + \int_{\Omega_1} \mu_1 \alpha_1 \, d\lambda_1 + \int_{\Omega_2} \mu_2 \alpha_2 \, d\lambda_2 \right) \quad (6)$$

Again, the inner optimization problem is quadratic w.r.t. ρ so that its solution is given by

$$\rho = -(\alpha_1 \oplus \alpha_2 - c)_-. \quad (7)$$

Inserted in (6), this results in the following dual problem

$$\left. \begin{array}{l} \min \quad \Phi(\alpha_1, \alpha_2) := \frac{1}{2} \|(\alpha_1 \oplus \alpha_2 - c)_+\|_{L^2(\Omega)}^2 \\ \quad \quad \quad \quad \quad \quad \quad \quad -\gamma \langle \alpha_1, \mu_1 \rangle - \gamma \langle \alpha_2, \mu_2 \rangle \\ \text{s.t.} \quad \alpha_i \in L^2(\Omega_i), i = 1, 2. \end{array} \right\} \quad (D)$$

To prove existence of solutions for this problem, we need to require the following

Assumption 1 *The domains Ω_1 and Ω_2 are compact. Moreover, the cost function c is in $L^2(\Omega)$ and fulfills $c \geq \underline{c} > -\infty$. Furthermore, the marginals μ_1 and μ_2 satisfy $\mu_i \in L^2(\Omega_i)$ and $\mu_i \geq \delta > 0$, $i = 1, 2$. In addition we assume that $\int_{\Omega_1} \mu_2 d\lambda_1 = \int_{\Omega_1} \mu_2 d\lambda_1 = 1$.*

Remark 2.3 The last assumption on the normalization of the marginals is just to ease the subsequent analysis and can be relaxed by $\int_{\Omega_1} \mu_2 d\lambda_1 = \int_{\Omega_2} \mu_2 d\lambda_1$, which is needed anyway to ensure the existence of a solution to the primal problem, see Lemma 2.1.

Remark 2.4 Note that there is an obvious source of non-uniqueness for the dual problem (D): We can add a constant to α_1 and subtract it from α_2 and this does not change the dual objective, i.e for any constant C it holds that $\Phi(\alpha_1 + C, \alpha_2 - C) = \Phi(\alpha_1, \alpha_2)$. This non-uniqueness will not cause trouble in the proofs and when convenient, we remove it, e.g. by demanding that $\int_{\Omega_2} \alpha_2 d\lambda_2 = 0$.

Observe that the objective Φ in (D) is also well defined for functions in $\alpha_i \in L^1(\Omega_i)$ with $(\alpha_1 \oplus \alpha_2 - c)_+ \in L^2(\Omega)$. This gives rise to the following auxiliary dual problem:

$$\left. \begin{array}{l} \min \quad \Phi(\alpha_1, \alpha_2) \\ \text{s.t.} \quad \alpha_i \in L^1(\Omega_i), i = 1, 2, \quad (\alpha_1 \oplus \alpha_2 - c)_+ \in L^2(\Omega). \end{array} \right\} \quad (D')$$

Our strategy to prove existence of solutions to (D) is now as follows:

1. First, we show that (D') admits a solution $(\alpha_1^*, \alpha_2^*) \in L^1(\Omega_1) \times L^1(\Omega_2)$, see Proposition 2.9.
2. Then, we prove that α_1^* and α_2^* possess higher regularity, namely that they are functions in $L^2(\Omega_i)$, $i = 1, 2$, cf. Theorem 2.10.
3. Thus, (α_1^*, α_2^*) is feasible for (D) and, since the feasible set of (D') contains the one of (D), while the objective of (D') restricted to L^2 -functions coincides with the objective in (D), this finally gives that (α_1^*, α_2^*) is indeed optimal for (D).

The reason to consider (D') is essentially that the objective Φ is not coercive in $L^2(\Omega)$, but only in $L^1(\Omega)$ (at least w.r.t. the negative part of α_i). Therefore, we have to deal with weakly* converging sequences in the space of Radon measures within the proof of existence of solutions. For this purpose, we need to extend the objective to a suitable set. To that end, let us define

$$G : L^2(\Omega) \ni w \mapsto \int_{\Omega} \frac{1}{2} w_+^2 - w\mu d\lambda \in \mathbb{R}. \quad (8)$$

Note that, thanks to $\int_{\Omega_1} \mu_2 d\lambda_1 = \int_{\Omega_1} \mu_2 d\lambda_1 = 1$, it holds

$$\Phi(\alpha_1, \alpha_2) = G(\alpha_1 \oplus \alpha_2 - c) - \int_{\Omega} c \mu d\lambda \quad \forall \alpha_i \in L^2(\Omega_i), i = 1, 2. \quad (9)$$

Of course, G is also well defined as a functional on the feasible set of (D') and we will denote this functional by the same symbol to ease notation. In order to extend G to the space of Radon measures, consider for a given measure $w \in \mathfrak{M}(\Omega)$, the Hahn-Jordan decomposition $w = w_+ - w_-$ and assume that $w_+ \in L^2(\Omega)$. Then, we set $G(w) = \int_{\Omega} \frac{1}{2} w_+^2 d\lambda - \int_{\Omega} \mu dw$. With a slight abuse of notation, we denote this mapping by G , too. Furthermore, for $w_+ \in L^2(\Omega)$, $-\int_{\Omega} w_+ \mu d\lambda$ is finite for $\mu \in L^2(\Omega)$ as in Assumption 1. Regarding, the negative part, we define $\int_{\Omega} \mu dw_- := \infty$, where this expression is not properly defined, as w_- and μ are both positive. Combining this, we obtain that $-\int_{\Omega} \mu dw \in \mathbb{R} \cup \{\infty\}$.

Note in this context that, if the singular part of w (w.r.t. the Lebesgue measure) vanishes, then also $w_+ \in L^1(\Omega)$ and $w_+(x) = \max\{0, w(x)\}$ λ -a.e. in Ω so that both functionals coincide on $L^2(\Omega)$, which justifies this notation. Furthermore, we also generalize the map \oplus to the measure space by setting

$$\alpha_1 \oplus \alpha_2 := \alpha_1 \otimes \lambda_2 + \lambda_1 \otimes \alpha_2, \quad \alpha_i \in \mathfrak{M}(\Omega_i), i = 1, 2.$$

Again, it is easily seen that, for $\alpha_i \in L^2(\Omega_i)$, $i = 1, 2$, this definition boils down to the one in (5). Also Remark 2.2 applies in that we can express $\alpha_1 \oplus \alpha_2$ in terms of the adjoints of P_1 and P_2 from (2) when defined appropriately.

The next lemma is rather obvious and covers the coercivity of G in $L^1(\Omega)$ as indicated above.

Lemma 2.5 *Let Assumption 1 hold and suppose that a sequence $\{w^n\} \subset L^2(\Omega)$ fulfills*

$$G(w^n) \leq C < \infty \quad \forall n \in \mathbb{N}.$$

Then, the sequences $\{w_+^n\}$ and $\{w_-^n\}$ are bounded in $L^2(\Omega)$ and $L^1(\Omega)$, respectively.

Proof We rewrite G as $G(w) = \int_{\Omega} \frac{1}{2} w_+^2 - w_+ \mu d\lambda + \int_{\Omega} w_- \mu d\lambda$. The positivity of μ then implies

$$\|w_+^n\|_{L^2(\Omega)}^2 = G(w^n) + \int_{\Omega} w_+^n \mu d\lambda - \int_{\Omega} w_-^n \mu d\lambda \leq C + \|\mu\|_{L^2(\Omega)} \|w_+^n\|_{L^2(\Omega)},$$

which gives the first assertion. To see the second one, we use $\mu \geq \delta$ to estimate

$$\begin{aligned} C \geq G(w^n) &= \int_{\Omega} \frac{1}{2} (w_+^n - \mu)^2 d\lambda - \int_{\Omega} \mu^2 / 2 d\lambda + \int_{\Omega} w_-^n \mu d\lambda \\ &\geq - \int_{\Omega} \mu^2 / 2 d\lambda + \delta \|w_-^n\|_{L^1(\Omega)}, \end{aligned}$$

which finishes the proof. \square

The next lemma provides a lower semicontinuity result for G w.r.t. weak* convergence in $\mathfrak{M}(\Omega)$. Note that, here, we need the extension of G as introduced above.

Lemma 2.6 *Let Assumption 1 be fulfilled and a sequence $\{w_n\} \subset L^2(\Omega)$ be given such that $w^n \rightharpoonup^* w^*$ in $\mathfrak{M}(\Omega)$ and $G(w^n) \leq C < \infty$ for all $n \in \mathbb{N}$. Then there holds $w_+^* \in L^2(\Omega)$ and*

$$G(w^*) \leq \liminf_{n \rightarrow \infty} G(w^n). \quad (10)$$

Proof By virtue of Lemma 2.5, $\{w_+^n\}$ is bounded in $L^2(\Omega)$ and thus, there is a subsequence of $\{w_+^n\}$, to ease notation denoted by the same symbol, that converges weakly in $L^2(\Omega)$ to some $\theta_+ \in L^2(\Omega)$. Since the set $\{v \in L^2(\Omega) : v \geq 0 \text{ a.e. in } \Omega\}$ is clearly weakly closed, we have $\theta_+ \geq 0$ a.e. in Ω . With a little abuse of notation, we denote the Radon measure induced by $C(\Omega) \ni \varphi \mapsto \int_{\Omega} \theta_+ \varphi \, d\lambda \in \mathbb{R}$ by θ_+ , too. If we define $\theta_- := \theta_+ - w^* \in \mathfrak{M}(\Omega)$, then $w^n = w_+^n - w^n \rightharpoonup^* \theta_-$ in $\mathfrak{M}(\Omega)$ with $\theta_- \geq 0$. Thus we have $w^* = \theta_+ - \theta_-$ with two positive Radon measures θ_+, θ_- . The maximality property of the Hahn-Jordan decomposition then implies $w_+^* \leq \theta_+$. Since θ_+ is absolutely continuous w.r.t. λ , the same thus holds for w_+^* , i.e. $w_+^* \in L^1(\Omega)$. Applying again $w_+^* \leq \theta_+$, which clearly also holds for the densities pointwise λ -almost everywhere, we moreover deduce from the weak convergence of w_+^n in $L^2(\Omega)$ that

$$\int_{\Omega} (w_+^*)^2 \, d\lambda \leq \int_{\Omega} (\theta_+)^2 \, d\lambda \leq \liminf_{n \rightarrow \infty} \int_{\Omega} (w_+^n)^2 \, d\lambda, \quad (11)$$

which implies $w_+^* \in L^2(\Omega)$ as claimed. Since the above reasoning applies to every subsequence w_+^n that is weakly converging in $L^2(\Omega)$, (11) holds for the whole sequence $\{w_+^n\}$, which together with the weak* convergence of w^n and the definition of G , gives (10). \square

Before we are in the position to prove existence for (D'), we need two additional results on the \oplus -operator in the space of Radon measures.

Lemma 2.7 *If $\alpha_i \in \mathfrak{M}(\Omega_i)$, $i = 1, 2$ and $\int_{\Omega_2} d\alpha_2 = 0$, then it holds that*

$$\|\alpha_1\|_{\mathfrak{M}} \leq \frac{1}{|\Omega_2|} \|\alpha_1 \oplus \alpha_2\|_{\mathfrak{M}} \quad \text{and} \quad \|\alpha_2\|_{\mathfrak{M}} \leq \frac{2}{|\Omega_1|} \|\alpha_1 \oplus \alpha_2\|_{\mathfrak{M}}$$

Proof We estimate

$$\begin{aligned} \|\alpha_1 \oplus \alpha_2\|_{\mathfrak{M}} &= \sup_{\|\phi\|_{\infty} \leq 1} \iint_{\Omega_1 \times \Omega_2} \phi(x_1, x_2) \, d(\alpha_1(x_1) + \alpha_2(x_2)) \\ &\geq \sup_{\substack{\|\phi_1\|_{\infty} \leq 1 \\ \|\phi_2\|_{\infty} \leq 1}} \iint_{\Omega_1 \times \Omega_2} \phi_1(x_1) \phi_2(x_2) \, d(\alpha_1(x_1) + \alpha_2(x_2)) \\ &= \sup_{\substack{\|\phi_1\|_{\infty} \leq 1 \\ \|\phi_2\|_{\infty} \leq 1}} \left[\iint_{\Omega_1 \times \Omega_2} \phi_1(x_1) \phi_2(x_2) \, d\alpha_1(x_1) \, d\lambda_2 \right. \\ &\quad \left. + \iint_{\Omega_1 \times \Omega_2} \phi_1(x_1) \phi_2(x_2) \, d\lambda_1 \, d\alpha_2(x_2) \right]. \end{aligned} \quad (12)$$

Taking $\phi_2 \equiv 1$ and using $\int_{\Omega_2} d\alpha_2(x_2) = 0$ gives

$$\begin{aligned} \|\alpha_1 \oplus \alpha_2\|_{\mathfrak{M}} &\geq \sup_{\|\phi_1\|_{\infty} \leq 1} \int_{\Omega_1} \phi_1(x_1) d\alpha_1(x_1) |\Omega_2| + \int_{\Omega_2} d\alpha_2(x_2) \int_{\Omega_1} \phi_1(x_1) d\lambda_1 \\ &= |\Omega_2| \|\alpha_1\|_{\mathfrak{M}}. \end{aligned}$$

Now we start again at (12) and estimate from below by taking $\phi_1 \equiv 1$ to get

$$\begin{aligned} \|\alpha_1 \oplus \alpha_2\|_{\mathfrak{M}} &\geq \sup_{\|\phi_2\|_{\infty} \leq 1} \int_{\Omega_1} d\alpha_1(x_1) \int_{\Omega_2} \phi_2(x_2) d\lambda_2 + \int_{\Omega_2} \phi_2(x_2) d\alpha(x_2) |\Omega_1| \\ &\geq -|\Omega_2| \int_{\Omega_1} d\alpha_1(x_1) + |\Omega_1| \|\alpha_2\|_{\mathfrak{M}} \end{aligned}$$

which implies

$$|\Omega_1| \|\alpha_2\|_{\mathfrak{M}} \leq \|\alpha_1 \oplus \alpha_2\|_{\mathfrak{M}} + |\Omega_2| \|\alpha_1\|_{\mathfrak{M}}$$

which completes the proof. \square

The next lemma will be used to show that the negative part of the minimizer of (D) does not have a singular part.

Lemma 2.8 *Let $c \in L^1(\Omega)$ and $\alpha_i \in \mathfrak{M}(\Omega_i)$ for $i \in \{1, 2\}$ with Lebesgue decompositions, $\alpha_i = f_i + \eta_i$ satisfying $f_i \ll \lambda$ and $\eta_i \perp \lambda$ for $i \in \{1, 2\}$.*

1. *It holds that*

$$(\alpha_1 \oplus \alpha_2 - c)_+ = (f_1 \oplus f_2 - c + (\eta_1)_+ \oplus (\eta_2)_+)_+. \quad (13)$$

2. *If $(\alpha_i)_+$ is absolutely continuous for $i = 1, 2$, then for $\tilde{\alpha}_i = \alpha_i - (\eta_i)_-$ for $i = 1, 2$, it holds that*

$$\Phi(\tilde{\alpha}_1, \tilde{\alpha}_2) \leq \Phi(\alpha_1, \alpha_2).$$

Proof We first proof point 1. The measures f_i, η_i exist by Lebesgue's decomposition theorem, see Theorem 1.155 in [12]. We combine these decompositions with $\alpha_1 \oplus \alpha_2 = \alpha_1 \otimes \lambda + \lambda \otimes \alpha_2$ to arrive at Lebesgue's decomposition of $\alpha_1 \oplus \alpha_2$ with respect to $\lambda \otimes \lambda$, namely

$$\alpha_1 \oplus \alpha_2 - c = f_1 \oplus f_2 - c + \eta_1 \oplus \eta_2 \quad (14)$$

$$f_1 \oplus f_2 - c \ll \lambda \otimes \lambda \quad (15)$$

$$\eta_1 \oplus \eta_2 \perp \lambda \otimes \lambda \quad (16)$$

(which holds true because $c \in L^1(\Omega) \hookrightarrow \mathfrak{M}(\Omega)$). Now, we consider the Hahn-Jordan decomposition of η_1 ,

$$\begin{aligned} \eta_1 &= (\eta_1)_+ - (\eta_1)_- \\ (\eta_1)_+ &\perp (\eta_1)_-, \end{aligned} \quad (17)$$

and obtain from (14) that

$$\begin{aligned} \alpha_1 \oplus \alpha_2 - c &= (f_1 + \eta_1) \oplus (f_2 + \eta_2) - c \\ &= f_1 \oplus f_2 + \eta_1 \oplus \eta_2 - c \\ &= f_1 \oplus f_2 + ((\eta_1)_+ - (\eta_1)_-) \oplus \eta_2 - c \\ &= f_1 \oplus f_2 + (\eta_1)_+ \otimes \lambda - (\eta_1)_- \otimes \lambda + \lambda \otimes \eta_2 - c \\ &= f_1 \oplus f_2 - c + (\eta_1)_+ \oplus \eta_2 - (\eta_1)_- \otimes \lambda. \end{aligned}$$

Furthermore,

$$(\eta_1)_- \otimes \lambda \perp f_1 \oplus f_2 - c + (\eta_1)_+ \oplus \eta_2$$

where the singularity with respect to $f_1 \oplus f_2 - c$ is due to (15) and (16) and the singularity with respect to $(\eta_1)_+ \oplus \eta_2$ is due to (17). Thus,

$$\begin{aligned} (\alpha_1 \oplus \alpha_2 - c)_- &= (f_1 \oplus f_2 - c + (\eta_1)_+ \oplus \eta_2)_- + (-(\eta_1)_- \otimes \lambda)_- \\ &= (f_1 \oplus f_2 - c + (\eta_1)_+ \oplus \eta_2)_- + (\eta_1)_- \otimes \lambda. \end{aligned}$$

as $(\eta_1)_- \otimes \lambda$ is a positive measure. Consequently,

$$(\alpha_1 \oplus \alpha_2 - c)_+ = (f_1 \oplus f_2 - c + (\eta_1)_+ \oplus \eta_2)_+.$$

Repeating this argument with the Hahn-Jordan decomposition of η_2 yields the claim.

The second part of the lemma is a direct consequence of the first: Since $(\alpha_1 \oplus \alpha_2 - c)_+ = (\tilde{\alpha}_1 \oplus \tilde{\alpha}_2 - c)_+$, the first summand in the functional Φ is equal for α_i and $\tilde{\alpha}_i$. However, the second summand in Φ can not decrease since $\tilde{\alpha}_i \leq \alpha_i$, $\mu_i \geq 0$ and $\gamma\langle(\eta_i)_-, \mu_i\rangle = \infty$ if the duality pairing does not exist. \square

Now we are ready to prove the existence result for (D'):

Proposition 2.9 *Under Assumption 1 the minimization problem (D') admits a solution $(\alpha_1^*, \alpha_2^*) \in L^1(\Omega_1) \times L^1(\Omega_2)$.*

Proof We proceed via the classical direct method of the calculus of variations. For this purpose, let $\{(\alpha_1^n, \alpha_2^n)\} \subset L^1(\Omega_1) \times L^1(\Omega_2)$ with $(\alpha_1^n \oplus \alpha_2^n - c)_+ \in L^2(\Omega)$ be a minimizing sequence for (D'), where we shift α_1 and α_2 by adding and subtracting constants such that we obtain $\int_{\Omega_2} \alpha_2 d\lambda_2 = 0$. Note that, due to its additive structure, this does not change the objective Φ in (D'), cf. Remark 2.4.

Next, let us define $w^n := \alpha_1^n \oplus \alpha_2^n - c$. Then, thanks to (9) and Lemma 2.5, the sequence $\{w^n\}$ is bounded in $L^1(\Omega)$. Hence, there is a weakly* converging subsequence, which we denote by the same symbol w.l.o.g., i.e. $w^n \rightharpoonup^* \tilde{w}$ in $\mathfrak{M}(\Omega)$. Now, Lemma 2.6 applies giving that

$$\tilde{w}_+ \in L^2(\Omega), \tag{18}$$

$$G(\tilde{w}) \leq \liminf_{n \rightarrow \infty} G(w^n). \tag{19}$$

Since $\{w^n\}$ is bounded in $\mathfrak{M}(\Omega)$, the same holds for $\{\alpha_1^n \oplus \alpha_2^n\}$ and, as α_2^n is normalized, Lemma 2.7 gives that $\{\alpha_i^n\}$ is bounded in $\mathfrak{M}(\Omega_i)$, $i = 1, 2$. Therefore, we can select a further (sub-)subsequence, still denoted by the same symbol to ease notation, such that

$$\alpha_i^n \rightharpoonup^* \tilde{\alpha}_i \quad \text{in } \mathfrak{M}(\Omega_i), \quad i = 1, 2.$$

Since the mapping $\mathfrak{M}(\Omega_1) \times \mathfrak{M}(\Omega_2) \ni (\alpha_1, \alpha_2) \mapsto \alpha_1 \oplus \alpha_2 \in \mathfrak{M}(\Omega)$ is the adjoint of the projection mapping $C(\Omega) \ni \varphi \mapsto (\int_{\Omega_2} \varphi d\lambda_2, \int_{\Omega_1} \varphi d\lambda_1) \in C(\Omega_1) \times C(\Omega_2)$, see Remark 2.2, it is weakly* continuous so that

$$\tilde{w} = \tilde{\alpha}_1 \oplus \tilde{\alpha}_2 - c. \tag{20}$$

Next, we investigate the singular parts of $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$. We start with the positive part and employ Lebesgue's decomposition of $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$:

$$\tilde{\alpha}_i = \alpha_i^* + \tilde{\eta}_i, \quad \alpha_i^* \ll \lambda_i, \quad \tilde{\eta}_i \perp \lambda_i, \quad i = 1, 2.$$

In the following we will see that the regular parts $\alpha_i^* \in L^1(\Omega_i)$, $i = 1, 2$, are exactly the solution of (D'). For this purpose, we first show that the positive parts of $\tilde{\eta}_1$ and $\tilde{\eta}_2$ vanish. We have $\alpha_1^* \oplus \alpha_2^* - c \ll \lambda$, $\tilde{\eta}_1 \oplus \tilde{\eta}_2 \perp \lambda$, and, by uniqueness of Lebesgue's decomposition, $\tilde{w}_s = \tilde{\eta}_1 \oplus \tilde{\eta}_2$. But from (18), we know that $(\tilde{w}_s)_+ = 0$. Combining this fact with Lemma 2.8, applied to the case $f_1 = 0$, $f_2 = 0$, and $c = 0$, we obtain

$$(\tilde{\eta}_1 \oplus \tilde{\eta}_2)_+ = (\tilde{\eta}_1)_+ \oplus (\tilde{\eta}_2)_+.$$

and consequently, $(\tilde{\eta}_i)_+ = 0$ for $i = 1, 2$ by positivity. Therefore, $(\tilde{\alpha}_i)_+$ are L^1 -functions rather than measures and

$$\tilde{w}_+ = (\tilde{\alpha}_1 \oplus \tilde{\alpha}_2 - c)_+ = (\alpha_1^* \oplus \alpha_2^* - c)_+ \quad (21)$$

Now Lemma 2.8 shows feasibility of (α_1^*, α_2^*) for (D') and we also see that

$$\begin{aligned} \Phi(\alpha_1^*, \alpha_2^*) &= \frac{1}{2} \int_{\Omega} (\alpha_1^* \oplus \alpha_2^* - c)_+^2 d\lambda - \gamma \int_{\Omega_1} \mu_1 \alpha_1^* d\lambda_1 - \gamma \int_{\Omega_2} \mu_2 \alpha_2^* d\lambda_2 \\ &\leq G(\tilde{\alpha}_1 \oplus \tilde{\alpha}_2 - c) - \int_{\Omega} c \mu d\lambda \\ &= G(\tilde{w}) - \int_{\Omega} c \mu d\lambda \\ &\leq \liminf_{n \rightarrow \infty} \Phi(\alpha_1^n, \alpha_2^n), \end{aligned} \quad (22)$$

which demonstrates the optimality of (α_1^*, α_2^*) . \square

In the following, we assume that $\int_{\Omega_2} \alpha_2^* d\lambda_2 = 0$. If this is not the case, then we can again shift α_1^* and α_2^* without changing the value of Φ , cf. Remark 2.4.

Theorem 2.10 *Let Assumption 1 hold. Then every optimal dual solution (α_1^*, α_2^*) from Proposition 2.9 satisfies $\alpha_i^* \in L^2(\Omega_i)$, $i = 1, 2$, and is therefore also a solution of the original dual problem (D). Moreover, the negative parts of α_i^* are bounded and the function $\frac{1}{\gamma} ((\alpha_1^* \oplus \alpha_2^*) - c)_+$ has marginals the μ_1 and μ_2 .*

Proof We again consider the positive and the negative part separately and start with $(\alpha_1^*)_+$. Let $\varphi \in C_c^\infty(\Omega_1)$ and $t > 0$ be fixed, but arbitrary. Then, thanks to

$$0 \leq ((\alpha_1^* + t\varphi) \oplus \alpha_2^* - c)_+ \leq (\alpha_1^* \oplus \alpha_2^* - c)_+ + t\varphi_+,$$

Proposition 2.9 implies that $((\alpha_1^* + t\varphi) \oplus \alpha_2^* - c)_+ \in L^2(\Omega)$ so that $(\alpha_1^* + t\varphi, \alpha_2^*)$ is feasible for (D'). Therefore, the optimality of (α_1^*, α_2^*) for (D') yields

$$\frac{1}{2} \int_{\Omega} \frac{1}{t} \left(((\alpha_1^* + t\varphi) \oplus \alpha_2^* - c)_+^2 - (\alpha_1^* \oplus \alpha_2^* - c)_+^2 \right) d\lambda - \gamma \int_{\Omega_1} \mu_1 \varphi d\lambda_1 \geq 0 \quad \forall t > 0.$$

Owing to the continuous differentiability of $\mathbb{R} \ni r \mapsto r_+^2 \in \mathbb{R}$, the first integrand converges to $2(\alpha_1^* \oplus \alpha_2^* - c)_+ \varphi$ λ -a.e. in Ω for $t \searrow 0$. Moreover, the Lipschitz continuity of the max-function gives that

$$\frac{1}{t} \left(((\alpha_1^* + t\varphi) \oplus \alpha_2^* - c)_+^2 - (\alpha_1^* \oplus \alpha_2^* - c)_+^2 \right) \leq |\varphi|^2 + 2|\varphi|(\alpha_1^* \oplus \alpha_2^* - c)_+ \quad \text{a.e. in } \Omega$$

holds for $0 < t \leq 1$. Hence, due to Lebesgue's dominated convergence theorem, we are allowed to pass to the limit $t \searrow 0$ and obtain in this way

$$\int_{\Omega_1} \left(\int_{\Omega_2} (\alpha_1^* \oplus \alpha_2^* - c)_+ d\lambda_2 - \gamma \mu_1 \right) \varphi d\lambda_1 \geq 0.$$

Since $\varphi \in C_c^\infty(\Omega)$ was arbitrary, the fundamental lemma of the calculus of variations thus gives

$$\int_{\Omega_2} (\alpha_1^* \oplus \alpha_2^* - c)_+ d\lambda_2 = \gamma \mu_1 \quad \lambda_1\text{-a.e. in } \Omega_1. \quad (23)$$

Next, define the following sequence of functions in $L^1(\Omega_2)$:

$$f_n(x_2) := (-n + \alpha_2^*(x_2) - \underline{c})_+, \quad n \in \mathbb{N},$$

where \underline{c} is the lower bound for c from Assumption 1. Then we have $f_n \geq 0$ λ_2 -a.e. Ω_2 and $f_n \searrow 0$ λ_2 -a.e. in Ω_2 so that the monotone convergence theorem gives

$$\int_{\Omega_2} (-n + \alpha_2^*(x_2) - \underline{c})_+ d\lambda_2 = \int_{\Omega_2} f_n(x_2) d\lambda_2 \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Thus there exists $N \in \mathbb{N}$ such that

$$\int_{\Omega_2} (-N + \alpha_2^*(x_2) - \underline{c})_+ d\lambda_2 < \gamma \delta, \quad (24)$$

where $\delta > 0$ is the threshold for μ_1 from Assumption 1. Now assume that $\alpha_1^* \leq -N$ λ_1 -a.e. on a set of $E \subset \Omega_1$ of positive Lebesgue measure. Then

$$\int_{\Omega_2} (\alpha_1^* \oplus \alpha_2^* - c)_+ d\lambda_2 \leq \int_{\Omega_2} (-N \oplus \alpha_2^* - \underline{c})_+ d\lambda_2 < \gamma \delta \leq \gamma \mu_1 \quad \lambda_1\text{-a.e. in } E,$$

which contradicts (23). Therefore, $\alpha_1^* > -N$ λ_1 -a.e. in Ω_1 , which even implies that $(\alpha_1^*)_+ \in L^\infty(\Omega_1)$. Concerning $(\alpha_2^*)_+$, one can argue in exactly the same way to conclude that $(\alpha_2^*)_+ \in L^\infty(\Omega_2)$, too.

For the positive parts we find

$$\begin{aligned} & |\Omega_2| \|\alpha_1^*\|_{L^2(\Omega_1)}^2 + |\Omega_1| \|\alpha_2^*\|_{L^2(\Omega_2)}^2 \\ &= \int_{\Omega} |\alpha_1^* \oplus \alpha_2^*|^2 d\lambda \quad \left(\text{since } \int_{\Omega_2} \alpha_2^* d\lambda_2 = 0 \right) \\ &= \int_{\Omega} (\alpha_1^* \oplus \alpha_2^*)_+^2 + (\alpha_1^* \oplus \alpha_2^*)_-^2 d\lambda \\ &\leq 2 \int_{\Omega} (\alpha_1^* \oplus \alpha_2^* - c)_+^2 + c_+^2 + (\alpha_1^*)_-^2 + (\alpha_2^*)_-^2 d\lambda < \infty, \end{aligned}$$

where we used (21) and the boundedness of the negative parts proven above. Note that the constant shift, potentially needed to ensure $\int_{\Omega_2} \alpha_2^* d\lambda_2 = 0$ has no effect on the equation in (21) due to the additive structure of \oplus .

We have thus shown that (α_1^*, α_2^*) is feasible for (D). Since (α_1^*, α_2^*) solves (D'), whose objective is the same as in (D), while its feasible set is larger, this implies that we have found a solution to (D). \square

We now show that, if π^* is of the form $\pi^* = \gamma^{-1}(\alpha_1^* \oplus \alpha_2^* - c)_+$ with two functions $\alpha_i^* \in L^2(\Omega_i)$, $i = 1, 2$, and has the marginals μ_1 and μ_2 , respectively, then it solves the necessary and sufficient optimality conditions of the primal problem (1) in form of the following variational inequality:

$$\pi^* \in \mathcal{F}, \quad \langle \gamma\pi^* + c, \pi - \pi^* \rangle_{L^2} \geq 0 \quad \forall \pi \in \mathcal{F}. \quad (\text{VI})$$

Herein, \mathcal{F} is the (convex) feasible set of (1), i.e.

$$\mathcal{F} := \left\{ \pi \in L^2(\Omega) : \pi \geq 0 \text{ } \lambda\text{-a.e. in } \Omega, \int_{\Omega_2} \pi d\lambda_2 = \mu_1 \text{ } \lambda_1\text{-a.e. in } \Omega_1, \int_{\Omega_1} \pi d\lambda_1 = \mu_2 \text{ } \lambda_2\text{-a.e. in } \Omega_2 \right\}.$$

For this purpose, let $\pi \in \mathcal{F}$ be fixed but arbitrary. Multiplying the equality constraints in \mathcal{F} with α_1^* and α_2^* , respectively, integrating the arising equations and add them yields

$$\begin{aligned} \int_{\Omega_1} \mu_1 \alpha_1^* d\lambda_1 + \int_{\Omega_2} \mu_2 \alpha_2^* d\lambda_2 &= \int_{\Omega} \pi (\alpha_1^* \oplus \alpha_2^*) d\lambda \\ &= \int_{\Omega} \pi [(\alpha_1^* \oplus \alpha_2^* - c)_+ + c] d\lambda - \int_{\Omega} \pi (\alpha_1^* \oplus \alpha_2^* - c)_- d\lambda \\ &\leq \int_{\Omega} \pi (\gamma\pi^* + c) d\lambda, \end{aligned} \quad (25)$$

where we used $\pi \geq 0$ for the last inequality. Using the feasibility of π^* , we find similarly

$$\begin{aligned} \int_{\Omega_1} \mu_1 \alpha_1^* d\lambda_1 + \int_{\Omega_2} \mu_2 \alpha_2^* d\lambda_2 &= \int_{\Omega} \pi^* [(\alpha_1^* \oplus \alpha_2^* - c) + c] d\lambda \\ &= \int_{\Omega} \gamma^{-1}(\alpha_1^* \oplus \alpha_2^* - c)_+ [(\alpha_1^* \oplus \alpha_2^* - c) + c] d\lambda \\ &= \int_{\Omega} \pi^* (\gamma\pi^* + c) d\lambda. \end{aligned} \quad (26)$$

Combining (25) and (26) now yields (VI). As (1) is a strictly convex minimization problem, this shows that, if π^* has the form $\pi^* = \gamma^{-1}(\alpha_1^* \oplus \alpha_2^* - c)_+$ with functions $\alpha_i^* \in L^2(\Omega_i)$ and satisfies $\pi^* \in \mathcal{F}$, then it is a solution of (1). On the other hand, we know from Theorem 2.10 that, under Assumption 1 (more or less needed for the existence of solutions of (1) anyway), there always exist $\alpha_i^* \in L^2(\Omega_i)$ so that $\pi^* = \gamma^{-1}(\alpha_1^* \oplus \alpha_2^* - c)_+$ satisfies the equality constraints in \mathcal{F} . Therefore, in summary we have deduced the following:

Theorem 2.11 (Necessary and Sufficient Optimality Conditions for (1)) *Under Assumption 1, $\pi^* \in L^2(\Omega)$ is a solution of (1) if and only if there exist functions $\alpha_i^* \in L^2(\Omega_i)$, $i = 1, 2$, such that the following optimality system is fulfilled:*

$$\pi^* - \frac{1}{\gamma} (\alpha_1^* \oplus \alpha_2^* - c)_+ = 0 \quad \lambda\text{-a.e. in } \Omega, \quad (27a)$$

$$\int_{\Omega_2} (\alpha_1^* \oplus \alpha_2^* - c)_+ d\lambda_2 = \gamma\mu_1 \quad \lambda_1\text{-a.e. in } \Omega_1, \quad (27b)$$

$$\int_{\Omega_1} (\alpha_1^* \oplus \alpha_2^* - c)_+ d\lambda_1 = \gamma\mu_2 \quad \lambda_2\text{-a.e. in } \Omega_2. \quad (27c)$$

The significance of Theorem 2.11 lies in the fact that we can characterize optimality of π by just two equalities in $L^2(\Omega_1)$ and $L^2(\Omega_2)$, respectively, namely (27b) and (27c). Thus, we effectively reduce the size of the problem from searching one function on $\Omega = \Omega_1 \times \Omega_2$ to searching two functions, one on Ω_1 and one on Ω_2 (similarly as for entropic regularization, cf. [4]). This will be exploited numerically in Section 3.

2.3 Regularization of the dual problem

As seen before, the dual problem in (D) is not uniquely solvable. One source of non-uniqueness is of course the kernel of the map $(\alpha_1, \alpha_2) \mapsto \alpha_1 \oplus \alpha_2$. This kernel is one-dimensional and is spanned by the function $(1, -1)$, which could be easily taken into account in an algorithmic framework. However, there is another source of non-uniqueness due to the max-operator that cuts off the negative part. Here is a simple example where dual solutions are not unique: For $\Omega_1 = \Omega_2 = [0, 1]$, $\mu_1 = \mu_2 \equiv 1$, and

$$c(x, y) := \begin{cases} C, & \text{if } \frac{1}{2} \leq x \leq 1, \frac{1}{2} \leq y \leq 1, \\ 0, & \text{else,} \end{cases} \quad \text{with } C > 4,$$

one can show by a straight forward calculation that, for every $\delta \in [0, \frac{C-4}{2}]$, the tuple

$$\alpha_1^*(x) = \begin{cases} 1 + \delta, & \text{if } x \in [0, \frac{1}{2}), \\ -1 - \delta, & \text{if } x \in [\frac{1}{2}, 1], \end{cases} \quad \alpha_2^*(y) = \begin{cases} 3 + \delta, & \text{if } y \in [0, \frac{1}{2}), \\ 1 - \delta, & \text{if } y \in [\frac{1}{2}, 1], \end{cases}$$

solves the optimality system (27b)–(27c). This shows that the potential structure of non-uniqueness might become fairly intricate. A situation like this can certainly happen in the discretized problem we will derive in Section 2.4 and can lead to problems when we derive algorithms for the discrete problem since non-unique solutions imply a degenerate Hessian at the optimum.

Therefore, we investigate the following regularization of the dual problem:

$$\left. \begin{array}{l} \min \quad \Phi_\varepsilon(\alpha_1, \alpha_2) := \Phi(\alpha_1, \alpha_2) + \frac{\varepsilon}{2} (\|\alpha_1\|_{L^2(\Omega_1)}^2 + \|\alpha_2\|_{L^2(\Omega_2)}^2) \\ \text{s.t.} \quad \alpha_i \in L^2(\Omega_i), i = 1, 2, \end{array} \right\} \quad (\text{D}_\varepsilon)$$

with a regularization parameter $\varepsilon > 0$. It is clear that the additional quadratic terms in the regularized objective Φ_ε yield that the latter is strictly convex and coercive in $L^2(\Omega_1) \times L^2(\Omega_2)$. Therefore, for every $\varepsilon > 0$, (D_ε) admits a unique solution.

Proposition 2.12 *Let $\{\varepsilon_n\} \subset \mathbb{R}^+$ be a sequence converging to zero and denote the solutions of (D_ε) with $\varepsilon = \varepsilon_n$ by $(\alpha_1^n, \alpha_2^n) \in L^2(\Omega_1) \times L^2(\Omega_2)$. Then the sequence $\{(\alpha_1^n, \alpha_2^n)\}$ admits a weak accumulation point, every weak accumulation point is also strong one and a solution of the original dual problem (D).*

Proof Let $(\alpha_1^*, \alpha_2^*) \in L^2(\Omega_1) \times L^2(\Omega_2)$ denote an arbitrary globally optimal solution of (D) (whose existence is guaranteed by Theorem 2.10). Then the optimality of (α_1^*, α_2^*) for (D) and of (α_1^n, α_2^n) for (D $_\varepsilon$) (with $\varepsilon = \varepsilon_n$) gives

$$\Phi(\alpha_1^*, \alpha_2^*) + \frac{\varepsilon_n}{2} (\|\alpha_1^n\|_{L^2(\Omega_1)}^2 + \|\alpha_2^n\|_{L^2(\Omega_2)}^2) \leq \Phi_{\varepsilon_n}(\alpha_1^n, \alpha_2^n) \leq \Phi_{\varepsilon_n}(\alpha_1^*, \alpha_2^*)$$

which implies

$$\|\alpha_1^n\|_{L^2(\Omega_1)}^2 + \|\alpha_2^n\|_{L^2(\Omega_2)}^2 \leq \|\alpha_1^*\|_{L^2(\Omega_1)}^2 + \|\alpha_2^*\|_{L^2(\Omega_2)}^2. \quad (28)$$

Thus, the boundedness of $\{(\alpha_1^n, \alpha_2^n)\}$ in $L^2(\Omega_1) \times L^2(\Omega_2)$. This in turn gives the existence of a weak accumulation point as claimed.

Now assume that $(\tilde{\alpha}_1, \tilde{\alpha}_2)$ is such a weak accumulation point, i.e.

$$(\alpha_1^n, \alpha_2^n) \rightharpoonup (\tilde{\alpha}_1, \tilde{\alpha}_2) \quad \text{in } L^2(\Omega_1) \times L^2(\Omega_2) \quad (29)$$

(for a subsequence). Using again the optimality of (α_1^*, α_2^*) and (α_1^n, α_2^n) , respectively, we obtain

$$\Phi(\alpha_1^*, \alpha_2^*) \leq \Phi(\alpha_1^n, \alpha_2^n) \leq \Phi_{\varepsilon_n}(\alpha_1^n, \alpha_2^n) \leq \Phi_{\varepsilon_n}(\alpha_1^*, \alpha_2^*) \rightarrow \Phi(\alpha_1^*, \alpha_2^*). \quad (30)$$

On the other hand, by convexity and weak lower semicontinuity of Φ we get from (29) and (30) that

$$\Phi(\tilde{\alpha}_1, \tilde{\alpha}_2) \leq \liminf_{n \rightarrow \infty} \Phi(\alpha_1^n, \alpha_2^n) = \lim_{n \rightarrow \infty} \Phi(\alpha_1^n, \alpha_2^n) = \Phi(\alpha_1^*, \alpha_2^*),$$

which gives in turn the optimality of the weak limit. Estimate (28) for the choice $(\alpha_1^*, \alpha_2^*) = (\tilde{\alpha}_1, \tilde{\alpha}_2)$ shows that

$$\|\alpha_1^n\|_{L^2(\Omega_1)}^2 + \|\alpha_2^n\|_{L^2(\Omega_2)}^2 \leq \|\tilde{\alpha}_1\|_{L^2(\Omega_1)}^2 + \|\tilde{\alpha}_2\|_{L^2(\Omega_2)}^2$$

and thus, we have

$$\liminf_{n \rightarrow \infty} \|\alpha_1^n\|_{L^2(\Omega_1)}^2 + \|\alpha_2^n\|_{L^2(\Omega_2)}^2 \leq \|\tilde{\alpha}_1\|_{L^2(\Omega_1)}^2 + \|\tilde{\alpha}_2\|_{L^2(\Omega_2)}^2,$$

but $(\alpha_1^n, \alpha_2^n) \rightharpoonup (\tilde{\alpha}_1, \tilde{\alpha}_2)$ would imply

$$\|\tilde{\alpha}_1\|_{L^2(\Omega_1)}^2 + \|\tilde{\alpha}_2\|_{L^2(\Omega_2)}^2 < \liminf_{n \rightarrow \infty} \|\alpha_1^n\|_{L^2(\Omega_1)}^2 + \|\alpha_2^n\|_{L^2(\Omega_2)}^2$$

and consequently, we have $(\alpha_1^n, \alpha_2^n) \rightarrow (\tilde{\alpha}_1, \tilde{\alpha}_2)$ in $L^2(\Omega_1) \times L^2(\Omega_2)$. \square

Theorem 2.13 *Let $\{\varepsilon_n\} \subset \mathbb{R}^+$ be a sequence converging to zero and denote the solutions of (D $_\varepsilon$) with $\varepsilon = \varepsilon_n$ again by $(\alpha_1^n, \alpha_2^n) \in L^2(\Omega_1) \times L^2(\Omega_2)$. Moreover, define*

$$\pi_n := \frac{1}{\gamma} (\alpha_1^n \oplus \alpha_2^n - c)_+. \quad (31)$$

Then π_n converges strongly in $L^2(\Omega)$ to the unique solution of (1).

Proof From (28), we know that $\{(\alpha_1^n, \alpha_2^n)\}$ is bounded and hence, $\{\pi_n\}$ is bounded in $L^2(\Omega)$. Thus,

$$\pi_n \rightharpoonup \tilde{\pi} \quad \text{in } L^2(\Omega) \quad (32)$$

for some subsequence. Now we show that $\tilde{\pi}$ is the optimal for (1). Weak closedness of $\{\pi \in L^2(\Omega) : \pi(x_1, x_2) \geq 0 \text{ a.e. in } \Omega\}$ implies $\tilde{\pi} \geq 0$. Integrating the first-order optimality conditions for (D_ε)

$$\int_{\Omega_2} (\alpha_1^n \oplus \alpha_2^n - c)_+ d\lambda_2 + \varepsilon \alpha_1^n = \gamma \mu_1 \quad \lambda_1\text{-a.e. in } \Omega_1 \quad (33)$$

$$\int_{\Omega_1} (\alpha_1^n \oplus \alpha_2^n - c)_+ d\lambda_1 + \varepsilon \alpha_2^n = \gamma \mu_2 \quad \lambda_2\text{-a.e. in } \Omega_2. \quad (34)$$

against some $\varphi_1 \in C_c^\infty(\Omega_1)$, inserting the definition of π_n , and integrating over Ω_1 yields

$$\int_{\Omega_1} \int_{\Omega_2} \pi_n d\lambda_2 \varphi_1 d\lambda_1 = \int_{\Omega_1} \mu_1 \varphi_1 d\lambda_1 - \frac{\varepsilon_n}{\gamma} \int_{\Omega_1} \alpha_1^n \varphi_1 d\lambda_1$$

Passing to the limit we obtain

$$\int_{\Omega_1} \int_{\Omega_2} \tilde{\pi} d\lambda_2 \varphi_1 d\lambda_1 = \int_{\Omega_1} \mu_1 \varphi_1 d\lambda_1,$$

and thus, $\tilde{\pi}$ satisfies the first equality constraint in (1). The second equality constraint can be verified analogously.

To show optimality of $\tilde{\pi}$, we test the optimality conditions in (33) and (34) with α_1^n and α_2^n , respectively, and get

$$\begin{aligned} \Phi_{\varepsilon_n}(\alpha_1^n, \alpha_2^n) &= \frac{\gamma^2}{2} \|\pi_n\|_{L^2(\Omega)}^2 - \gamma \int_{\Omega} \pi_n (\alpha_1^n \oplus \alpha_2^n) d\lambda - \frac{\varepsilon_n}{2} \|\alpha_1^n\|_{L^2(\Omega_1)}^2 - \frac{\varepsilon_n}{2} \|\alpha_2^n\|_{L^2(\Omega_2)}^2 \\ &= -\frac{\gamma^2}{2} \|\pi_n\|_{L^2(\Omega)}^2 - \gamma \int_{\Omega} c \pi_n d\lambda - \frac{\varepsilon_n}{2} \|\alpha_1^n\|_{L^2(\Omega_1)}^2 - \frac{\varepsilon_n}{2} \|\alpha_2^n\|_{L^2(\Omega_2)}^2 \\ &= -\gamma E_\gamma(\pi_n) - \frac{\varepsilon_n}{2} \|\alpha_1^n\|_{L^2(\Omega_1)}^2 - \frac{\varepsilon_n}{2} \|\alpha_2^n\|_{L^2(\Omega_2)}^2, \end{aligned}$$

where E_γ is the primal objective from (3). Similarly, we get

$$\Phi(\alpha_1^*, \alpha_2^*) = -\gamma E_\gamma(\pi^*),$$

where $\pi^* \in L^2(\Omega)$ is the unique solution of (1) and $(\alpha_1^*, \alpha_2^*) \in L^2(\Omega_1) \times L^2(\Omega_2)$ solves the dual problem (D). Now, putting everything so far together, we obtain

$$\begin{aligned} \lim_{n \rightarrow \infty} E_\gamma(\pi_n) &= \lim_{n \rightarrow \infty} \left(-\frac{1}{\gamma} \Phi_{\varepsilon_n}(\alpha_1^n, \alpha_2^n) - \frac{\varepsilon_n}{2\gamma} \|\alpha_1^n\|_{L^2(\Omega_1)}^2 - \frac{\varepsilon_n}{2\gamma} \|\alpha_2^n\|_{L^2(\Omega_2)}^2 \right) \\ &= -\frac{1}{\gamma} \Phi(\alpha_1^*, \alpha_2^*) = E_\gamma(\pi^*). \end{aligned}$$

On the other hand, E_γ is weakly lower semicontinuous, and therefore

$$E_\gamma(\tilde{\pi}) \leq \liminf_{n \rightarrow \infty} E_\gamma(\pi_n) = E_\gamma(\pi^*).$$

This gives the optimality of $\tilde{\pi}$ and by strict convexity also uniqueness, i.e. $\tilde{\pi} = \pi^*$. Thus, the weak limit is unique and a well known argument by contradiction therefore implies the weak convergence of the whole sequence $\{\pi_n\}$ to π^* . Finally, strong convergence follows from a standard argument. \square

2.4 The discrete dual problem

We show a simple discretization of the quadratically regularized optimal transport problem (1) by piecewise constant approximation in Appendix A. To keep the notation concise, we state the corresponding discrete optimal transport problem and illustrate the duality already here. This will be the basis of our algorithms we derive in Section 3. A discrete version of the continuous problem (1) is the finite-dimensional problem

$$\min_{\pi \in \mathbb{R}^{M \times N}} \langle \pi, c \rangle + \frac{\gamma}{2} \|\pi\|_F^2 \quad \text{s.t.} \quad \pi^T \mathbf{1}_M = \mu, \quad \pi \mathbf{1}_N = \nu, \quad \pi \geq 0 \quad (35)$$

where $\mathbf{1}_N \in \mathbb{R}^N$ denotes the vector of all ones, $\mu \in \mathbb{R}^N$, $\nu \in \mathbb{R}^M$ denote the discretized marginals with $\sum_{j=1}^N \mu_j = \sum_{i=1}^M \nu_i$, and $c \in \mathbb{R}^{M \times N}$ denotes the discretized cost. Note that we slightly changed the notation from μ_1 and μ_2 to μ and ν , respectively. For the discrete form of the optimality system (27) we further replace the Lagrange multipliers α_1 and α_2 by α and β , respectively, and get

$$\pi = \frac{1}{\gamma} (\alpha \oplus \beta - c)_+ \quad (36a)$$

$$\sum_{i=1}^M (\alpha_i + \beta_j - c_{ij})_+ = \gamma \mu_j, \quad j = 1, \dots, N \quad (36b)$$

$$\sum_{j=1}^N (\alpha_i + \beta_j - c_{ij})_+ = \gamma \nu_i, \quad i = 1, \dots, M \quad (36c)$$

where $\alpha \in \mathbb{R}^M$, $\beta \in \mathbb{R}^N$ and $(\alpha \oplus \beta)_{i,j} = \alpha_i + \beta_j$ is the ‘‘outer sum’’. The discrete counterpart of Φ from (D) is

$$\Phi(\alpha, \beta) = \frac{1}{2} \|(\alpha \oplus \beta - c)_+\|_F^2 - \gamma \langle \nu, \alpha \rangle - \gamma \langle \mu, \beta \rangle$$

where $\|\cdot\|_F$ denotes the Frobenius norm.

We write the optimality condition (36b)-(36c) as a non-smooth equation $F(\alpha, \beta) = 0$ in \mathbb{R}^{M+N} with

$$F(\alpha, \beta) = \begin{pmatrix} F_1(\alpha, \beta) \\ F_2(\alpha, \beta) \end{pmatrix} = \begin{pmatrix} (\sum_{j=1}^N (\alpha_i + \beta_j - c_{ij})_+ - \gamma \nu_i)_{i=1, \dots, M} \\ (\sum_{i=1}^M (\alpha_i + \beta_j - c_{ij})_+ - \gamma \mu_j)_{j=1, \dots, N} \end{pmatrix} \quad (37)$$

(note that $F_1 = \partial_\alpha \Phi$ and $F_2 = \partial_\beta \Phi$). Since F is the composition of Lipschitz continuous and semismooth functions, we have the following result (for the chain rule for semismooth functions, see e.g. [14, Thm. 2.10]):

Lemma 2.14 *The function F (and thus, the gradient of Φ) is (globally) Lipschitz continuous and semismooth.*

3 Algorithms

The optimality system (36b), (36c) for the smooth and convex problem (D) can be solved by different methods. In [3] the authors propose to use a generic L-BFGS solver and also derive an alternating minimization scheme, which is similar to the non-linear Gauss-Seidel method in the next section, but differs slightly in the numerical realization and [20] also uses an off-the-shelf solver. Here we propose methods that exploit the special structure of the optimality system: A non-linear Gauss-Seidel method and a semismooth Newton method.

3.1 Non-linear Gauss-Seidel

The method in this section is similar to the one described in the Appendix of [3], but we describe it here for the sake of completeness. A close look at the optimality system

$$\sum_{j=1}^N (\alpha_i + \beta_j - c_{ij})_+ = \gamma \nu_i, \quad i = 1, \dots, M. \quad (38a)$$

$$\sum_{i=1}^M (\alpha_i + \beta_j - c_{ij})_+ = \gamma \mu_j, \quad j = 1, \dots, N \quad (38b)$$

shows that we can solve all M equations in (38a) for the α_i in parallel (for fixed β) since the i th equation depends on α_i only. Similarly, all N equations in (38b) can be solved for the β_j if α is fixed. Hence, we can perform a non-linear Gauss-Seidel method for these non-smooth equations (also known as alternating minimization, nonlinear SOR or coordinate descent method for Φ [6, 25]), i.e. alternately solving the equations (38a) for α (for fixed β) and then the equations (38b) for β (for fixed α). The whole method is stated in Algorithm 1. Since Φ is convex with Lipschitz continuous gradient (cf. Lemma 2.14) the convergence of the algorithm follows from results in [2].

Algorithm 1 Non-linear Gauss-Seidel for quadratically regularized optimal transport

Initialize: $\beta^0 \in \mathbb{R}^N$, set $k = 0$

repeat

 Set α^{k+1} to be the solution of (38a) with $\beta = \beta^k$.

 Set β^{k+1} to be the solution of (38b) with $\alpha = \alpha^{k+1}$.

$k \leftarrow k + 1$

until some stopping criterion

Each equation for an α_i or β_j is just a single scalar equation for a scalar quantity and the structure of the equation is of the following form: For a given vector $y \in \mathbb{R}^n$ and right hand side $b \in \mathbb{R}$, solve

$$f(x) := \sum_{j=1}^n (x - y_j)_+ = b. \quad (39)$$

Of course, one can solve this problem by bisection, but here are two other, more efficient methods to solve equations of the type (39):

Direct search. If we denote by $y_{[j]}$ the j -th smallest entry of y (i.e. we sort y in an ascending way), we get that

$$f(x) = \sum_{j=1}^n (x - y_{[j]})_+ = \begin{cases} 0, & x \leq y_{[1]} \\ kx - \sum_{j=1}^k y_{[j]}, & y_{[k]} \leq x \leq y_{[k+1]}, \quad k = 1, \dots, n-1 \\ nx - \sum_{j=1}^n y_{[j]}, & x \geq y_{[n]}. \end{cases}$$

To obtain the solution of (39) we evaluate f at the break points $y_{[j]}$ until we find the interval $]y_{[k]}, y_{[k+1}[$ in which the solution lies (by finding k such that $f(y_{[k]}) \leq b < f(y_{[k+1]})$), and then setting

$$x = \frac{b + \sum_{j=1}^k y_{[j]}}{k}.$$

The complexity of the method is dominated by the sorting of the vector y , its complexity is $\mathcal{O}(n \log(n))$.

Semismooth Newton. Although f is non-smooth, we may perform Newton's method here. The function f is piecewise linear and on each interval $]y_{[j]}, y_{[j+1}[$ is has the slope j (a simple situation with $n = 3$ is shown in Figure 1). At the break points we may define $f'(y_{[j]}) = j$ and then we iterate

$$x^{k+1} = x^k - \frac{f(x^k)}{f'(x^k)}.$$

If we start with $x^0 \geq y^{[n]} = \max_k y_k$, the method will produce a monotonically decreasing sequence which converges in at most n steps. Actually, we can initialize the method with any x^0 that is strictly larger than $y_{[1]} = \min_k y_k$. Note that we do not need to sort the values of y_k to calculate the derivative since we have $f'(x) = \#\{i : x \geq y_i\}$. In practice, the method usually needs much less iterations than n .

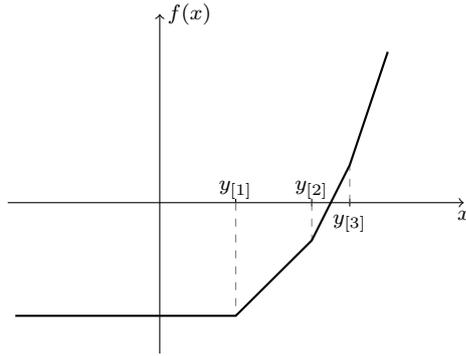


Fig. 1 Illustration of the non-smooth function f from (39).

3.2 Semismooth Newton

As seen in Lemma 2.14, the mapping F is semismooth and hence, we may use a semismooth Newton method [5, 7].

A simple calculation proves the following lemma.

Lemma 3.1 *A Newton derivative of F from (37) at (α, β) is given by*

$$G = \begin{pmatrix} \text{diag}(\sigma \mathbf{1}_N) & \sigma \\ \sigma^T & \text{diag}(\sigma^T \mathbf{1}_M) \end{pmatrix} \in \mathbb{R}^{(M+N) \times (M+N)}$$

where $\sigma \in \mathbb{R}^{M \times N}$ is given by

$$\sigma_{ij} = \begin{cases} 1 & \alpha_i + \beta_j - c_{ij} \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

A step of the semismooth Newton method for the solution of $F(\alpha, \beta) = 0$ would consist of setting

$$\begin{pmatrix} \alpha^{k+1} \\ \beta^{k+1} \end{pmatrix} = \begin{pmatrix} \alpha^k \\ \beta^k \end{pmatrix} - \begin{pmatrix} \delta\alpha^k \\ \delta\beta^k \end{pmatrix} \quad \text{where} \quad F(\alpha^k, \beta^k) = G \begin{pmatrix} \delta\alpha^k \\ \delta\beta^k \end{pmatrix}.$$

However, the next lemma shows, that G has a non-trivial kernel.

Lemma 3.2 *Let G be the Newton derivative of F at (α, β) defined in Lemma 3.1. Then the following holds true:*

1. $G \in \mathbb{R}^{(M+N) \times (M+N)}$ is symmetric,
2. G is positive semi-definite,
3. $(a, b) \in \text{kern}(G)$ if and only if $\sigma_{ij}(a_i + b_j) = 0$ for all $1 \leq i \leq M$, $1 \leq j \leq N$.

Proof Symmetry of G is clear by construction. To see that G is positive semi-definite we calculate

$$\begin{aligned} (a, b)^\top G(a, b) &= \sum_{j=1}^N \sum_{i=1}^M \sigma_{ij} a_i^2 + \sum_{j=1}^N \sum_{i=1}^M \sigma_{ij} b_j^2 + 2 \sum_{j=1}^N \sum_{i=1}^M \sigma_{ij} a_i b_j \\ &= \sum_{j=1}^N \sum_{i=1}^M \sigma_{ij} (a_i + b_j)^2 \geq 0. \end{aligned}$$

Due to the non-negativity of σ , this also shows the last point. \square

The third point of the lemma shows that the kernel of G may have a high dimension, depending on the matrix σ . Hence we resort to a quasi Newton method where we regularize the Newton step arising from the dual problem from Section 2.2 by setting

$$\begin{pmatrix} \alpha^{k+1} \\ \beta^{k+1} \end{pmatrix} = \begin{pmatrix} \alpha^k \\ \beta^k \end{pmatrix} - \begin{pmatrix} \delta\alpha^k \\ \delta\beta^k \end{pmatrix} \quad \text{where} \quad F(\alpha^k, \beta^k) = (G + \varepsilon I) \begin{pmatrix} \delta\alpha^k \\ \delta\beta^k \end{pmatrix}$$

with a small $\varepsilon > 0$. By [5], the method still converges, but only a local linear rate is guaranteed. We note that we have not applied the semismooth Newton method to the regularized dual problem from Section 2.3. This would also be possible, but lead not only to the regularized Newton matrix from above but we would also have to adapt the objective F in the computation of the update.

Let us make a few remarks on the the regularized Newton step and its numerical treatment.

- The matrix σ (and hence the Newton matrix G) is usually very sparse. The closer α and β are to the optimal ones, the closer $(\alpha_i + \beta_j - c_{ij})_+$ is to the optimal regularized transport plan π and for small γ this usually very sparse.
- Since G is positive semi-definite, the regularized step could be done by the method of conjugate gradients. However, any linear solver that can exploit the sparsity of G can be used.

As usual, the regularized semismooth Newton method may not converge globally. A simple globalization technique is an Armijo linesearch in the Newton direction. The full method is described in Algorithm 2.

Algorithm 2 Globalized and regularized semismooth Newton method quadratically regularized optimal transport

Initialize: $\alpha^0 \in \mathbb{R}^M$, $\beta^0 \in \mathbb{R}^N$, set $k = 0$, choose regularization parameter $\varepsilon > 0$, Armijo parameters $\theta, \kappa \in]0, 1[$, and a tolerance $\tau > 0$

repeat

 Calculate

$$P_{ij} = \alpha_i^k + \beta_j^k - c_{ij}, \quad \sigma_{ij} = \begin{cases} 1 & P_{ij} \geq 0 \\ 0 & \text{otherwise} \end{cases}, \quad \text{and} \quad \pi_{ij} = \max(P_{ij}, 0)/\gamma.$$

 Calculate $\delta\alpha$ and $\delta\beta$ by solving

$$\left(\begin{pmatrix} \text{diag}(\sigma \mathbf{1}_N) & \sigma \\ \sigma^T & \text{diag}(\sigma^T \mathbf{1}_M) \end{pmatrix} + \varepsilon I \right) \begin{pmatrix} \delta\alpha \\ \delta\beta \end{pmatrix} = -\gamma \begin{pmatrix} \pi \mathbf{1}_N - \nu \\ \pi^T \mathbf{1}_M - \mu \end{pmatrix}$$

 Set $t = 1$ and compute the directional derivative

$$d = D_{(\delta\alpha, \delta\beta)} \Phi(\alpha^k, \beta^k) = \gamma \sum_{ij} \pi_{ij} (\delta\alpha_i + \delta\beta_j) - \gamma (\langle \delta\alpha, \nu \rangle + \langle \delta\beta, \mu \rangle).$$

while $\Phi(\alpha^k + t\delta\alpha, \beta^k + t\delta\beta) \geq \Phi(\alpha^k, \beta^k) + t\theta d$ **do**

$t \leftarrow \kappa t$

end while

 Set $\alpha^{k+1} = \alpha^k - t\delta\alpha$, $\beta^{k+1} = \beta^k - t\delta\beta$

$k \leftarrow k + 1$

until $\|\pi \mathbf{1}_N - \nu\|_\infty, \|\pi^T \mathbf{1}_M - \mu\|_\infty \leq \tau$

4 Numerical examples

4.1 Illustration of $\gamma \rightarrow 0$

In our first numerical example we illustrate the how the solutions π^* of the regularized problem converge for vanishing regularization parameter $\gamma \rightarrow 0$. We generate some marginals, fix a transport cost and compute solutions of the discretized transport problems (35) for a sequence $\gamma_n \rightarrow 0$ and illustrate the optimal transport plans (and the related regularized transport costs). Our marginals are non-negative functions sampled at equidistant points x_i, y_i in the interval $[0, 1]$ and we used $M = N = 400$ and the cost $c_{ij} = (x_i - y_j)^2$ is the squared distance between the sampling points. The results are shown in Figure 2. One observes that the optimal transport plans converge to a measure that is singular and is supported on the graph of a monotonically increasing function, exactly as the fundamental theorem of optimal transport [1] predicts.

We repeat the same experiment where the cost is the (non-squared) distance $c_{ij} = |x_i - y_j|$. Here we had to choose larger regularization parameters as it turned out that values similar to Figure 2 would lead to almost undistinguishable results. The results are shown in Figure 3. Note the different structure of the transport plan (which is again in agreement with the predicted results from the fundamental theorem of optimal transport). In Figure 4 we show the results for the concave but increasing cost $c_{ij} = \sqrt{|x_i - y_j|}$ and again observe the expected effect that a concave transport cost encouraged that as much mass as possible stays in place (as can be seen by the concentration of mass along the diagonal of the transport plan).

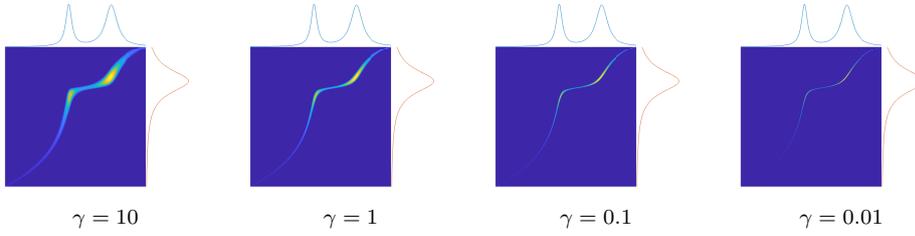


Fig. 2 Visualization of transport plans of the quadratically regularized optimal transport problem with $M = N = 400$ and quadratic transport cost $c_{ij} = (x_i - y_j)^2$.

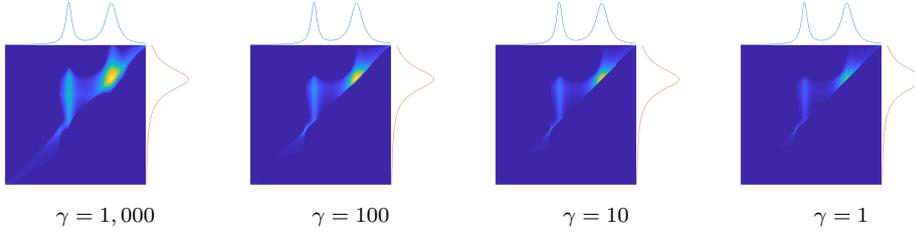


Fig. 3 Visualization of transport plans of the quadratically regularized optimal transport problem with $M = N = 400$ and metric transport cost $c_{ij} = |x_i - y_j|$.

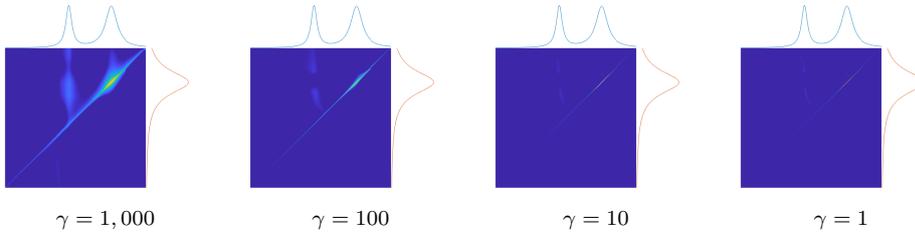


Fig. 4 Visualization of transport plans of the quadratically regularized optimal transport problem with $M = N = 400$ and concave increasing transport cost $c_{ij} = \sqrt{|x_i - y_j|}$.

4.2 Mesh independence and comparison of SSN and NLGS

While we did not analyze our algorithms in the continuous case, we made an experiment to see how the methods converge when we change the mesh size of the discretization. To that end, we did a simple piecewise constant approximation of the marginals, the cost and the transport plan as described in Appendix A. This derivation shows that one has to scale up the marginals for finer discretization (or, equivalently, scale down the regularization parameter γ) to get consistent results. We also took care to adapt the termination criteria so that we terminate the algorithms when the continuous counterpart of the termination criteria is satisfied (again, see Table 1 in Appendix A for details).

We used marginals $\mu^\pm : [0, 1] \rightarrow [0, \infty[$ of the form

$$\mu(x) = r \frac{1}{1+m(x-a)^2}, \quad \nu(x) = s \left(\frac{1}{1+m_1(x-a_1)^2} + \frac{1}{1+m_2(x-a_2)^2} \right)$$

with varying $m, m_1, m_2 > 0$, $0 < a, a_1, a_2 < 1$ and appropriate normalization factors r, s and quadratic cost $c(x, y) = (x - y)^2$ and discretized each instance of the problem with $M = N$ varying from 10 to 1,000. We solved the problem for each size for regularization parameter $\gamma = 0.001$ with the semismooth Newton method from Algorithm 2 (with parameters $\epsilon = 10^{-6}$ and Armijo parameters $\kappa = 0.5$ and $\theta = 0.1$) up to tolerance 10^{-3} and report the number of iterations needed in Figure 5. As can be observed, the number of iterations is comparable for each instance of the problem. Moreover, it seems that the number of iterations does not grow with finer discretization (however, the number of iterations seems to oscillate unpredictable for coarse discretization). This would hint at mesh independence of the method and one could hope to prove this in future research. We performed a similar experiment for the nonlinear Gauss-Seidel method from Algorithm 1 (with larger regularization parameter $\gamma = 0.05$ and only up to $M = N = 500$ and show the results in Figure 6. We see an overall increase of the number of iterations but only very slightly (with several instances where the number of iterations does not increase with finer discretization).

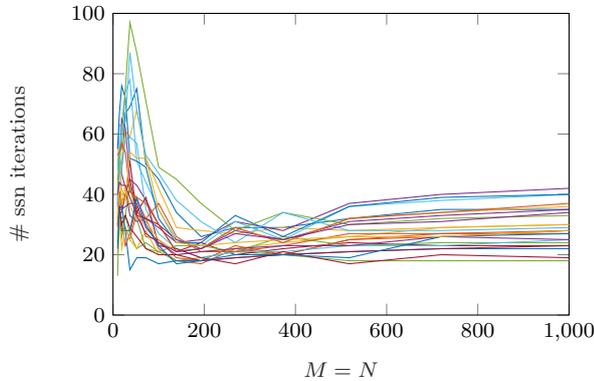


Fig. 5 Number of iteration for the semismooth Newton method to achieve a desired accuracy. Each graph corresponds to one instance of the problem.

4.3 Optimal transport between empirical distributions

As an example in two space dimensions, we consider two distributions μ, ν . Instead of using these as marginals, we consider empirical distributions, i.e. we generate samples $(x_i)_{i=1, \dots, N}$, sampled from μ and $(y_j)_{j=1, \dots, M}$, sampled from ν . These samples give empirical approximations

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^N \delta_{x_i}, \quad \hat{\nu} = \frac{1}{M} \sum_{j=1}^M \delta_{y_j}.$$

The optimal transport problem (1) with these two marginals does not fulfill Assumption 1, since the marginals are not L^2 -functions. However, we can consider it as a discrete problem optimal transport problem in the form (35) when we denote

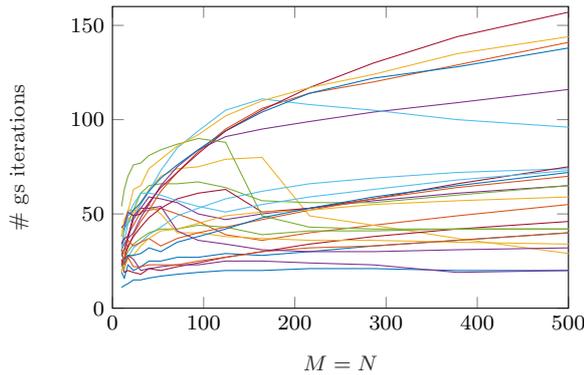


Fig. 6 Number of iteration for the nonlinear Gauss-Seidel semismooth method to achieve a desired accuracy. Each graph corresponds to one instance of the problem.

$c_{ij} = c(x_i, y_j)$ (for some cost c) and marginals $\mathbf{1}_M$ and $\mathbf{1}_N$, respectively. We solve this discrete optimal transport problem and obtain a transport plan π^* . Since we use quadratic regularization, the plan will be sparse and hence, we can visualize it by plotting arrows from x_i to y_j and we make the thickness of the arrows proportional to the size of the entry π_{ij}^* . In other words: The thickness of the arrow from x_i to y_j indicates how much of the mass in x_i has been transported to y_j . In Figure 7 we show the result for $N = 80$ samples from an anisotropic Gaussian distribution (centered at the origin) and $M = 120$ samples from a uniform distribution on a segment of an annulus. We used $c(x_i, y_j) = \|x_i - y_j\|^2$ with the Euclidean norm and regularization parameter $\gamma = 1$. The resulting plan π^* has 212 non-zero entries. For a comparison we show the result of entropically regularized optimal transport in the same situation in Figure 8. We used $\gamma = 0.05$ (which is the smallest value for which our naive implementation of Sinkhorn algorithms is still stable). The resulting plan has 6730 nonzero entries and we only plot lines for the transport which are larger than 1% of the largest entry in the optimal transport plan.

5 Conclusion

We analyzed the quadratically regularized optimal transport problem in Kantorovich form. While it is straight forward to derive the dual problem, our proof of existence of dual optima is quite intricate. We note that we are not aware of any proof of existence of the dual of other regularized transport problems in the continuous case besides the very recent [8] for entropic regularization. We derived two algorithms to solve the dual problems, both of which converge by standard results. It turns out that the semismooth quasi-Newton methods converges fast in all cases and that it behaves stably with respect to the regularization parameter in our numerical experiments. We even observe mesh independence of the method in the experiments. One drawback of the semismooth Newton method is (compared with, e.g. the Sinkhorn iteration [9]), is that we need to assemble the Newton matrix in each step. While this matrix is usually very sparse, one still needs to check MN cases, which may be too large for large scale problems. We did not

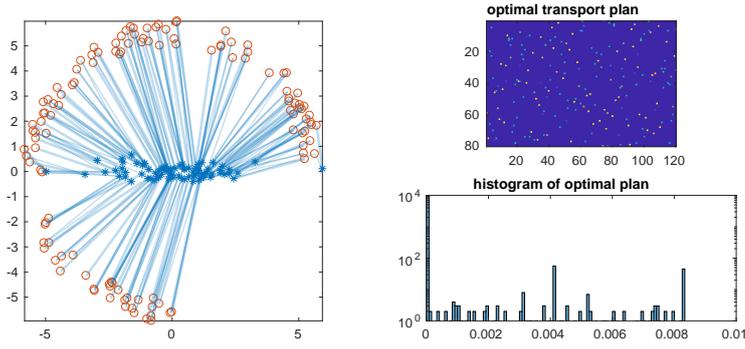


Fig. 7 Illustration of the quadratically regularized optimal transport between empirical distributions. Left: Source distribution $\hat{\mu}$ denoted by blue stars and target distribution $\hat{\nu}$ denoted by red circles together with lines that indicate the transport. Right: The transport plan and its histogram in semi-log scale.

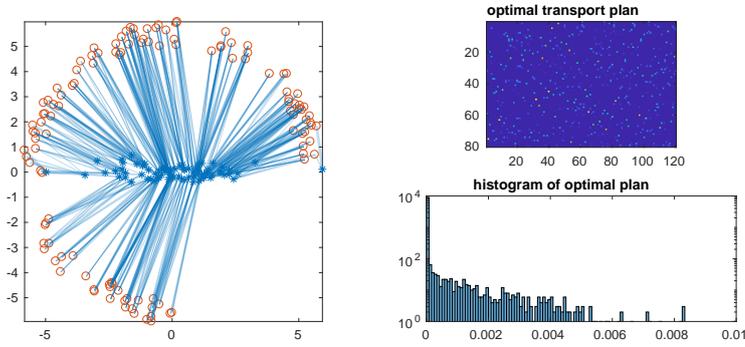


Fig. 8 Illustration of the entropically regularized optimal transport between empirical distributions. Left: Source distribution $\hat{\mu}$ denoted by blue stars and target distribution $\hat{\nu}$ denoted by red circles together with lines that indicate the transport. Right: The transport plan and its histogram in semi-log scale.

investigate, how special structure of the cost function c may help to reduce the cost to assemble the sparse matrix σ .

Acknowledgements We would like to thank the reviewer for helpful suggestions that lead to an improvement presentation and also Stephan Walther (TU Dortmund) for helping with the construction of the counterexample in Section 2.3.

A Discretization with piecewise-constant ansatz functions

For sake of brevity, we just consider an equidistant discretization of $[0, 1]$ into N intervals using piecewise constant ansatz functions, i.e.

$$\pi(x, y) := \sum_{i,j=0}^{N-1} \pi_{ij} \chi_{(\frac{i}{N}, \frac{i+1}{N}) \times (\frac{j}{N}, \frac{j+1}{N})}(x, y),$$

for coefficients π_{ij} and assume analogous definitions for the quantities c , μ^+ , μ^- , α and β . They have to coincide on average over the intervals. Again, we study this for π and obtain that the identity

$$\begin{aligned} \int_{\frac{i}{N}}^{\frac{i+1}{N}} \int_{\frac{j}{N}}^{\frac{j+1}{N}} \pi(x, y) \, dy dx &= \int_{\frac{i}{N}}^{\frac{i+1}{N}} \int_{\frac{j}{N}}^{\frac{j+1}{N}} \sum_{i,j=0}^{N-1} \pi_{ij} \chi_{(\frac{i}{N}, \frac{i+1}{N}) \times (\frac{j}{N}, \frac{j+1}{N})}(x, y) \, dy dx \\ &= \frac{1}{N^2} \pi_{ij} \end{aligned}$$

holds. Again, analogous identities hold for the quantities c , μ^+ , μ^- , α and β . The ones with one-dimensional domain are scaled by $\frac{1}{N}$ instead of $\frac{1}{N^2}$.

Now, we consider the discrete Algorithm 2, which operates on discrete quantities and establish a consistent mapping of the quantities from the discretization to the ones of the solver. We denote its input quantities by \bar{c}_{ij} , $\bar{\mu}_i^-$, $\bar{\mu}_i^+$ and its output quantities by $\bar{\alpha}_i$, $\bar{\beta}_j$, \bar{p}_{ij} , and \bar{E} . It solves for

$$\sum_{j=0}^{N-1} \bar{\pi}_{ij} = \gamma \bar{\mu}_i^+,$$

which we desire to correspond to

$$\int_{\frac{i}{N}}^{\frac{i+1}{N}} \int_0^1 \pi(x, y) \, dy dx = \int_{\frac{i}{N}}^{\frac{i+1}{N}} \mu^+(x) \, dx.$$

We plug in the ansatz functions and obtain the identity

$$\frac{1}{N^2} \sum_{j=0}^{N-1} \pi_{ij} = \frac{1}{N} \mu_i^+.$$

We set $\bar{\pi}_{ij} := \gamma \pi_{ij}$ and obtain

$$\sum_{j=0}^{N-1} \bar{\pi}_{ij} = \sum_{j=0}^{N-1} \pi_{ij} = N \mu_i^-.$$

Thus, the choice $\bar{\mu}_i^- := N \mu_i^-$ gives a consistent conversion. Similarly, we obtain $\bar{\mu}_j^+ := N \mu_j^+$. We proceed with the objective. Plugging in the ansatz functions into the continuous objective gives

$$E = \frac{\gamma}{2} \frac{1}{N^2} \sum_{i,j=0}^{N-1} \pi_{ij}^2 - \frac{1}{N} \left(\sum_{i=0}^{N-1} \alpha_i \mu_i^- + \sum_{j=0}^{N-1} \beta_j \mu_j^+ \right).$$

The solver computes

$$\bar{E} = \frac{1}{2} \sum_{i,j=0}^{N-1} \bar{\pi}_{ij}^2 - \gamma \left(\sum_{i=0}^{N-1} \bar{\alpha}_i \bar{\mu}_i^- + \sum_{j=0}^{N-1} \bar{\beta}_j \bar{\mu}_j^+ \right),$$

Plugging in $N \mu_i^- = \bar{\mu}_i^-$, $N \mu_j^+ = \bar{\mu}_j^+$ and $\gamma \pi_{ij} = \bar{\pi}_{ij}$ gives

$$\bar{E} = \gamma^2 \frac{1}{2} \sum_{i,j=0}^{N-1} \pi_{ij}^2 - \gamma N \left(\sum_{i=0}^{N-1} \bar{\alpha}_i \mu_i^- + \sum_{j=0}^{N-1} \bar{\beta}_j \mu_j^+ \right).$$

Thus, the consistent identity $E = \frac{1}{\gamma N^2} \bar{E}$ follows if we choose $\bar{\alpha}_i := \alpha_i$ and $\bar{\beta}_i := \beta_i$. The solver computes $\bar{\alpha}_i$ as the solution of

$$\sum_{j=0}^{N-1} (\bar{\alpha}_i + \bar{\beta}_j - \bar{c}_{ij})_+ = \gamma \bar{\mu}_i^-,$$

whereas the discretization of the corresponding continuous equation reads

$$\frac{1}{N} \sum_{j=0}^{N-1} (\alpha_i + \beta_j - c_{ij})_+ = \gamma \mu_i^-$$

in terms of the coefficients. Plugging in the choices $\alpha_i = \bar{\alpha}_i$, $\beta_j = \bar{\beta}_j$, $c_{ij} = \bar{c}_{ij}$ and $N\mu_i^- = \bar{\mu}_i^-$ yields equivalence of the latter equation to

$$\frac{1}{N} \sum_{j=0}^{N-1} (\bar{\alpha}_i + \bar{\beta}_j - \bar{c}_{ij})_+ = \gamma \frac{1}{N} \bar{\mu}_i^-,$$

which is equivalent to the equation that is solved by Algorithm 2. The argument for $\bar{\mu}_j^+$ is carried out analogously.

Regarding termination, the solver checks the criteria

$$\left| \frac{1}{\gamma} \sum_{j=0}^{M-1} \bar{\pi}_{ij} - \bar{\mu}_i^- \right| < \tau \quad \text{and} \quad \left| \frac{1}{\gamma} \sum_{i=0}^{M-1} \bar{\pi}_{ij} - \bar{\mu}_j^+ \right| < \tau.$$

We only consider the first and plug the identity $\gamma\pi_{ij} = \bar{\pi}_{ij}$ into it, which gives equivalence to

$$\left| \sum_{j=0}^{M-1} \pi_{ij} - N\mu_i^- \right| < \tau.$$

This in turn is equivalent to

$$N^2 \left| \sum_{j=0}^{M-1} \int_{\frac{j}{N}}^{\frac{j+1}{N}} \int_{\frac{j}{N}}^{\frac{j+1}{N}} \pi(x, y) \, dy dx - \int_{\frac{j}{N}}^{\frac{j+1}{N}} \mu^-(x) \, dx \right| < \tau,$$

$$\left| \int_{\frac{j}{N}}^{\frac{j+1}{N}} \int_0^1 \pi(x, y) \, dy - \mu^-(x) \, dx \right| < \frac{\tau}{N^2}.$$

Moreover, the ansatz functions for π and μ^- are constant on $\left(\frac{j}{N}, \frac{j+1}{N}\right)$, which induces equivalence to

$$\int_{\frac{j}{N}}^{\frac{j+1}{N}} \left| \int_0^1 \pi(x, y) \, dy - \mu^-(x) \right| dx < \frac{\tau}{N^2}.$$

This implies that if the solver terminates, we have

$$\left\| \int_0^1 \pi(\cdot, y) \, dy - \mu^-(\cdot) \right\|_{L^1((0,1))} < \frac{\tau}{N}.$$

We summarize the choices for the consistent mapping of quantities arising from the discretization to quantities the solver operates on in Table 1. Finally, we make a note on the calculation of the coefficients c_{ij} for the cost function $c(x, y) := (x - y)^2$:

$$c_{ij} = N^2 \int_{\frac{j}{N}}^{\frac{j+1}{N}} \int_{\frac{j}{N}}^{\frac{j+1}{N}} (x - y)^2 \, dy dx = \dots = \frac{1}{N^2} \left((i - j)^2 + \frac{1}{6} \right).$$

Conflict of Interest: The authors declare that they have no conflict of interest.

Table 1 Mapping discretization quantities to solver quantities.

Coefficient	Solver Quantity	Conversion
π_{ij}	$\bar{\pi}_{ij}$	$\bar{\pi}_{ij} = \gamma\pi_{ij}$
c_{ij}	\bar{c}_{ij}	$\bar{c}_{ij} = c_{ij}$
μ_i^-	$\bar{\mu}_i^-$	$\bar{\mu}_i^- = N\mu_i^-$
μ_j^+	$\bar{\mu}_j^+$	$\bar{\mu}_j^+ = N\mu_j^+$
α_i	$\bar{\alpha}_i$	$\bar{\alpha}_i = \alpha_i$
β_j	$\bar{\beta}_j$	$\bar{\beta}_j = \beta_j$
J	\bar{J}	$\bar{J} = JN^2\gamma$
τ	$\bar{\tau}$	$\bar{\tau} = \tau N$

References

- Luigi Ambrosio and Nicola Gigli. A user’s guide to optimal transport. In *Modelling and optimisation of flows on networks*, pages 1–155. Springer, 2013.
- Dimitri P. Bertsekas. *Nonlinear programming*. Athena Scientific Optimization and Computation Series. Athena Scientific, Belmont, MA, third edition, 2016.
- Mathieu Blondel, Vivien Seguy, and Antoine Rolet. Smooth and sparse optimal transport. In Amos Storkey and Fernando Perez-Cruz, editors, *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*, volume 84 of *Proceedings of Machine Learning Research*, pages 880–889, Playa Blanca, Lanzarote, Canary Islands, 09–11 Apr 2018. PMLR.
- Guillaume Carlier, Vincent Duval, Gabriel Peyré, and Bernhard Schmitzer. Convergence of entropic schemes for optimal transport and gradient flows. *SIAM Journal on Mathematical Analysis*, 49(2):1385–1418, 2017.
- Xiaojun Chen. Superlinear convergence of smoothing quasi-newton methods for nonsmooth equations. *Journal of Computational and Applied Mathematics*, 80(1):105 – 126, 1997.
- Xiaojun Chen. On convergence of SOR methods for nonsmooth equations. *Numer. Linear Algebra Appl.*, 9(1):81–92, 2002.
- Xiaojun Chen, Zuhair Nashed, and Liqun Qi. Smoothing methods and semismooth methods for nondifferentiable operator equations. *SIAM J. Numer. Anal.*, 38(4):1200–1216, 2000.
- Christian Clason, Dirk A Lorenz, Hinrich Mahler, and Benedikt Wirth. Entropic regularization of continuous optimal transport problems. *arXiv preprint arXiv:1906.01333*, 2019.
- Marco Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in neural information processing systems*, pages 2292–2300, 2013.
- Marco Cuturi and Gabriel Peyré. A smoothed dual approach for variational Wasserstein problems. *SIAM J. Imaging Sci.*, 9(1):320–343, 2016.
- Montacer Essid and Justin Solomon. Quadratically regularized optimal transport on graphs. *SIAM Journal on Scientific Computing*, 40(4):A1961–A1986, 2018.
- Irene Fonseca and Giovanni Leoni. *Modern methods in the calculus of variations: L^p spaces*. Springer Monographs in Mathematics. Springer, New York, 2007.
- Aude Genevay, Gabriel Peyre, and Marco Cuturi. Learning generative models with sinkhorn divergences. In *International Conference on Artificial Intelligence and Statistics*, pages 1608–1617, 2018.
- Michael Hinze, René Pinnau, Michael Ulbrich, and Stefan Ulbrich. *Optimization with PDE constraints*, volume 23. Springer Science & Business Media, 2008.
- Leonid V. Kantorovič. On the translocation of masses. *C. R. (Doklady) Acad. Sci. URSS (N.S.)*, 37:199–201, 1942.
- Nicolas Papadakis, Gabriel Peyré, and Edouard Oudet. Optimal transport with proximal splitting. *SIAM J. Imaging Sci.*, 7(1):212–238, 2014.
- Gabriel Peyré and Marco Cuturi. Computational optimal transport. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607, 2019.
- Svetlozar T. Rachev and Ludger Rüschendorf. *Mass transportation problems. Vol. I. Probability and its Applications* (New York). Springer-Verlag, New York, 1998. Theory.

19. Svetlozar T. Rachev and Ludger Rüschendorf. *Mass transportation problems. Vol. II. Probability and its Applications* (New York). Springer-Verlag, New York, 1998. Applications.
20. Lucas Roberts, Leo Razoumov, Lin Su, and Yuyang Wang. Gini-regularized optimal transport with an application to spatio-temporal forecasting. arXiv preprint arXiv:1712.02512, 2017.
21. Filippo Santambrogio. *Optimal transport for applied mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser/Springer, Cham, 2015. Calculus of variations, PDEs, and modeling.
22. Fredi Tröltzsch. Regular Lagrange multipliers for control problems with mixed pointwise control-state constraints. *SIAM Journal on Optimization*, 15:616–634, 2005.
23. Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
24. Cédric Villani. *Optimal transport. Old and new*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009.
25. Stephen J. Wright. Coordinate descent algorithms. *Math. Program.*, 151(1, Ser. B):3–34, 2015.