Calc. Var. (2012) 43:529–536
DOI 10.1007/s00526-011-0421-z

Calculus of Variations

On the local structure of optimal measures in the multi-marginal optimal transportation problem

Brendan Pass

Received: 10 June 2010 / Accepted: 15 March 2011 / Published online: 30 March 2011

Abstract We consider an optimal transportation problem with more than two marginals. We use a family of semi-Riemannian metrics derived from the mixed, second order partial derivatives of the cost function to provide upper bounds for the dimension of the support of the solution.

Mathematics Subject Classification (2000) 49N99 · 58E17 · 91B25

1 Introduction

The optimal transportation problem (with two marginals) asks what is the most efficient way to transform one distribution of mass to another relative to a given cost function. The problem was originally posed by Monge in 1781 [19]. In 1942, Kantorovich [12] proposed a relaxed version of the problem; roughly speaking, he allowed a piece of mass to be split between two or more target points. Since then, these problems have been studied extensively by many authors and have found applications in such diverse fields as geometry, fluid mechanics, statistics, economics, shape recognition, inequalities, meteorology, etc.

Here we study a multi-marginal generalization of the above; how do we align m distributions of mass with maximal efficiency, again relative to a prescribed cost function. Precisely, given Borel probability measures μ_i on smooth manifolds M_i of respective dimensions n_i, for $i = 1, 2, \dots, m$ and a continuous cost function $c : M_1 \times M_2 \times \dots \times M_m \to \mathbb{R}$, we would like to minimize

$$C(\gamma) = \int\limits_{M_1 \times M_2 \times \dots \times M_m} c(x_1, x_2, \dots, x_m) d\gamma$$

Communicated by L. Ambrosio.

B. Pass (✉)
Department of Mathematics, University of Toronto, Toronto, ON M5S 2E3, Canada
e-mail: bpass@math.utoronto.ca

among all measures γ on $M_1 \times M_2 \times \cdots \times M_m$ which project to the μ_i under the canonical projections; that is, for any Borel subset $A \subset M_i$,

$$\gamma(M_1 \times M_2 \times \cdots \times M_{i-1} \times A \times M_{i+1} \cdots \times M_m) = \mu_i(A).$$

When $m = 2$, we recover Kantorovich's formulation of the classical optimal transportation problem.

Under mild conditions, a minimizer γ will exist. The support of γ, which we will denote by $spt(\gamma)$, is defined as the smallest closed subset of $M_1 \times M_2 \times \cdots \times M_m$ of full mass. When $m = 2$ and the cost function satisfies a twist condition[1] the solution γ is unique and is contained in the graph of a function from M_1 to M_2, provided the first marginal is suitably regular; this function then solves the original problem posed by Monge [2, 3, 8, 9, 16]. Assuming M_1 and M_2 are both C^2 smooth manifolds of common dimension n, the present author, together with McCann and Warren, has shown that under a related non-degeneracy condition on c, $spt(\gamma)$ must be contained in an n-dimensional Lipschitz submanifold of $M_1 \times M_2$ [17].

Whereas the two marginal problem is relatively well understood, results concerning the structure of these optimal measures have thus far been elusive for $m > 2$. Much of the progress to date has been in the special case where the M_i's are all Euclidean domains of common dimension n and the cost function is given by $c(x_1, x_2, \ldots, x_m) = \sum_{i \neq j} |x_i - x_j|^2$, or equivalently $c(x_1, x_2, \ldots, x_m) = -|(\sum_i x_i)|^2$. When $m = 3$, partial results for this cost were obtained by Olkin and Rachev [20], Knott and Smith [15] and Rüschendorf and Uckelmann [23], before Gangbo and Święch [10] proved that for a general m, under a mild regularity condition on the first marginal, there is unique solution to the Kantorovich problem and it is concentrated on the graph of a function over x_1, hence inducing a solution to a Monge type problem; an alternate proof of Gangbo and Święch's theorem was subsequently found by Rüschendorf and Uckelmann [24][2]. This result was then extended by Heinich [11] to cost functions of the form $c(x_1, x_2, \ldots, x_m) = h(\sum_i x_i)$ where h is strictly concave and, in the case when the domains M_i are all 1-dimensional, by Carlier [4] to cost functions satisfying a strict 2-monotonicity condition. More recently, Carlier and Nazaret [6] studied the related problem of maximizing the determinant (or its absolute value) of the matrix whose columns are the elements $x_1, x_2, \ldots, x_n \in \mathbb{R}^n$; unlike the results in [4, 10, 11], the solution in this problem may not be concentrated on the graph of a function over one of the x_i's and may not be unique. The proofs of many of these results exploit a duality theorem, proved in the multi-marginal setting by Kellerer [13]. Although duality holds for general cost functions, it alone says little about the structure of the optimal measure; the proofs of each of the aforementioned results rely heavily on the special forms of the cost.

In the present manuscript, we establish an upper bound on $\dim(spt(\gamma))$. This bound will depend on the signatures of a family of semi-Riemannian metrics derived from the mixed, second order partial derivatives of c, reminiscent of the pseudo-metric introduced by Kim and McCann [14] to study the regularity of optimal maps in two marginal problems[3]. In the case when the n_i's are equal to some common value n, this bound may or may not be n; we show by example that when the bound is larger than n, the solution may be supported on a

[1] The twist condition asserts that for fixed x_1, the mapping $x_2 \mapsto D_{x_1}c(x_1, x_2)$ is injective, where $D_{x_1}c(x_1, x_2)$ denotes the differential of c with respect to x_1; see, for example, Villani's text [26].

[2] See also a recent paper by Agueh and Carlier [1], relating this problem to barycenters in Wasserstein space.

[3] For the purposes of this paper, the term semi-Riemannian metric will refer to a symmetric, covariant 2-tensor (which is not necessarily non-degenerate). The term pseudo-Riemannian metric will be reserved for semi-Riemannian metrics which are also non-degenerate.

higher dimensional submanifold and may not be unique. In fact, the costs in these examples satisfy naive, multi-marginal extensions of both the twist and non-degeneracy conditions; given the aforementioned results in the two marginal case, we found it surprising that higher dimensional solutions can exist for twisted, non-degenerate costs. On the other hand, if the support of at least one of the measures μ_i has Hausdorff dimension n, the dimension of $spt(\gamma)$ must be at least n; therefore, in cases where our upper bound is n, the support is exactly n-dimensional, in which case we show it is actually n-rectifiable.

Like our work in [17] and in contrast to the results of Gangbo and Święch [10], Heinich [11], and Carlier [4], our results here only concern the local structure of the optimizer γ and cannot be easily used to assert uniqueness of γ or the existence of a solution to an appropriate Monge type problem. We address these problems in a separate paper [21].

In the following section, we prove our main result, while in Sect. 3 we briefly discuss applications of this result to several example cost functions.

2 Dimension of the support

Before stating our main result, we must introduce some notation. Suppose that $c \in C^2(M_1 \times M_2 \times \cdots \times M_m)$. Consider the set P of all partitions of the set $\{1, 2, 3, \dots, m\}$ into two disjoint, nonempty subsets; note that P has $2^{m-1}-1$ elements. For any partition $p \in P$, label the corresponding subsets p_+ and p_-; thus, $p_+ \cup p_- = \{1, 2, 3, \dots, m\}$ and $p_+ \cap p_-$ is empty. For each $p \in P$, define the following symmetric, bi-linear form on $M_1 \times M_2 \times \cdots \times M_m$

$$g_p = \sum_{j \in p_+, k \in p_-} \frac{\partial^2 c}{\partial x_j^{\alpha_j} \partial x_k^{\alpha_k}} \left(dx_j^{\alpha_j} \otimes dx_k^{\alpha_k} + dx_k^{\alpha_k} \otimes dx_j^{\alpha_j} \right) \tag{1}$$

where, in accordance with the Einstein summation convention, summation on the α_k and α_j is implicit.

Definition 2.1 We will say that a subset S of $M_1 \times M_2 \times \cdots \times M_m$ is c-monotone with respect to a partition p if for all $y = (y_1, y_2, \dots, y_m)$ and $\tilde{y} = (\tilde{y_1}, \tilde{y_2}, \dots, \tilde{y_m})$ in S we have

$$c(y) + c(\tilde{y}) \leq c(z) + c(\tilde{z}),$$

where

$$\begin{aligned} z_i &= y_i \quad \text{and} \quad \tilde{z_i} = \tilde{y_i}, \quad \text{if } i \in p_+, \\ z_i &= \tilde{y_i} \quad \text{and} \quad \tilde{z_i} = y_i, \quad \text{if } i \in p_-, \end{aligned}$$

The following lemma, which is well known when $m = 2$, provides the link between c-monotonicity and optimal transportation.

Lemma 2.2 *Suppose γ is an optimizer and $C(\gamma) < \infty$. Then the support of γ is c-monotone with respect to every partition $p \in P$.*

Proof Define $M_{p_+} = \otimes_{i \in p_+} M_i$ and $M_{p_-} = \otimes_{i \in p_-} M_i$. Note that we can identify $M_1 \times M_2 \times \cdots \times M_m$ with $M_{p_+} \times M_{p_-}$ and let γ_{p_+} and γ_{p_-} be the projections of γ onto M_{p_+} and M_{p_-}, respectively. Consider the two marginal problem

$$\inf \int\limits_{M_{p_+} \times M_{p_-}} c(x_1, x_2, \dots, x_m) d\lambda,$$

where the infimum is taken over all measures λ whose projections onto M_{p_+} and M_{p_-} are γ_{p_+} and γ_{p_-}, respectively. Then γ is optimal for this problem and, as c is continuous, the result follows from c-monotonicity for two marginal problems; see, for example [25]. □

We are now ready to state our main result:

Theorem 2.3 *Let g be a convex combination of the g_p's defined in Eq. 1; that is $g = \sum_{p\in P} t_p g_p$ where $0 \leq t_p \leq 1$ for all $p \in P$ and $\sum_{p\in P} t_p = 1$. Suppose γ is an optimizer and $C(\gamma) < \infty$; choose a point $x = (x_1, x_2, \ldots, x_m) \in M_1 \times M_2 \times \cdots \times M_m$. Let $N = \sum_{i=1}^m n_i$. Suppose the $(+, -, 0)$ signature of g at $(x_1, x_2, \ldots, x_m)$ is $(q_+, q_-, N - q_+ - q_-)$ (i.e., the corresponding matrix has q_+ positive eigenvalues, q_- negative eigenvalues and a zero eigenvalue with multiplicity $N - q_+ - q_-$). Then there is a neighbourhood O of $(x_1, x_2, \ldots, x_m)$ such that the intersection of the support of γ with O is contained in a Lipschitz submanifold of dimension $N - q_-$. If the support is differentiable at x, it is timelike for g; that is, $g(v, v) \leq 0$ for all $v \in T_x(spt(\gamma))$.*

Our proof is an adaptation of our argument with McCann and Warren in [17]. When $m = 2$, after choosing appropriate coordinates, we rotated the coordinate system and showed that c-monotonicity implied that the solution was concentrated on a Lipschitz graph over the diagonal, a trick dating back to Minty [18]. When passing to the multi-marginal setting, however, it is not immediately clear how to choose coordinates that make an analogous rotation possible; unlike in the two marginal case, it is not possible in general to choose coordinates around a point $(x_1, x_2, \ldots, x_m)$ such that $D^2_{x_i x_j} c(x_1, x_2, \ldots, x_m) = I$ for all $i \neq j$. The key to resolving this difficulty is the observation that Minty's trick amounts to diagonalizing the pseudo-metric of Kim and McCann and that this approach generalizes to $m \geq 3$.

Proof Choose a point $x = (x_1, x_2, \ldots, x_m) \in M_1 \times M_2 \times \cdots \times M_m$. Choose local coordinates around x_i on each M_i and set $A_{ij} = D^2_{x_i x_j} c(x_1, x_2, \ldots, x_m)$. For any $\epsilon > 0$, there is a neighbourhood O of $(x_1, x_2, \ldots, x_m)$ which is convex in these coordinates such that for all $(y_1, y_2, \ldots, y_m) \in O$ we have $||A_{ij} - D^2_{x_i x_j} c(y_1, y_2, \ldots, y_m)|| \leq \epsilon$, for all $i \neq j$.

Let G be the matrix of g at x in our chosen coordinates. There exists some invertible N by N matrix U such that

$$UGU^T = H := \begin{bmatrix} I & 0 & 0 \\ 0 & -I & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

where the diagonal I, $-I$ and 0 blocks have sizes determined by the signature of g.

Define new coordinates in O by $u := Uy$, where $y = (y_1, y_2, \ldots, y_m)$ and let $u = (u_1, u_2, u_3)$ be the obvious decomposition. We will show that the optimizer is locally contained in a Lipschitz graph in these coordinates.

Choose $y = (y_1, y_2, \ldots, y_m)$ and $\tilde{y} = (\tilde{y_1}, \tilde{y_2}, \ldots, \tilde{y_m})$ in the intersection of $spt(\gamma)$ and O. Set $\Delta y = y - \tilde{y}$. Define $z = (z_1, z_2, \ldots z_m)$ and $\tilde{z} = (\tilde{z_1}, \tilde{z_2}, \ldots, \tilde{z_m})$ as in Lemma 2.2; we then have

$$c(y) + c(\tilde{y}) \leq c(z) + c(\tilde{z})$$

or

$$\int_0^1 \int_0^1 \sum_{j\in p_+, i\in p_-} (\Delta y_i)^T D^2_{x_i x_j} c(y(s,t)) \Delta y_j dt ds \leq 0,$$

where

$$y_i(s,t) = \begin{cases} y_i + s(\Delta y_i) & \text{if } i \in p_+, \\ y_i + t(\Delta y_i) & \text{if } i \in p_-. \end{cases}$$

This implies that

$$\sum_{j\in p_+, i\in p_-} (\Delta y_i)^T A_{ij} \Delta y_j \leq \epsilon \sum_{j\in p_+, i\in p_-} |\Delta y_i||\Delta y_j|.$$

Hence,

$$\sum_{p\in P} t_p \sum_{j\in p_+, i\in p_-} (\Delta y_i)^T A_{ij} \Delta y_j \leq \epsilon \sum_{p\in P} t_p \sum_{j\in p_+, i\in p_-} |\Delta y_i||\Delta y_j|.$$

But this means

$$(\Delta y)^T G \Delta y \leq \epsilon \sum_{p\in P} t_p \sum_{j\in p_+, i\in p_-} |\Delta y_i||\Delta y_j|. \tag{2}$$

With $\Delta u = U\Delta y$ and $\Delta u = (\Delta u_1, \Delta u_2, \Delta u_3)$ being the obvious decomposition, this becomes:

$$\begin{aligned} |\Delta u_1|^2 - |\Delta u_2|^2 = (\Delta u)^T H \Delta u = (\Delta y)^T G \Delta y \\ \leq \epsilon \sum_{p\in P} t_p \sum_{j\in p_+, i\in p_-} |\Delta y_i||\Delta y_j| \\ \leq \epsilon m^2 ||U^{-1}||^2 \sum_i^3 |\Delta u_i|^2, \end{aligned}$$

where the last line follows because for each i and j we have

$$\begin{aligned} |\Delta y_i||\Delta y_j| &\leq |\Delta y|^2 \\ &\leq ||U^{-1}||^2 |\Delta u|^2 \\ &= ||U^{-1}||^2 \sum_{i=1}^3 |\Delta u_i|^2. \end{aligned}$$

Choosing ϵ sufficiently small, we have

$$|\Delta u_1|^2 - |\Delta u_2|^2 \leq \frac{1}{2} \sum_i^3 |\Delta u_i|^2.$$

Rearranging yields

$$\frac{1}{2}|\Delta u_1|^2 \leq \frac{3}{2}|\Delta u_2|^2 + \frac{1}{2}|\Delta u_3|^2.$$

Together with Kirzbraun's theorem, the above inequality implies that the support of γ is locally contained in a Lipschitz graph of u_1 over u_2 and u_3.

If $spt(\gamma)$ is differentiable at x, the timelike implication follows from taking $y = x$ in (2), then noting that we can take $\epsilon \to 0$ as $\tilde{y} \to x$. □

3 Examples

In this section, we outline some applications of Theorem 2.3. The proofs of the following results are straightforward calculations; they will appear, along with some additional applications, in [22]. Throughout this section, we restrict our attention to the special case where the weights t_p in Theorem 2.3 are all $\frac{1}{2^{m-1}-1}$; in this case, we obtain the special semi-metric

$$\overline{g} = \frac{2^{m-2}}{2^{m-1}-1} \sum_{j \neq k} \frac{\partial^2 c}{\partial x_j^{\alpha_j} \partial x_k^{\alpha_k}} \left(dx_j^{\alpha_j} \otimes dx_k^{\alpha_k} + dx_k^{\alpha_k} \otimes dx_j^{\alpha_j} \right) \tag{3}$$

3.1 Applications to two marginal problems

We first consider the classical, two marginal problem. In this case, $\overline{g}$ is the Kim–McCann pseudo-metric; assuming $n_1 = n_2 := n$ and $D^2_{x_1 x_2} c$ is non-degenerate, its signature is $(n, n, 0)$ [14]. Theorem 2.3 is then precisely the main result of the present author, together with McCann and Warren, in [17].

Theorem 2.3 actually generalizes this result even when $m = 2$, as we assume here neither non-degeneracy, nor even equality of the dimensions $n_1 = n_2$.

Proposition 3.1.1 *Suppose* $m = 2$*. If the rank of* $D^2_{x_1 x_2} c$ *is* r*, the signature of* $\overline{g}$ *in Eq. (3) is* $(r, r, n_1 + n_2 - 2r)$*.*

Theorem 2.3 then implies that $spt(\gamma)$ is at most $(n_1 + n_2 - r)$-dimensional. It is worth noting that the topology of many important manifolds prohibits the non-degeneracy condition from holding everywhere. Suppose, for example, that $M_1 = M_2 = S^1$, the unit circle. Then periodicity in x_1 of $\frac{\partial c}{\partial x_2}(x_1, x_2)$ implies

$$\int_{S^1} \frac{\partial^2 c}{\partial x_1 \partial x_2}(x_1, x_2) dx_1 = 0.$$

It follows that for every x_2 there is at least one x_1 such that $\frac{\partial^2 c}{\partial x_1 \partial x_2}(x_1, x_2) = 0$.

3.2 Functions of the sum: $c(x_1, x_2, \ldots, x_m) = h\left(\sum_{i=1}^m x_i\right)$

Next we consider the case where $M_i = \mathbb{R}^n$ for all i and $c(x_1, x_2, \ldots, x_m) = h(\sum_{i=1}^m x_i)$.

Proposition 3.2.1 *Suppose* $M_i = \mathbb{R}^n$ *for all* i *and* $c(x_1, x_2, \ldots, x_m) = h(\sum_{i=1}^m x_i)$*. Denote the signature of* $D^2 h$ *by* $(q_+, q_-, n - q_+ - q_-)$*; then the signature of* $\overline{g}$ *in Eq. (3) is* $\big(q_+ + (m-1)(q_-), q_- + (m-1)q_+, m(n - q_+ - q_-)\big)$*.*

Remark 3.2.2 When $D^2 h$ is negative definite (corresponding to a uniformly concave h), the signature of $\overline{g}$ reduces to $((m-1)n, n, 0)$; combined with Theorem 2.3, this implies that the support of any optimal measure γ is contained in an n-dimensional submanifold. This is consistent with the results of Gangbo and Święch [10] and Heinich [11], who show that if the first marginal assigns measure zero to every set of Hausdorff dimension $n - 1$, then $spt(\gamma)$ is contained in the graph of a function over x_1.

On the other hand, when $D^2 h$ is not negative definite, the signature of $\overline{g}$ has more than n timelike directions. In this case, Theorem 2.3 does not preclude optimal measures with higher dimensional support. The next result verifies that this can in fact occur.

Proposition 3.2.3 *Let* $c(x_1, x_2, \ldots, x_m) = h(\sum_{i=1}^m x_i)$, *where the signature of* D^2h *is* $(q, n-q, 0)$. *Then there exist optimal measures whose support has dimension* $(n - q + q(m-1))$.

Finally, the following proposition implies that when the dimension of $spt(\gamma)$ is larger than n, the solution may not be unique.

Proposition 3.2.4 *Set* $m = 4$ *and* $c(x, y, z, w) = h(x + y + z + w)$ *for h strictly convex. Suppose all four marginals* μ_i *are Lebesgue measure on the unit cube* I^n *in* $\mathbb{R}^n$. *Then the optimal measure is not unique.*

It is worth noting that this cost is twisted; that is, the maps $x_i \mapsto D_{x_j} c(x_1, x_2, \ldots, x_m)$ are injective for all $i \neq j$ where x_k is held fixed for all $k \neq i$. In the two marginal case, the twist condition and mild regularity on μ_1 suffice to imply the uniqueness of the solution γ [16]; this example demonstrates that this is no longer true for $m \geq 3$.

3.3 Hedonic pricing costs

Our next example has an economic motivation. Chiappori et al. [7] and Carlier and Ekeland [5] introduced a hedonic pricing model based on a multi-marginal optimal transportation problem with cost function of the form $c(x_1, x_2, \ldots, x_m) = \inf_{y \in Y} \sum_{i=1}^m f_i(x_i, y)$. Combined with Theorem 2.3, the following result demonstrates that, assuming all the dimensions $n_i = n$ are equal, the support of the opimizer is at most n-dimensional.

Proposition 3.3.1 *Suppose* $n_i = n$ *for all i and let* $c(x_1, x_2, \ldots, x_m) = \inf_{y \in Y} \sum_{i=1}^m f_i(x_i, y)$, *where Y is a* C^2, *n-dimensional manifold. Assume the following conditions:*

1. *For all i,* f_i *is* C^2 *and the matrix* $D^2_{x_i y} f_i$ *of mixed , second order partial derivatives is everywhere non-singular.*
2. *For each* $(x_1, x_2, \ldots, x_m)$ *the infinum is attained by a unique* $y(x_1, x_2, \ldots, x_m) \in Y$.
3. $\sum_{i=1}^m D^2_{yy} f_i(x_i, y(x_1, x_2, \ldots, x_m))$ *is non-singular.*

Then the signature of $\overline{g}$ *in Eq. (3) is* $((m-1)n, n, 0)$.

Acknowledgments The author is indebted to his advisor Robert McCann, for originally suggesting this problem to him, for many fruitful discussions and for his continuous support. He would also like to thank the members of his PhD supervisory committee, Jim Colliander and Bob Jerrard, for providing constructive feedback, and Man-Duen Choi and Chandler Davis for useful discussions concerning the signatures of symmetric matrices. The author was supported in part by an NSERC postgraduate scholarship. This work was completed in partial fulfillment of the requirements of a doctoral degree in mathematics at the University of Toronto.

References

1. Agueh, M., Carlier, G.: Barycenters in the Wasserstein space. SIAM J. Math. Anal. (to appear)
2. Brenier, Y.: Decomposition polaire et rearrangement monotone des champs de vecteurs. C.R. Acad. Sci. Paris I Math. **305**, 805–808 (1987)
3. Caffarelli, L.: Allocation maps with general cost funtions. In: Marcellini, P., Talenti, G., Vesintini, E. (eds.) Partial Differential Equations and Applications. Lecture Notes in Pure and Applied Mathematics, vol. 177, pp. 29–35. Dekker, New York (1996)
4. Carlier, G.: On a class of multidimensional optimal transportation problems. J. Convex Anal. **10**(2), 517–529 (2003)
5. Carlier, G., Ekeland, I.: Matching for teams. Econ. Theory **42**(2), 397–418 (2010)
6. Carlier, G., Nazaret, B.: Optimal transportation for the determinant. ESAIM Control Optim. Calc. Var. **14**(4), 678–698 (2008)

7. Chiappori, P.-A., McCann, R., Nesheim, L.: Hedonic price equilibria, stable matching and optimal transport; equivalence, topology and uniqueness. Econ. Theory **42**(2), 317–354 (2010)
8. Gangbo, W.: Habilitation Thesis, Universite de Metz (1995)
9. Gangbo, W., McCann, R.: The geometry of optimal transportation. Acta Math. **177**, 113–161 (1996)
10. Gangbo, W., Święch, A.: Optimal maps for the multidimensional Monge-Kantorovich problem. Commun. Pure Appl. Math. **51**(1), 23–45 (1998)
11. Heinich, H.: Probleme de Monge pour *n* probabilities. C.R. Math. Acad. Sci. Paris **334**(9), 793–795 (2002)
12. Kantorovich, L.: On the translocation of masses. C.R. (Dokl.) Acad. Sci. URSS (N.S.) **37**, 199–201 (1942)
13. Kellerer, H.G.: Duality theorems for marginal problems. Z. Wahrsch. Verw. Gebiete **67**, 399–432 (1984)
14. Kim, Y.-H., McCann, R.: Continuity, curvature, and the general covariance of optimal transportation. J. Eur. Math. Soc. **12**, 1009–1040 (2010)
15. Knott, M., Smith, C.C.: On a generalization of cyclic monotonicity and distances among random vectors. Linear Algebra Appl. **199**, 363–371 (1994)
16. Levin, V.: Abstract cyclical monotonicity and Monge solutions for the general Monge-Kantorovich problem. Set-Valued Anal. **7**(1), 7–32 (1999)
17. McCann, R., Pass, B., Warren, M.: Rectifiability of optimal transportation plans. Preprint available at arXiv:1003.4556v1
18. Minty, G.J.: Monotone (nonlinear) operators in Hilbert space. Duke Math. J. **29**, 341–346 (1962)
19. Monge, G.: Memoire sur la theorie des deblais et de remblais. In: Histoire de l'Academie Royale des Sciences de Paris, avec les Memoires de Mathematique et de Physique pour la meme annee, pp. 666–704. (1781)
20. Olkin, I., Rachev, S.T.: Maximum submatrix traces for positive definite matrices. SIAM J. Matrix Anal. Appl. **14**, 390–397 (1993)
21. Pass, B.: Uniqueness and Monge solutions in the multi-marginal optimal transportation problem. http://arxiv.org/abs/1007.0424
22. Pass, B.: PhD Thesis, University of Toronto (2011)
23. Rüschendorf, L., Uckelmann, L.: On optimal multivariate couplings. In: Benes, V., Stepan, I. (eds.) Distributions with Given Marginals and Moment Problems (Prague, Czech Republic, 1996), pp. 261–274. Kluwer Academic publishers, Dordrecht (1997)
24. Rüschendorf, L., Uckelmann, L.: On the *n* coupling problem. J. Multivar. Anal. **81**, 242–258 (2002)
25. Smith, C., Knott, M.: On Hoeffding-Frechet bounds and cyclic monotone relations. J. Multivar. Anal. **40**, 328–334 (1992)
26. Villani, C.: Optimal transport: old and new, volume 338 of Grundlehren der mathematischen Wissenschaften. Springer, New York (2009)