SIAM J. MATH. ANAL.
Vol. 43, No. 6, pp. 2758–2775

UNIQUENESS AND MONGE SOLUTIONS IN THE MULTIMARGINAL OPTIMAL TRANSPORTATION PROBLEM*

BRENDAN PASS†

Abstract. We study a multimarginal optimal transportation problem. Under certain conditions on the cost function and the first marginal, we prove that the solution to the relaxed, Kantorovich version of the problem induces a solution to the Monge problem and that the solutions to both problems are unique. We also exhibit several examples of cost functions under which our conditions are satisfied, including one arising in a hedonic pricing model in mathematical economics.

Key words. optimal transportation, Monge–Kantorovich problem, multimarginal problem, Monge solutions, uniqueness, hedonic pricing

AMS subject classifications. 49J30, 49K30, 49J52, 58C35, 58E17, 91B68

DOI. 10.1137/100804917

1. Introduction. Given two Borel probability measures μ_1 and μ_2 on spaces M_1 and M_2, respectively, and a cost function $c : M_1 \times M_2 \to \mathbb{R}$, Monge's optimal transportation problem asks how to most efficiently map μ_1 onto μ_2, where efficiency is measured relative to c. We say that a function $G : M_1 \to M_2$ pushes μ_1 forward to μ_2 and write $G_\#\mu_1 = \mu_2$ if $\mu_1(G^{-1}(B)) = \mu_2(B)$ for all measurable $B \subseteq M_2$. Monge's problem is then to minimize $\int_{M_1} c(x_1, G(x_1))d\mu_1$ among all G that push μ_1 forward to μ_2. Due to both its deep connections to other areas of mathematics and its applicability in other fields, optimal transportation has grown into a thriving field of research over the past 20 years; we refer the interested reader to the books by Villani for references [24], [25].

In this paper, we are interested in a multimarginal version of the preceding problem. Given Borel probability measures μ_i on n-dimensional, smooth manifolds M_i for $i = 1, 2, \dots, m$, and a *cost* function $c : M_1 \times M_2 \times \cdots \times M_m \to \mathbb{R}$, the multimarginal version of Monge's optimal transportation problem is to minimize

$$\text{(M)} \qquad C(G_2, G_3, \dots, G_m) := \int_{M_1} c(x_1, G_2(x_1), G_3(x_1), \dots, G_m(x_1))d\mu_1$$

among all $(m-1)$-tuples of measurable maps $(G_2, G_3, \dots, G_m)$, where $G_i : M_1 \to M_i$ pushes μ_1 forward to μ_i for all $i = 2, 3, \dots, m$. The Kantorovich formulation of the problem is to minimize

$$\text{(K)} \qquad C(\mu) := \int_{M_1 \times M_2 \times \cdots \times M_m} c(x_1, x_2, x_3, \dots, x_m)d\mu$$

among all positive Borel measures μ on $M_1 \times M_2 \times \cdots \times M_m$ such that the canonical projection

$$\pi_i : M_1 \times M_2 \times \cdots \times M_m \to M_i$$

*Received by the editors August 10, 2010; accepted for publication (in revised form) October 20, 2011; published electronically December 22, 2011. The author was supported in part by an NSERC postgraduate scholarship. This work was completed in partial fulfillment of the requirements for a doctoral degree in mathematics at the University of Toronto.

http://www.siam.org/journals/sima/43-6/80491.html

†Department of Mathematics, University of Toronto, Toronto, Ontario, Canada M5S 2E4 (bpass@math.utoronto.ca). Current address: Department of Mathematical and Statistical Sciences, 632 CAB, University of Alberta, Edmonton, Alberta, Canada T6G 2G1 (pass@ualberta.ca).

pushes μ forward to μ_i for all i. For any $(m-1)$-tuple $(G_2, G_3, \dots, G_m)$ such that $G_{i\#}\mu_1 = \mu_i$ for all $i = 2, 3, \dots, m$, we can define the measure $\mu = (Id, G_2, G_3, \dots, G_M)_{\#}\mu_1$ on $M_1 \times M_2 \times \cdots \times M_m$, where $Id : M_1 \to M_1$ is the identity map. Then μ projects to μ_i for all i and $C(G_2, G_3, \dots, G_m) = C(\mu)$; therefore, (K) can be interpreted as a relaxed version of (M). Roughly speaking, the difference between the two formulations is that in (M) almost every point $x_1 \in M_1$ is coupled with exactly one point $x_i \in M_i$ for each $i = 2, 3, \dots, m$, whereas in (K) an element of mass at x_i is allowed to be *split* between two or more target points in M_i for $i = 2, 3, \dots, m$.

Assuming that c is continuous, it is not hard to show that a solution to (K) exists. When $m = 2$, under a regularity condition on μ_1 and a twist condition on c, which we will define in the next section, one can show that this solution is concentrated on the graph of a function over x_1 [15], [8], [2], [9], [3]. It is then straightforward to show that this function solves (M) and to establish uniqueness results for both (M) and (K). When $m \geq 3$, however, existence and uniqueness in (M) as well as uniqueness in (K) are still largely open. In their seminal paper, Gangbo and Święch [10] used a duality theorem of Kellerer [12] to resolve these questions for the quadratic cost function, $c(x_1, x_2, x_3, \dots, x_m) = -|\sum_{i=1}^m x_i|^2$ on $\mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n$, extending partial results for the same cost by Olkin and Rachev [19], Knott and Smith [14], and Rüschendorf and Uckelmann [22]. Gangbo and Święch's theorem was then reproved using a different argument by Rüschendorf and Uckelmann [23] and generalized by Heinich [11] to cost functions of the form $c(x_1, x_2, x_3, \dots, x_m) = h(\sum_{i=1}^m x_i)$, where $h : \mathbb{R}^n \to \mathbb{R}$ is strictly concave. Carlier [4] extended these results to a still wider class of cost functions, but only in the case when all the domains M_i are one dimensional. A result with a somewhat different flavor was obtained by Carlier and Nazaret [6]. Assuming $n = m$ and $M_i = \mathbb{R}^n$ for all i, their objective was to maximize a convex function of the determinant of the matrix whose columns are the vectors x_i; for this cost function, they demonstrate that the maximizer may not be concentrated on the graph of a function over x_1 and may not be unique. The aim of the present article is to identify general conditions on the cost function under which both (K) and (M) admit unique solutions.

With one exception, the conditions we impose on c will look similar to standard conditions which arise when studying the two marginal problem. Our lone novel hypothesis is that a certain covariant 2-tensor on the product space $M_2 \times M_2 \times \cdots \times M_{m-1}$ should be negative definite. Tensors have recently become very relevant to optimal transportation, owing to their essential role in the study of the regularity of optimal maps. Ma, Trudinger, and Wang [17] showed that a certain fourth order differential condition on c, called (A3w) in the literature, dictates the regularity of the solution to (M) in the two marginal case. The tensoriality of (A3w) was observed by Loeper [16], and Kim and McCann [13] reinterpreted this condition, relating it to the sectional curvature of a certain pseudometric. In [20], the present author showed that the dimension of the support of the optimizers for multimarginal problems is related to a family of symmetric, bilinear forms. One surprising consequence of that work is a class of counterexamples demonstrating that the obvious generalization of the twist condition to the multimarginal setting is sufficient neither to guarantee uniqueness of minimizers for (K) nor to ensure that the solution to (K) induces a solution to (M). In these examples, the solutions were concentrated on submanifolds of dimension greater than n; motivated in part by this observation, [20] identified local conditions on c under which the support of the optimal measure is at most n-dimensional. Our condition here is a little different. Whereas the question about the dimension of the support of a solution μ to (K) is purely local, showing that μ gives rise to a solution to (M) is a global issue: for almost all $x_1 \in M_1$ we must show that there is exactly one $(x_2, x_3, \dots, x_m) \in M_2 \times M_3 \times \cdots \times M_m$ which get coupled to x_1 by μ. Our tensor here is designed to capture this global aspect of the problem. Our proof also requires a geometric condition on the domains; interestingly, this condition turns out to be somewhat reminiscent of the c-convexity condition introduced by Ma, Trudinger, and Wang [17] (see Definition 2.9 and condition IV in Theorem 3.1).

Regularity of the optimal map for multimarginal problems remains an interesting open question. Our tensorial condition is in some ways similar to (A3w), which implies regularity in the two marginal context. However, (A3w) involves fourth order derivatives of the cost function, whereas our condition involves only second order derivatives. We therefore suspect that still stronger conditions are required to ensure regularity in the multimarginal setting.

In the next section we recall relevant concepts from the theory of optimal transportation and formulate the conditions we will need. In section 3 we state and prove our main result, and in section 4 we exhibit several examples of cost functions which satisfy the criteria of our main theorem.

2. Preliminaries and definitions. We will assume that each M_i can be smoothly embedded in some larger manifold in which its closure $\overline{M_i}$ is compact and that the cost $c \in C^2(\overline{M_1} \times \overline{M_2} \times \cdots \times \overline{M_m})$. In addition, we will assume that M_i is a Riemannian manifold for $i = 2, 3, \dots, m-1$ and that any two points can be joined by a smooth, length minimizing geodesic,[1] although no such assumptions will be needed on M_1 or M_m. The requirement of a Riemannian structure is related to the global nature of (M) that we alluded to in the introduction; a Riemannian metric gives us a natural way to connect any pair of points, namely geodesics. In fact, as we alluded to in the introduction, we will impose a somewhat stronger geometric assumption on these domains later on (see Definition 2.9). We will also impose conditions on c which, much like the twist and nondegeneracy conditions for smooth costs in two marginal problems, imply that none of the M_i can be a compact manifold without boundary (see Remark 2.8). Altogether, this means that our domains are all curved analogues of bounded open sets in $\mathbb{R}^n$, with additional convexity-type assumptions on M_i, for $i = 2, 3, \dots, m-1$.

We will denote by $D_{x_i}c(x_1, x_2, \dots, x_m)$ the differential of c with respect to x_i. For $i \neq j$, the bilinear form $D^2_{x_i x_j}c(x_1, x_2, \dots, x_m)$ on $T_{x_i}M_i \times T_{x_j}M_j$ was introduced in [13]; in local coordinates, it is defined by

$$D^2_{x_i x_j} c \left\langle \frac{\partial}{\partial x_i^{\alpha_i}}, \frac{\partial}{\partial x_j^{\alpha_j}} \right\rangle = \frac{\partial^2 c}{\partial x_i^{\alpha_i} \partial x_j^{\alpha_j}}.$$

As M_i is Riemannian for $i = 2, \dots, m-1$, Hessians or unmixed, second order partial derivatives with respect to these coordinates make sense, and we will denote them by $Hess_{x_i}c(x_1, x_2, \dots, x_m)$; note, however, that no Riemannian structure is necessary to ensure the tensoriality of the mixed second order partials $D^2_{x_i x_j}c(x_1, x_2, \dots, x_m)$, as was observed in [13].

Given a Borel probability measure μ on a topological space M, the support of μ, which we will denote by $spt(\mu)$, is defined to be the smallest closed subset of M such that $\mu(spt(\mu)) = 1$.

[1] Note that we do *not* assume M_i is complete, however, as we do not wish to exclude, for example, bounded, convex domains in $\mathbb{R}^n$.

The dual problem to (K) is to maximize

$$\sum_{i=1}^{m}\int_{M_i} u_i(x_i)d\mu_i \tag{D}$$

among all m-tuples $(u_1, u_2, \dots, u_m)$ of functions $u_i \in L^1(\mu_i)$ for which $\sum_{i=1}^m u_i(x_i) \leq c(x_1, \dots, x_m)$ for all $(x_1, \dots, x_m) \in M_1 \times M_2 \times \cdots \times M_m$.

There is a special class of functions satisfying the constraint in (D) that will be of particular interest to us.

DEFINITION 2.1. *We say that an m-tuple of functions $(u_1, u_2, \dots, u_m)$ is c-conjugate if for all i,*

$$u_i(x_i) = \inf_{\substack{x_j \in M_j \\ j \neq i}} \Big(c(x_1, x_2, \dots, x_m) - \sum_{j \neq i} u_j(x_j) \Big).$$

Whenever $(u_1, u_2, \dots, u_m)$ is c-conjugate, the u_i are semiconcave and hence have superdifferentials $\overline{\partial} u_i(x_i)$ at each point $x_i \in M_i$. By compactness, for each $x_i \in M_i$ we can find $x_j \in \overline{M_j}$ for all $j \neq i$ such that $u(x_i) = c(x_1, x_2, \dots, x_m) - \sum_{j \neq i} u_j(x_j)$; furthermore, as long as $|u_i(x_i)| < \infty$ for at least one x_i, u_i is locally Lipschitz [18].

The following theorem makes explicit the link between the Kantorovich problem and its dual.

THEOREM 2.2. *There exists a solution μ to the Kantorovich problem and a c-conjugate solution $(u_1, u_2, \dots, u_m)$ to its dual. Furthermore, the maximum value in* (D) *coincides with the minimum value in* (K)*. Finally, for any solution μ to* (K)*, any c-conjugate solution $(u_1, u_2, \dots, u_m)$ to* (D) *and any $(x_1, \dots, x_m) \in spt(\mu)$, we have $\sum_{i=1}^m u_i(x_i) = c(x_1, \dots, x_m)$.*

This result is well known in the two marginal case; for $m \geq 3$, the existence of solutions to (K) and (D) as well as the equality of their extremal values was proved in [12]. The remaining conclusions were proved for a special cost by Gangbo and Święch [10] and for a general, continuous cost when each $M_i = \mathbb{R}^n$ by Carlier and Nazaret [6]. The same proof applies for more general spaces M_i; we reproduce it below in the interest of completeness.

Proof. As mentioned above, a proof of the existence of solutions μ to (K) and $(v_1, v_2, \dots, v_m)$ to (D) as well as the equality

$$\sum_{i=1}^{m}\int_{M_i} v_i(x_i)d\mu_i = \int_{M_1 \times M_2 \times \cdots \times M_m} c(x_1, x_2, x_3, \dots, x_m)d\mu \tag{1}$$

can be found in [12]. We use a convexification trick, also found in [10] and [6], to build a c-conjugate solution to (D).

Define

$$u_1(x_1) = \inf_{\substack{x_j \in M_j \\ j \geq 2}} \left(c(x_1, x_2, \dots, x_m) - \sum_{j=2}^{m} v_j(x_j) \right)$$

and u_i inductively by

$$u_i(x_i) = \inf_{\substack{x_j \in M_j \\ j \neq i}} \left(c(x_1, x_2, \dots, x_m) - \sum_{j=1}^{i-1} u_j(x_j) - \sum_{j=i+1}^{m} v_j(x_j) \right).$$

As

$$u_m(x_m) = \inf_{\substack{x_j \in M_j \\ j \neq i}} \left(c(x_1, x_2, \dots, x_m) - \sum_{j=1}^{m-1} u_j(x_j) \right),$$

we immediately obtain

$$u_i(x_i) \leq \inf_{\substack{x_j \in M_j \\ j \neq i}} \left(c(x_1, x_2, \dots, x_m) - \sum_{j \neq i} u_j(x_j) \right). \tag{2}$$

The definition of u_{i-1} implies that for all $(x_1, x_2, \dots, x_m)$,

$$v_i(x_i) \leq c(x_1, x_2, \dots, x_m) - \sum_{j=1}^{i-1} u_j(x_j) - \sum_{j=i+1}^{m} v_j(x_j).$$

Therefore, $v_i(x_i) \leq u_i(x_i)$. It then follows that

$$\begin{aligned} u_i(x_i) &= \inf_{\substack{x_j \in M_j \\ j \neq i}} \left(c(x_1, x_2, \dots, x_m) - \sum_{j=1}^{i-1} u_j(x_j) - \sum_{j=i+1}^{m} v_j(x_j) \right) \\ &\geq \inf_{\substack{x_j \in M_j \\ j \neq i}} \left(c(x_1, x_2, \dots, x_m) - \sum_{j \neq i} u_j(x_j) \right), \end{aligned}$$

which, together with (2), implies that $(u_1, u_2, \dots, u_m)$ is c-conjugate. Now, we have

$$\begin{aligned} \sum_{i=1}^{m} \int_{M_i} v_i(x_i) d\mu_i &\leq \sum_{i=1}^{m} \int_{M_i} u_i(x_i) d\mu_i \\ &= \sum_{i=1}^{m} \int_{M_1 \times M_2 \times \cdots \times M_m} u_i(x_i) d\mu \\ &\leq \int_{M_1 \times M_2 \times \cdots \times M_m} c(x_1, x_2, x_3, \dots, x_m) d\mu, \end{aligned}$$

and so by (1) we must have

$$\sum_{i=1}^{m} \int_{M_i} u_i(x_i) d\mu_i = \sum_{i=1}^{m} \int_{M_1 \times M_2 \times \cdots \times M_m} u_i(x_i) d\mu = \int_{M_1 \times M_2 \times \cdots \times M_m} c(x_1, x_2, x_3, \dots, x_m) d\mu.$$

But because $\sum_{i=1}^{m} u_i(x_i) \leq c(x_1, x_2, x_3, \dots, x_m)$, we must have equality μ almost everywhere. Continuity then implies that equality holds on $spt(\mu)$. □

As a corollary to the duality theorem, we now prove a uniqueness result for the solution to (D). When $m = 2$, this result, under the weak conditions on c stated below, is due to Chiappori, McCann, and Nesheim [7]; for certain special, multimarginal costs, it was proven by Gangbo and Święch [10] and Carlier and Nazaret [6]. Although this result is tangential to the main goals of this article, we prove it here to emphasize that, whereas uniqueness in (K) requires certain structure conditions on the cost, uniqueness in (D) depends only on the differentiability of c.

COROLLARY 2.3. *Suppose the domains M_i are all connected, that c is continuously differentiable, and that each μ_i is absolutely continuous with respect to local coordinates with a strictly positive density. If $(v_1, v_2, \dots, v_m)$ and $(\overline{v_1}, \overline{v_2}, \dots, \overline{v_m})$ solve (D), then there exist constants t_i for $i = 1, 2, \dots, m$ such that $\sum_{i=1}^m t_i = 0$ and $v_i = \overline{v_i} + t_i$, μ_i almost everywhere for all i.*

Proof. Using the convexification trick in the proof of Theorem 2.2, we can find c-conjugate solutions $(u_1, u_2, \dots, u_m)$ and $(\overline{u_1}, \overline{u_2}, \dots, \overline{u_m})$ to (D) such that $v_i(x_i) \leq u_i(x_i)$ and $\overline{v_i}(x_i) \leq \overline{u_i}(x_i)$ for all $x_i \in M_i$. Now, as

$$\sum_{i=1}^{m} \int_{M_i} v_i(x_i) d\mu_i = \sum_{i=1}^{m} \int_{M_i} u_i(x_i) d\mu_i,$$

we must have $v_i = u_i$, μ_i almost everywhere. Similarly, $\overline{v_i} = \overline{u_i}$, μ_i almost everywhere. Now, choose $x_i \in M_i$, where u_i and $\overline{u_i}$ are differentiable. Then there exists x_j for all $j \neq i$ such that

$$(x_1, x_2, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_m) \in spt(\mu);$$

Theorem 2.2 then yields

$$u_i(x_i) - c(x_1, x_2, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_m) = -\sum_{j \neq i} u_j(x_j).$$

Because

$$u_i(z_i) - c(x_1, x_2, \dots, x_{i-1}, z_i, x_{i+1}, \dots, x_m) \leq -\sum_{j \neq i} u_j(x_j),$$

for all other $z_i \in M_i$ we must have

$$Du_i(x_i) = D_{x_i} c(x_1, x_2, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_m).$$

Similarly,

$$D\overline{u_i}(x_i) = D_{x_i} c(x_1, x_2, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_m),$$

and hence $Du_i(x_i) = D\overline{u_i}(x_i)$. As this equality holds for almost all x_i, we conclude that $u_i(x_i) = \overline{u_i}(x_i) + t_i$ for some constant t_i. Choosing any $(x_1, x_2, \dots, x_m) \in \mathrm{spt}(\mu)$ and noting that

$$\sum_{i=1}^{m} u_i(x_i) = c(x_1, x_2, \dots, x_m) = \sum_{i=1}^{m} \overline{u_i}(x_i),$$

we obtain $\sum_{i=1}^m t_i = 0$. ∎

The next two definitions are straightforward generalizations of concepts borrowed from the two marginal setting.

DEFINITION 2.4. *For $i \neq j$, we say that c is (i,j)-twisted if the map $x_j \in M_j \mapsto D_{x_i} c(x_1, x_2, \dots, x_m) \in T^*_{x_i} M_i$ is injective for all fixed x_k, $k \neq j$.*

DEFINITION 2.5. *We say that c is (i,j)-nondegenerate if $D^2_{x_i x_j} c(x_1, x_2, \dots, x_m)$, considered as a map from $T_{x_j} M_j$ to $T^*_{x_i} M_i$, is injective for all $(x_1, x_2, \dots, x_m)$.*

In local coordinates, nondegeneracy simply means that the corresponding matrix of mixed, second order partial derivatives has a nonzero determinant. When this condition holds, inverse map $T^*_{x_i} M_i \to T_{x_j} M_j$ will be denoted by $(D^2_{x_i x_j} c)^{-1}(x_1, x_2, \dots, x_m)$.

When $m = 2$, the nondegeneracy condition is not needed to ensure the existence of an optimal map (although it plays an important role in studying the regularity of that map). On the other hand, the twist condition plays an essential role in showing that Monge's problem has a solution; it ensures that a first order differential condition arising from the duality theorem can be solved uniquely for one variable as a function of the other [15] (see also [2], [9], and [3]). In light of this, one might expect that, for $m \geq 3$, if c is (i, j)-twisted for all $i \neq j$, then the Kantorovich solution μ induces a Monge solution. This is not true, as our example in [20] demonstrates. In the multi-marginal problem, duality yields m first order conditions; our strategy in this paper is to show that if we fix the first variable, these equations can be uniquely solved for the other $m-1$ variables. In the problems considered by Gangbo and Święch [10] and Heinich [11], these equations turn out to have a particularly simple form and can be solved explicitly. For more general cost functions, this becomes a much more subtle issue. Our proof will combine a second order differential condition with tools from convex analysis and will require that the tensor T, defined below, is negative definite.

DEFINITION 2.6. *Suppose c is $(1, m)$-nondegenerate. Let $\vec{y} = (y_1, y_2, \dots, y_m) \in M_1 \times M_2 \times \cdots \times M_m$. For each $i := 2, 3, \dots, m-1$ choose a point $\vec{y}(i) = (y_1(i), y_2(i), \dots, y_m(i)) \in \overline{M_1} \times \overline{M_2} \times \cdots \times \overline{M_m}$ such that $y_i(i) = y_i$. Define the following bilinear maps on $T_{y_2}M_2 \times T_{y_3}M_3 \times \cdots \times T_{y_{m-1}}M_{m-1}$:*

$$S_{\vec{y}} = -\sum_{j=2}^{m-1} \sum_{\substack{i=2 \\ i \neq j}}^{m-1} D^2_{x_i x_j} c(\vec{y}) + \sum_{i,j=2}^{m-1} (D^2_{x_i x_m} c (D^2_{x_1 x_m} c)^{-1} D^2_{x_1 x_j} c)(\vec{y}),$$

$$H_{\vec{y}, \vec{y}(2), \vec{y}(3), \dots, \vec{y}(m-1)} = \sum_{i=2}^{m-1} (Hess_{x_i} c(\vec{y}(i)) - Hess_{x_i} c(\vec{y})),$$

$$T_{\vec{y}, \vec{y}(2), \vec{y}(3), \dots, \vec{y}(m-1)} = S_{\vec{y}} + H_{\vec{y}, \vec{y}(2), \vec{y}(3), \dots, \vec{y}(m-1)}.$$

REMARK 2.7. *Note that $D^2_{x_i x_j} c(x_1, x_2, \dots, x_m), Hess_{x_i} c(x_1, x_2, \dots, x_m)$, and the composition $\big(D_{x_i x_m} c (D^2_{x_1 x_m} c)^{-1} D^2_{x_1 x_j} c\big)(x_1, x_2, \dots, x_m)$ are actually bilinear maps on the spaces $T_{x_i}M_i \times T_{x_j}M_j$, $T_{x_i}M_i \times T_{x_i}M_i$, and $T_{x_i}M_i \times T_{x_j}M_j$, respectively, but we can extend them to maps on the product space $(T_{x_2}M_2 \times T_{x_3}M_3 \times \cdots \times T_{x_{m-1}}M_{m-1})^2$ by considering only the appropriate components of the tangent vectors.*

REMARK 2.8. *In the two marginal setting, it is well known that the twist condition cannot hold if M_1 is a compact manifold [1], [7]. Analogously, we show here that we cannot have $T < 0$ if any of the M_i are compact manifolds.*

First, note that taking $\vec{y}(i) = \vec{y}$ for all $i = 2, 3, \dots, m-1$ implies that

$$H_{\vec{y}, \vec{y}(2), \vec{y}(3), \dots, \vec{y}(m-1)} = 0,$$

and so $T_{\vec{y}, \vec{y}(2), \vec{y}(3), \dots, \vec{y}(m-1)} = S_{\vec{y}}$. The diagonal entries of $S_{\vec{y}}$ are

$$(D^2_{x_i x_m} c (D^2_{x_1 x_m} c)^{-1} D^2_{x_1 x_i} c)(\vec{y});$$

if $S_{\vec{y}} < 0$, we must have

$$(D^2_{x_i x_m} c (D^2_{x_1 x_m} c)^{-1} D^2_{x_1 x_i} c)(\vec{y}) < 0.$$

This implies that $D^2_{x_ix_m}c$ and $D^2_{x_1x_i}c$ must be invertible; i.e., c is (i,m) and $(1,i)$-nondegenerate.

We now show that, if M_i is compact, c cannot be (i,j)-nondegenerate for any j. In light of the preceding calculation, this will imply that, if $T<0$, no M_i can be compact. To see this, fix local coordinates at some $x_j\in M_j$. Fix x_k for all $k\neq i$ and consider the first coordinate of $D_{x_j}c(x_1,x_2,\dots,x_m)$, $\frac{\partial c}{\partial x^1_j}(x_1,x_2,\dots,x_m)$. If M_i is compact, the continuous mapping $x_i\mapsto\frac{\partial c}{\partial x^1_j}(x_1,x_2,\dots,x_m)$ has a global minimum. At this minimum, its differential vanishes. This differential is the first row of the matrix corresponding to $D^2_{x_jx_i}c(x_1,x_2,\dots,x_m)$ in local coordinates; it follows that this matrix cannot be invertible at that point, so c cannot be (i,j)-nondegenerate.

Though T looks complicated, it appears naturally in our argument. The condition $T<0$ is in one sense analogous to the twist and nondegeneracy conditions that are so important in the two marginal problem. Like the nondegeneracy condition, negativity of S is an inherently local property on $M_1\times M_2\times\cdots\times M_m$; under this condition, one can show that our system of equations is locally uniquely solvable. To show that the solution is actually globally unique requires something more; in the two marginal case, this is the twist condition, which can be seen as a global extension of nondegeneracy. In our setting, requiring that the sum $T=S+H<0$ turns out to be enough to ensure that the locally unique solution is in fact globally unique.

To prove our main theorem, we will also need a geometric condition on the domains. This condition is an analogue of convexity and is somewhat reminiscent of the notion of c-convexity of domains, introduced by Ma, Trudinger, and Wang [17]; more precisely, we will require that the set $Y^c_{x_1,p_1}\subseteq M_2\times M_3\times\cdots\times M_{m-1}$ defined below be geodesically convex. That is, whenever $(x_2,x_3,\dots,x_{m-1}),(\overline{x_2},\overline{x_3},\dots,\overline{x_{m-1}})\in Y^c_{x_1,p_1}$ and $\gamma_i(t)$ are geodesics joining x_i and $\overline{x_i}$, for $i=2,3,\dots,m-1$, we have $(\gamma_2(t),\gamma_3(t),\dots,\gamma_{m-1}(t))\in Y^c_{x_1,p_1}$ for all $t\in[0,1]$.

DEFINITION 2.9. *Let $x_1\in M_1$ and $p_1\in T^*_{x_1}M_1$. We define $Y^c_{x_1,p_1}\subseteq M_2\times M_3\times\cdots\times M_{m-1}$ by*

$$Y^c_{x_1,p_1}=\{(x_2,x_3,\dots,x_{m-1})|\ \exists\ x_m\in M_m\ \text{such that}\ D_{x_1}c(x_1,x_2,\dots,x_m)=p_1\}.$$

To build some intuition about this condition, we consider the costs of Gangbo and Święch [10] and Heinich [11]. In this case, convexity of the domains $M_2,M_3,\dots,M_m\subseteq\mathbb{R}^n$ implies geodesic convexity of the sets $Y^c_{x_1,p_1}$, as we demonstrate now.

PROPOSITION 2.10. *Let $M_i\subseteq\mathbb{R}^n$ for all i and $c(x_1,x_2,\dots,x_m)=h(x_1+x_2+\cdots+x_m)$ for a strictly convex $h\in C^1(\mathbb{R}^n)$. Then, if the domains M_i are convex for $i=2,3,\dots,m$, the sets $Y^c_{x_1,p_1}$ are all geodesically convex.*

Proof. Select $x_1\in M_1\subseteq\mathbb{R}^n$ and $p_1\subseteq T^*_{x_1}M_1\approx\mathbb{R}^n$. Choose $(x_2,x_3,\dots,x_{m-1})$, $(\overline{x_2},\overline{x_3},\dots,\overline{x_{m-1}})\in Y^c_{x_1,p_1}$; as we are working in $\mathbb{R}^n$, the geodesic joining x_i and $\overline{x_i}$ is simply the straight line $\gamma_i(t)=(\overline{x_i}-x_i)t+x_i$. We must show that $(\gamma_2(t),\gamma_3(t),\dots,\gamma_{m-1}(t))\in Y^c_{x_1,p_1}$ for all $t\in[0,1]$; that is, there exists some $x_m(t)\in M_m$ such that

$$D_{x_1}c(x_1,\gamma_2(t),\gamma_3(t),\dots,\gamma_{m-1}(t),x_m(t))=p_1.$$

Now, for each t, the equation

$$D_{x_1}c(x_1,\gamma_2(t),\gamma_3(t),\dots,\gamma_{m_1}(t),x_m(t))=Dh(x_1+\gamma_2(t)+\gamma_3(t)+\cdots+\gamma_{m_1}(t)+x_m(t))=p_1$$

has the unique solution

$$x_m(t) = Dh^*(p_1) - x_1 - \gamma_2(t) - \gamma_3(t) - \cdots - \gamma_{m-1}(t)$$
$$= Dh^*(p_1) - x_1 - \sum_{i=2}^{m-1}(\overline{x_i} - x_i)t + \sum_{i=2}^{m-1} x_i,$$

where h^* is the Legendre transform of h. Therefore, $x_m(t)$ is linear in t. Now, $x_m(0), x_m(1) \in M_m$ because $(x_2, x_3, \ldots, x_{m-1}), (\overline{x_2}, \overline{x_3}, \ldots, \overline{x_{m-1}}) \in Y^c_{x_1,p_1}$, and so $x_m(t) \in M_m$ by the convexity of M_m. ∎

Note that, assuming the marginals μ_i are all compactly supported, we are free to take each domain M_i to be a large ball containing $spt(\mu_i)$, and so the convexity assumption is not at all restrictive here. This illustrates a contrast between the role of c-geodesically convex domains here and the role of c-convexity in the regularity theory of Ma, Trudinger, and Wang; in our setting, we will not need to make any regularity assumptions on the measures μ_i for $i \geq 2$, and so we are free to replace a domain M_i with a larger domain, if necessary, to ensure that this condition holds. On the other hand, to prove regularity results in the two marginal setting requires a lower bound on the marginal μ_2. Replacing M_2 with a larger domain would violate this condition.

3. Monge solutions. We are now in a position to precisely state our main theorem.

THEOREM 3.1. *Suppose that*

I. *c is $(1, m)$-nondegenerate;*

II. *c is $(1, m)$-twisted;*

III. *for all choices of $\vec{y} = (y_1, y_2, \ldots, y_m) \in M_1 \times M_2 \times \cdots \times M_m$ and of $\vec{y}(i) = (y_1(i), y_2(i), \ldots, y_m(i)) \in \overline{M_1} \times \overline{M_2} \times \cdots \times \overline{M_m}$ such that $y_i(i) = y_i$ for $i = 2, \ldots, m-1$, we have*

$$T_{\vec{y},\vec{y}(2),\vec{y}(3),\ldots,\vec{y}(m-1)} < 0;$$

IV. *for all $x_1 \in M_1$ and $p_1 \in T^*_{x_1}M_1$, $Y^c_{x_1,p_1}$ is geodesically convex;*

V. *the first marginal μ_1 does not charge sets of Hausdorff dimension less than or equal to $n-1$. Then any solution μ to the Kantorovich problem is concentrated on the graph of a function; that is, there exist functions $G_i : M_1 \to M_i$ such that $graph(\vec{G}) = \{(x_1, G_2(x_1), G_3(x_1), \ldots, G_m(x_1))\}$ satisfies $\mu(graph(\vec{G})) = 1$*

Proof. Let u_i be a c-conjugate solution to the dual problem. Now u_1 is semi-concave and hence differentiable off a set of Hausdorff dimension $n-1$; as μ_1 vanishes on every set of Hausdorff dimension less than or equal to $n-1$, by Theorem 2.2 it suffices to show that for every $x_1 \in M_1$, where u_1 is differentiable, there is at most one $(x_2, x_3, \ldots, x_m) \in M_2 \times M_3 \times \cdots \times M_m$ such that $\sum_{i=1}^m u_i(x_i) = c(x_1, x_2, x_3, \ldots, x_m)$. Note that this equality implies that $D_{x_i}c(x_1, x_2, \ldots, x_m) \in \overline{\partial} u_i(x_i)$ for all $i = 1, 2, \ldots, m$; in particular, as u_1 is differentiable at x_1, $Du_1(x_1) = D_{x_1}c(x_1, x_2, \ldots, x_m)$. Our strategy will be to show that these inclusions can hold for at most one $(x_2, x_3, \ldots, x_m)$.

Fix a point x_1 where u_1 is differentiable. Twistedness implies that the equation $Du_1(x_1) = D_{x_1}c(x_1, x_2, \ldots, x_m) := p_1$ defines x_m as a function $x_m = F_{x_1}(x_2, \ldots, x_{m-1})$ of the variables $x_2, x_3, \ldots, x_{m-1}$ for $x_2, x_3, \ldots, x_{m-1} \in Y^c_{x_1,p_1}$; nondegeneracy and the implicit function theorem then imply that F_{x_1} is continuously differentiable with respect to $x_2, x_3, \ldots, x_{m-1}$ and

$$D_{x_i}F_{x_1}(x_2, \ldots, x_{m-1}) = -\big((D^2_{x_1x_m}c)^{-1}D^2_{x_1x_i}c\big)(x_1, x_2, \ldots, F_{x_1}(x_2, \ldots, x_{m-1})\big)$$

for $i = 2, \dots, m-1$. We will show there exists at most one point $(x_2, x_3, \dots, x_{m-1}) \in M_2 \times M_3 \times \cdots \times M_{m-1}$ such that

$$D_{x_i}c(x_1, x_2, \dots, F_{x_1}(x_2, \dots, x_{m-1})) \in \overline{\partial} u_i(x_i)$$

for all $i = 2, \dots, m-1$.

The proof is by contradiction; suppose there are two such points, $(x_2, x_3, \dots, x_{m-1})$ and $(\overline{x_2}, \overline{x_3}, \dots, \overline{x_{m-1}})$. For $i = 2, \dots, m-1$, we can choose Riemannian geodesics $\gamma_i(t)$ in M_i such that $\gamma_i(0) = x_i$ and $\gamma_i(1) = \overline{x_i}$. By the geodesic convexity of $Y^c_{x_1,p_1}$, we have $(\gamma_2(t), \gamma_3(t), \dots, \gamma_{m-1}(t)) \in Y^c_{x_1,p_1}$, and so $F_{x_1}(\gamma_2(t), \gamma_3(t), \dots, \gamma_{m-1}(t))$ is well defined. Take a measurable selection of covectors $V_i(t) \in \partial u_i(\gamma_i(t))$. We will show that $f(1) < f(0)$, where

$$f(t) := \sum_{i=2}^{m-1} [V_i(t) - D_{x_i}c(x_1, \vec{\gamma}(t))]\left\langle \frac{d\gamma_i}{dt} \right\rangle,$$

and we use $(x_1, \vec{\gamma}(t))$ as shorthand for $(x_1, \gamma_2(t), \dots, \gamma_{m-1}(t), F_{x_1}(\gamma_2(t), \dots, \gamma_{m-1}(t)))$ and $a\langle b\rangle$ to denote the duality pairing between a 1-form a and a vector b. This will clearly imply the desired result.

For each t and each $i = 2, \dots, m-1$, by c-conjugacy of u_i and the compactness of $\overline{M_j}$ for $j \neq i$, we can choose $\vec{y}(i;t) = (y_1(i;t), y_2(i;t), \dots, y_m(i:t)) \in \overline{M_1} \times \overline{M_2} \times \cdots \times \overline{M_m}$ so that $y_i(i;t) = \gamma_i(t)$ and

$$\sum_{j=1}^{m} u_j(y_j(i;t)) = c(y_1(i;t), y_2(i;t), \dots, y_m(i;t)).$$

Note that $V_i(t)\langle \frac{d\gamma_i}{dt}\rangle$ supports the semiconcave function $T \in [0,1] \mapsto u_i(\gamma_i(t))$. But $u_i(\gamma_i(t))$ is twice differentiable almost everywhere, and hence we have $V_i(t)\langle \frac{d\gamma_i}{dt}\rangle = \frac{d(u_i(\gamma_i(t)))}{dt}$ for almost all t and, by semiconcavity, $V_i(1)\langle \frac{d\gamma_i}{dt}\rangle - V_i(0)\langle \frac{d\gamma_i}{dt}\rangle \leq \int_0^1 \frac{d^2(u_i(\gamma_i(t)))}{dt^2}dt$. Now, for any $t, s \in [0,1]$,

$$u_i(\gamma_i(t)) \leq c(y_1(i;s), y_2(i;s), \dots, y_{i-1}(i;s), \gamma_i(t), y_{i+1}(i;s) \dots, y_m(i;s)) - \sum_{j \neq i} u_j(y_j(i;s))$$

and we have equality when $t = s$, as $\gamma_i(s) = y_i(i;s)$. Hence, whenever $\frac{d^2(u_i(\gamma_i(t)))}{dt^2}$ exists, we have

$$\begin{aligned} \left.\frac{d^2(u_i(\gamma_i(t)))}{dt^2}\right|_{t=s} &\leq \left.\frac{d^2(c(y_1(i;s), y_2(i;s), \dots, y_{i-1}(i;s), \gamma_i(t), y_{i+1}(i;s) \dots, y_m(i;s)))}{dt^2}\right|_{t=s} \\ &= Hess_{x_i}c(y_1(i;s), y_2(i;s), \dots, y_m(i;s))\left\langle \frac{d\gamma_i}{ds}, \frac{d\gamma_i}{ds} \right\rangle. \end{aligned}$$

We conclude that

(3)

$$V_i(1)\left\langle \frac{d\gamma_i}{dt} \right\rangle - V_i(0)\left\langle \frac{d\gamma_i}{dt} \right\rangle \leq \int_0^1 Hess_{x_i}c(y_1(i;t), y_2(i;t), \dots, y_m(i;t))\left\langle \frac{d\gamma_i}{dt}, \frac{d\gamma_i}{dt} \right\rangle dt.$$

Turning now to the other term in $f(1)-f(0)$, we have

$$\begin{aligned}
&D_{x_i}c(x_1,\vec{\gamma(1)})\Big\langle\frac{d\gamma_i}{dt}\Big\rangle - D_{x_i}c(x_1,\vec{\gamma(0)})\Big\langle\frac{d\gamma_i}{dt}\Big\rangle\\
&=\int_0^1 \frac{d}{dt}\Big(D_{x_i}c(x_1,\vec{\gamma}(t))\Big\langle\frac{d\gamma_i}{dt}\Big\rangle\Big)dt\\
&=\int_0^1 \Bigg(\sum_{\substack{j=2\\ j\neq i}}^{m-1}\Big(D^2_{x_ix_j}c(x_1,\vec{\gamma}(t))\Big)\Big\langle\frac{d\gamma_i}{dt},\frac{d\gamma_j}{dt}\Big\rangle \;+Hess_{x_i}c(x_1,\vec{\gamma}(t))\Big\langle\frac{d\gamma_i}{dt},\frac{d\gamma_i}{dt}\Big\rangle\\
&\quad+\sum_{j=2}^{m-1}\Big(D^2_{x_ix_m}c(x_1,\vec{\gamma}(t))D_{x_j}F_{x_1}(\vec{\gamma}(t))\Big)\Big\langle\frac{d\gamma_i}{dt},\frac{d\gamma_j}{dt}\Big\rangle\Bigg)dt\\
&=\int_0^1 \Bigg(\sum_{\substack{j=2\\ j\neq i}}^{m-1}\Big(D^2_{x_ix_j}c(x_1,\vec{\gamma}(t))\Big)\Big\langle\frac{d\gamma_i}{dt},\frac{d\gamma_j}{dt}\Big\rangle +Hess_{x_i}c(x_1,\vec{\gamma}(t))\Big\langle\frac{d\gamma_i}{dt},\frac{d\gamma_i}{dt}\Big\rangle\\
&\quad-\sum_{j=2}^{m-1}\Big((D^2_{x_ix_m}c(D^2_{x_1x_m}c)^{-1}D^2_{x_1x_j}c)(x_1,\vec{\gamma}(t))\Big)\Big\langle\frac{d\gamma_i}{dt},\frac{d\gamma_j}{dt}\Big\rangle\Bigg)dt. \qquad (4)
\end{aligned}$$

Combining (3) and (4) yields

$$\begin{aligned}
f(1)-f(0)&\leq \int_0^1 T_{(x_1,\vec{\gamma}(t)),\vec{y}(2;t),\vec{y}(3;t),\dots,\vec{y}(m-1;t)}\Big\langle\frac{d\vec{\gamma}}{dt},\frac{d\vec{\gamma}}{dt}\Big\rangle dt\\
&<0. \qquad \square
\end{aligned}$$

REMARK 3.2. *The assumption $T<0$ can be weakened somewhat. In our proof, we argue that a certain function $f(t)$ is decreasing, essentially by showing that its derivative is bounded above by $T=S+H<0$. f is the contraction of certain covectors with the direction vectors of the geodesics $\gamma_i(t)$, and the Hessian terms H arise as unmixed second derivatives of c along this path. It is possible to construct similar arguments by using a class of curves other than geodesics.*

Instead of geodesic convexity of the sets $Y^c_{x_1,p_1}$ and negativity of T, we could assume the existence of a family of smooth curves $\mathcal{C}$ on $M_2\times M_3\times\cdots\times M_{m-1}$ satisfying the two conditions below; the proof is directly analogous to the proof of Theorem 3.1.

1. *For every pair of points $(x_2,x_3,\dots,x_{m-1}),(\overline{x_2},\overline{x_3},\dots,\overline{x_{m-1}})\in M_2\times M_3\times\cdots\times M_{m-1}$ there is a curve $\vec{\gamma(\cdot)}=(\gamma_2(\cdot),\gamma_3(\cdot),\dots,\gamma_{m-1}(\cdot))\in\mathcal{C}$ connecting these two points, with the following property: whenever we have $x_1\in M_1,p_1\in T^*_{x_1}M_1$, and $x_m,\overline{x_m}\in M_m$ such that*

$$D_{x_1}c(x_1,x_2,x_3,\dots,x_m)=p_1=D_{x_1}c(x_1,\overline{x_2},\overline{x_3},\dots,\overline{x_m}),$$

   then, for all $t\in[0,1]$, there is some $x_m(t)\in M_m$ such that

$$D_{x_1}c(x_1,\gamma_2(t),\gamma_3(t),\dots,x_m(t))=p_1.$$

2. *For all points $\vec{y}=(y_1,y_2,\dots,y_m)\in M_1\times M_2\times\cdots\times M_m$, $\vec{y}(i)=(y_1(i),y_2(i),\dots,y_m(i))\in\overline{M_1}\times\overline{M_2}\times\cdots\times\overline{M_m}$ such that $y_i(i)=y_i$ for $i=2,3,\dots,m-1$*

and curves $(\gamma_2(\cdot), \gamma_3(\cdot), \dots, \gamma_{m-1}(\cdot))$ *in* $\mathcal{C}$ *passing through* $(y_2, y_3, \dots, y_{m-1})$, *we have*

$$S_{\vec{y}}\left\langle \frac{d\vec{\gamma}}{dt}, \frac{d\vec{\gamma}}{dt} \right\rangle - \sum_{i=2}^{m-1} \frac{\partial^2 c}{\partial t^2}(y_1, y_2, \dots, y_{i-1}, \gamma_i(t), y_{i+1}, \dots, y_m)$$
$$+ \sum_{i=2}^{m-1} \frac{\partial^2 c}{\partial t^2}(y_1(i), y_2(i), , \dots, y_{i-1}(i), \gamma_i(t), y_{i+1}(i), \dots, y_m(i))$$
$$< 0.$$

Corollary 3.3. *Under the same conditions as Theorem* 3.1*, the Monge problem* (M) *admits a unique solution and the solution to the Kantorovich problem* (K) *is unique.*

Proof. We first show that the G_i defined in Theorem 3.1 push μ_1 to μ_i for all $i = 2, 3, \dots, m$. Pick a Borel set $B \subseteq M_i$. We have

$$\begin{aligned}
\mu_i(B) &= \mu(M_1 \times M_2 \times \dots \times M_{i-1} \times B \times M_{i+1} \times \dots \times M_m) \\
&= \mu\Big((M_1 \times M_2 \times \dots \times M_{i-1} \times B \times M_{i+1} \times \dots \times M_m) \cap \text{graph}(\vec{G})\Big) \\
&= \mu\Big(\{(x_1, G_2(x_1), \dots, G_m(x_1)) | G_i(x_1) \in B\}\Big) \\
&= \mu\Big((G_i^{-1}(B) \times M_2 \times \dots \times M_m) \cap \text{graph}(\vec{G})\Big) \\
&= \mu(G_i^{-1}(B) \times M_2 \times \dots \times M_m) \\
&= \mu_1(G_i^{-1}(B)).
\end{aligned}$$

This implies that $(G_2, G_3, \dots, G_m)$ solves (M). To prove uniqueness of μ, note that any other optimizer $\overline{\mu}$ must also be concentrated on graph$(\vec{G})$, which in turn implies $\overline{\mu} = (\text{Id}, G_2, \dots, G_m)_{\#}\mu_1 = \mu$. Uniqueness of $(G_2, G_3, \dots, G_m)$ now follows immediately; if $(\overline{G_2}, \overline{G_3}, \dots, \overline{G_m})$ is another solution to (M), then $(\text{Id}, \overline{G_2}, \overline{G_3}, \dots, \overline{G_m})_{\#}\mu_1$ is another solution to (K), which must then be concentrated on graph$(\vec{G})$. This means that $G_i = \overline{G}_i$, μ_1 almost everywhere. ∎

4. Examples. In this section, we discuss several types of cost functions to which Theorem 3.1 applies. In these examples, the complicated tensor T simplifies considerably.

In the first subsection, we consider certain classes of cost functions whose form is very similar to those found in either Gangbo and Święch [10] and Heinich [11]. In the second subsection, we consider a type of cost function arising in a hedonic pricing model.

4.1. Generalizations of the costs of Gangbo and Święch and Heinich. In this subsection, we consider three classes of cost functions which directly generalize, in different ways, those found in [10] and [11].

Example 4.1 (perturbations of concave functions of the sum). *Recall that Gangbo and Święch and Heinich treated cost functions defined on* $(\mathbb{R}^n)^m$ *by* $c(x_1, x_2, \dots, x_m) = h(\sum_{k=1}^m x_k)$*, where* $h : \mathbb{R}^n \to \mathbb{R}$ *is strictly concave. Here, we make the slightly stronger assumption that* h *is* C^2 *with* $D^2 h < 0$*. Assuming each* μ_i *is compactly supported, we can take* M_i *to be a bounded, convex domain in* $\mathbb{R}^n$ *for* $i = 2, 3, \dots, m$*, in which case the sets* $Y^c_{x_1,p_1}$ *are all geodesically convex, by Proposition* 2.10*. Now,* $D_{x_i} c(x_1, x_2, \dots, x_m) = Dh(\sum_{k=1}^m x_k)$ *and* $D^2_{x_i x_j} c(x_1, x_2, \dots, x_m) = D^2 h(\sum_{k=1}^m x_k)$, *where we have made the obvious identification between tangent spaces at different points. c is then clearly $(1,m)$-twisted and $(1,m)$-nondegenerate. Furthermore, the bilinear map $S_{\vec{y}}$ on $(\mathbb{R}^n)^{m-2}$ is block diagonal, and each of its diagonal blocks is*

$$D^2h\left(\sum_{k=1}^{m} y_k\right).$$

Similarly, as $Hess_{x_i}c(\vec{y}(i)) = D^2h(\sum_{k=1}^{m} y_k(i))$ and $Hess_{x_i}c(\vec{y}) = D^2h(\sum_{k=1}^{m} y_k)$, $H_{\vec{y},\vec{y}(2),\vec{y}(3),\dots,\vec{y}(m-1)}$ is block diagonal, and its ith diagonal block is

$$D^2h\left(\sum_{k=1}^{m} y_k(i)\right) - D^2h\left(\sum_{k=1}^{m} y_k\right).$$

Therefore, $T_{\vec{y},\vec{y}(2),\vec{y}(3),\dots,\vec{y}(m-1)}$ is block diagonal, and its ith diagonal block is

$$D^2h\left(\sum_{k=1}^{m} y_k(i)\right).$$

This is clearly negative definite. Furthermore, C^2 perturbations of this cost function will also satisfy $T_{\vec{y},\vec{y}(2),\vec{y}(3),\dots,\vec{y}(m-1)} < 0$; this shows that the results of Gangbo and Święch and Heinich are robust with respect to perturbations of the cost function.

EXAMPLE 4.2 (bilinear costs). *We now turn to bilinear costs; suppose $c : (\mathbb{R}^n)^m \to \mathbb{R}$ is given by $c(x_1, x_2, \dots, x_m) = \sum_{i\neq j}(x_i)^T A_{ij} x_j$ for n by n matrices A_{ij}. If A_{1m} is nonsingular, c is $(1,m)$-twisted and $(1,m)$-nondegenerate, and, if the domains $M_2, M_3, \dots, M_m$ are all convex, the sets $Y^c_{x_1,p_1}$ are all geodesically convex, by an argument similar to the proof of Proposition* 2.10. *Now, the Hessian terms in T vanish and so the condition $T < 0$ becomes a condition on the A_{ij}. For example, when $m = 3$, we have $T = A_{21}(A_{31})^{-1}A_{32}$; $T < 0$ is the same condition that ensures the solution to* (K) *is contained in an n-dimensional submanifold in* [20] *and* [21].

Note that after changing coordinates in x_2 and x_3, we can assume any bilinear three marginal cost is of the form

$$c(x_1, x_2, x_3) = x_1 \cdot x_2 + x_1 \cdot x_3 + x_2^T A x_3.$$

In these coordinates, the threefold product $A_{21}(A_{31})^{-1}A_{32} = A^T$. Applying the linear change of coordinates

$$\begin{aligned} x_1 &\mapsto U_1x_1,\\ x_2 &\mapsto U_2x_2,\\ x_3 &\mapsto U_3x_3 \end{aligned}$$

yields

$$c(x_1, x_2, x_3) = x_1^T U_1^T U_2 x_2 + x_1^T U_1^T U_3 x_3 + x_2^T U_2^T A U_3 x_3.$$

If A is negative definite and symmetric, then we can choose $U_3 = U_2$ such that $U_2^T A U_3 = -I$ and $U_1 = -(U_2^T)^{-1}$ to obtain

$$c(x_1, x_2, x_3) = -x_1^T x_2 - x_1^T x_3 - x_2^T x_3,$$

which is equivalent[2] *to the cost of Gangbo and Święch. As the symmetry of* $D^2_{x_2x_1}c(D^2_{x_3x_1}c)^{-1}D^2_{x_3x_2}c$ *is independent of our choice of coordinates, we conclude that* c *is equivalent to Gangbo and Święch's cost if and only if* $A_{21}(A_{31})^{-1}A_{32}$ *is symmetric and negative definite. Thus, when* $m = 3$*, our result restricted to bilinear costs generalizes Gangbo and Święch's theorem from costs for which* $A_{21}(A_{31})^{-1}A_{32}$ *is symmetric and negative definite to costs for which it is only negative definite.*

EXAMPLE 4.3. *There is another class of three marginal problems to which Theorem* 3.1 *applies: on* $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$*, set*

$$c(x_1, x_2, x_3) = g(x_1, x_3) + \frac{|x_1 - x_2|^2}{2} + \frac{|x_3 - x_2|^2}{2}.$$

If $g(x_1, x_3) = \frac{|x_1 - x_3|^2}{2}$*, this is equivalent to the cost of Gangbo and Święch. More generally, if* g *is* $(1, 3)$*-twisted and nondegenerate, then* c *is as well. Moreover, if we make the usual identification between tangent spaces at different points in* $\mathbb{R}^n$*, we have*

$$T_{\vec{y},\vec{y}(2)} = \left(D^2_{x_1x_3}g(y_1, y_3)\right)^{-1}.$$

Hence, if $D^2_{x_1x_3}g(y_1, y_3) < 0$*, we have* $T_{\vec{y},\vec{y}(2)} < 0$*. This will be the case if, for example,* $g(x_1, x_3) = h(x_1 - x_3)$ *for* h *uniformly convex or* $g(x_1, x_3) = h(x_1 + x_3)$ *for* h *uniformly concave.*

It is also straightforward to show that, if M_2 *is convex, geodesic convexity of* $Y^c_{x_1,p_1}$ *follows from ordinary convexity of* $D_{x_1}g(x_1, M_3)$*. This latter condition is well known in the regularity theory of optimal transportation with two marginals* [17]*; if it holds for all* x_1*,* M_3 *is said to be g-convex.*

4.2. Hedonic pricing costs. Chiappori, McCann, and Nesheim [7] and Carlier and Ekeland [5] showed that finding equilibrium in a certain hedonic pricing model is equivalent to solving a multimarginal optimal transportation problem with a cost function of the form

$$c(x_1, x_2, \ldots, x_m) = \inf_{z \in Z} \sum_{i=1}^{m} f_i(x_i, z).$$

Let us assume the following:

1. Z is a C^2 smooth n-dimensional manifold.
2. For all k, f_k is C^2 and nondegenerate.
3. For each $(x_1, x_2, \ldots, x_m)$ the infimum is attained by a unique $z(x_1, x_2, \ldots, x_m) \in Z$.
4. $\sum_{k=1}^{m} D^2_{zz}f_k(x_k, z(x_1, x_2, \ldots, x_m))$ is nonsingular.

In [20] and [21], we showed that these conditions implied that c is C^2 and (i, j)-nondegenerate for all $i \neq j$; we then showed that the support of any optimizer is contained in an n-dimensional Lipschitz submanifold of the product $M_1 \times M_2 \times \cdots \times M_m$. Our aim is to examine conditions on the f_i under which the other hypotheses of Theorem 3.1 are satisfied. First, we recall that Theorem 3.1 requires $(1, m)$-twistedness. Below, we show that the following conditions ensure that c is (i, j)-twisted:

5. f_i is (x_i, z)-twisted (that is, $z \mapsto D_{x_i}f_i(x_i, z)$ is injective).

[2] We say cost functions c and $\overline{c}$ are equivalent if $\overline{c}(x_1, x_2, \ldots, x_m) = c(x_1, x_2, \ldots, x_m) + \sum_{i=1}^{m} g_i(x_i)$. As the effect of the g_i's is to shift the functionals $C(G_2, G_3, \ldots, G_m)$ and $C(\mu)$ by the constant $\sum_{i=1}^{m} \int_{M_i} g_i(x_i)d\mu_i$, studying c is essentially equivalent to studying $\overline{c}$.

6. f_j is (z,x_j)-twisted (that is, $x_j \mapsto D_z f_j(x_j,z)$ is injective).

PROPOSITION 4.2.1. *Assume conditions* 1–4 *and that* 5 *and* 6 *hold for a fixed* $i \neq j$. *Then* c *is* (i,j)*-twisted.*

Proof. Note that $c(x_1,x_2,\dots,x_m) \leq \sum_{i=1}^m f_i(x_i,z)$ with equality when $z = z(x_1,x_2,\dots,x_m)$; therefore,

$$D_{x_i}c(x_1,x_2,\dots,x_m) = D_{x_i}f(x_i,z(x_1,x_2,\dots,x_m)). \tag{5}$$

Therefore, for fixed x_k for all $k \neq j$, the map $x_j \mapsto D_{x_i}c(x_1,x_2,\dots,x_m)$ is the composition of the maps $x_j \mapsto z(x_1,x_2,\dots,x_m)$ and $z \mapsto D_{x_i}f(x_i,z)$. The latter map is injective by assumption. Now, note that

$$\sum_{k=1}^m D_z f_k(x_k,z(x_1,x_2,\dots,x_m)) = 0;$$

hence,

$$D_z f_j(x_j,z(x_1,x_2,\dots,x_m)) = -\sum_{k\neq j} D_z f_k(x_k,z(x_1,x_2,\dots,x_m)).$$

Twistedness of f_j now immediately implies injectivity of the first map. □

We now investigate the form of the tensor T. For convenience, define

$$A(x_1,x_2,\dots,x_m) := \sum_{i=1}^m D^2_{zz}f_i(x_i,z(x_1,x_2,\dots,x_m)).$$

PROPOSITION 4.2.2. *Under conditions* 1–4, $T_{\vec{y},\vec{y}(2),\vec{y}(3),\dots,\vec{y}(m-1)}$ *is block diagonal, and its ith block is*

$$-\Big[(D_{x_iz}f_i)A^{-1}(D^2_{zx_i}f_i)\Big]\Big(\vec{y}(i),z\big(\vec{y}(i)\big)\Big)+Hess_{x_i}f_i\Big(y_i,z\big(\vec{y}(i)\big)\Big)-Hess_{x_i}f_i\Big(y_i,z(\vec{y})\Big). \tag{6}$$

Proof. As $A(x_1,x_2,\dots,x_m)$ is nonsingular by assumption, the implicit function theorem implies that $z(x_1,x_2,\dots,x_m)$ is differentiable and

$$D_{x_i}z(x_1,x_2,\dots,x_m) = -\big(A(x_1,x_2,\dots,x_m)\big)^{-1}D^2_{zx_i}f_i(x_i,z(x_1,x_2,\dots,x_m)).$$

Furthermore, note that as A is positive semidefinite by the minimality of $z \mapsto \sum_{i=1}^m f_i(x_i,z)$ at $z(x_1,x_2,\dots,x_m)$, the nonsingular assumption implies that it is in fact positive definite.

Differentiating (5) with respect to x_i for $i = 2,3,\dots,m-1$ yields

$$Hess_{x_i}c = -(D^2_{x_iz}f_i)D_{x_i}z + Hess_{x_i}f_i = -(D^2_{x_iz}f_i)A^{-1}(D^2_{zx_i}f_i) + Hess_{x_i}f_i,$$

where we have suppressed the arguments $x_1,x_2,\dots,x_m$ and $z(x_1,x_2,\dots,x_m)$. A similar calculation yields, for all $i \neq j$,

$$D^2_{x_ix_j}c = (D^2_{x_iz}f_i)D_{x_j}z = -(D^2_{x_iz}f_i)A^{-1}(D^2_{zx_j}f_i).$$

Thus, for all $i \neq j$, a straightforward calculation yields

$$D^2_{x_ix_m}c(D^2_{x_1x_m}c)^{-1}D^2_{x_1x_j}c = -(D^2_{x_iz}f_i)A^{-1}(D^2_{zx_j}f_i) = D^2_{x_ix_j}c.$$

Hence, $S_{\vec{y}}$ is block diagonal. Furthermore, another simple calculation implies that its ith diagonal block is

$$\Big[D^2_{x_i x_m} c (D^2_{x_1 x_m} c)^{-1} D^2_{x_1 x_i} c\Big]\Big(\vec{y}\Big) = -\Big[(D^2_{x_i z} f_i) A^{-1} (D^2_{z x_i} f_i)\Big]\Big(\vec{y}, z(\vec{y})\Big). \tag{7}$$

In addition, $H_{\vec{y},\vec{y}(2),\vec{y}(3),\dots,\vec{y}(m-1)}$ is block diagonal, and its ith block is

$$\begin{aligned} -\Big[(D^2_{x_i z} f_i) A^{-1} (D^2_{z x_i} f_i)\Big]\Big(\vec{y}(i), z\big(\vec{y}(i)\big)\Big) + Hess_{x_i} f_i\Big(y_i, z\big(\vec{y}(i)\big)\Big) \\ + \Big[(D^2_{x_i z} f_i) A^{-1} (D^2_{z x_i} f_i)\Big]\Big(\vec{y}, z(\vec{y})\Big) - Hess_{x_i} f_i\Big(y_i, z(\vec{y})\Big). \end{aligned} \tag{8}$$

The result now follows by combining (7) and (8). ∎

Therefore, $T_{\vec{y},\vec{y}(2),\vec{y}(3),\dots,\vec{y}(m-1)}$ is negative definite if and only if each of its diagonal blocks is. Now, A is symmetric and positive definite; therefore A^{-1} is as well. The first term in the ith block of (6) is therefore negative definite; the entire block will be negative definite if this term dominates the difference of the Hessian terms. This is the case if, for example, $M_i = \mathbb{R}^n$ and f_i takes the form

$$f_i(x_i, z) = \alpha_i(z) \cdot x_i + \beta_i(x_i) + \lambda_i(z) \tag{9}$$

for all $i = 2, 3, \dots, m-1$, in which case $Hess_{x_i} f_i\big(y_i, z\big(\vec{y}(i)\big)\big) = Hess_{x_i} f_i\big(y_i, z(\vec{y})\big)$.

Finally, we consider the geodesic convexity of $Y^c_{x_1,p_1}$. For simplicity, we will consider only the simple form (9) for the functions f_i for $i = 2, 3, \dots, m-1$.

Proposition 4.2.3. *Assume conditions* 1–4*, and that* 5 *and* 6 *hold for* $i = 1, j = m$. *Suppose that* f_i *takes the form* (9) *for* $i = 2, 3, \dots, m-1$ *and that the domians* $M_2, M_3, \dots, M_{m-1} \subseteq \mathbb{R}^n$ *are convex. Fix* $x_1 \in M_1$ *and* $p_1 \in T^*_{x_1} M_1$ *and suppose there exists a (necessarily unique by condition* 5*)* $z \in Z$ *such that* $D_{x_1} f_1(x_1, z) = p_1$. *Then, if the set* $D_z f_m(M_m, z)$ *is convex,* $Y^c_{x_1,p_1}$ *is geodesically convex.*

One should note that, if there is no $z \in Z$ such that $D_{x_1} f_1(x_1, z) = p_1$, then by (5), $Y^c_{x_1,p_1}$ is empty. Therefore, in the terminology of [17], if M_m is f_m-convex, then $Y^c_{x_1,p_1}$ is geodesically convex for all x_1, p_1.

Proof. Fix $(x_2, x_3, \dots, x_{m-1})$ and $(\overline{x_2}, \overline{x_3}, \dots, \overline{x_{m-1}})$ in $Y^c_{x_1,p_1}$; there exist $x_m, \overline{x_m}$ such that

$$D_{x_1} c(x_1, x_2, x_3, \dots, x_{m-1}, x_m) = p_1 = D_{x_1} f_1(x_1, z) = D_{x_1} c(x_1, \overline{x_2}, \overline{x_3}, \dots, \overline{x_{m-1}}, \overline{x_m}).$$

By (5) and (x_1, z)-twistedness of f_1, we have

$$z(x_1, x_2, \dots, x_{m-1}, x_m) = z = z(x_1, \overline{x_2}, \overline{x_3}, \dots, \overline{x_{m-1}}, \overline{x_m}),$$

and so

$$D_z f_1(x_1, z) + \sum_{i=2}^{m} D_z f_i(x_i, z) = 0 = D_z f_1(x_1, z) + \sum_{i=2}^{m} D_z f_i(\overline{x_i}, z). \tag{10}$$

As $M_i = \mathbb{R}^n$ for $i = 2, 3, \dots, m-1$, the geodesics joining x_i and $\overline{x_i}$ are straight lines; that is, $\gamma_i(t) := (\overline{x_i} - x_i)t + x_i$.

Now, consider the equation

$$D_z f_1(x_1, z) + \sum_{i=2}^{m-1} D_z f_i(\gamma_i(t), z) + D_z f_m(x_m(t), z) = 0.$$

Using (9), this reduces to

$$D_z f_m(x_m(t), z) = -D_z f_1(x_1, z) - \sum_{i=2}^{m-1} [D\alpha_i(z) \cdot \gamma_i(t) - D\lambda_i(z)]. \tag{11}$$

Now, (10) implies that (11) has a (unique by the (z, x_m)-twistedness of f_m) solution $x_m(t) \in M_m$ for $t = 0$ and $t = 1$; namely, $x_m(0) = x_m$ and $x_m(1) = \overline{x_m}$. By the linearity of the $\gamma_i(t)$'s and the convexity of $D_z f_m(M_m, z)$, we conclude that there is a unique solution $x_m(t) \in M_m$ to (11) for each $t \in [0, 1]$. Now, (11) implies $z = z(x_1, \gamma_2(t), \gamma_3(t), \ldots, \gamma_{m-1}(t), x_m(t))$ so that, by (5),

$$D_{x_1} c(x_1, \gamma_2(t), \gamma_3(t), \ldots, \gamma_{m-1}(t), x_m(t)) = D_{x_1} f_1(x_1, z) = p_1.$$

Therefore, $(\gamma_2(t), \gamma_3(t), \ldots, \gamma_{m-1}(t)) \in Y^c_{x_1,p_1}$, as desired. ◻

Acknowledgment. It is a pleasure to thank my advisor, Robert McCann, for his support and for many useful discussions throughout the duration of this work.

REFERENCES

[1] N. Ahmad, H. K. Kim, and R. J. McCann, *Optimal transportation, topology and uniqueness,* Bull. Math. Sci., 1 (2011), pp. 13–32.

[2] Y. Brenier, *Decomposition polaire et rearrangement monotone des champs de vecteurs*, C. R. Acad. Sci. Paris Ser. I Math., 305 (1987), pp. 805–808.

[3] L. Caffarelli, *Allocation maps with general cost functions*, in Partial Differential Equations and Applications, P. Marcellini, G. Talenti, and E. Vesintini, eds., Lecture Notes in Pure and Appl. Math. 177, Dekker, New York, 1996, pp. 29–35.

[4] G. Carlier, *On a class of multidimensional optimal transportation problems*, J. Convex Anal., 10 (2003), pp. 517–529.

[5] G. Carlier and I. Ekeland, *Matching for teams*, Econom. Theory, 42 (2010), pp. 397–418.

[6] G. Carlier and B. Nazaret, *Optimal transportation for the determinant*, ESAIM Control Optim. Calc. Var., 14 (2008), pp. 678–698.

[7] P.-A. Chiappori, R. McCann, and L. Nesheim, *Hedonic price equilibria, stable matching, and optimal transport: Equivalence, topology and uniqueness*, Econom. Theory, 42 (2010), pp. 317–354.

[8] W. Gangbo, *Quelques problemes d'analyse nonconvex,* Habilitation thesis, Universite de Metz, Metz, France, 1995.

[9] W. Gangbo and R. J. McCann, *The geometry of optimal transportation*, Acta Math., 177 (1996), pp. 113–161.

[10] W. Gangbo and A. Święch, *Optimal maps for the multidimensional Monge-Kantorovich problem*, Comm. Pure Appl. Math., 51 (1998), pp. 23–45.

[11] H. Heinich, *Problème de Monge pour n probabilitiés*, C. R. Math. Acad. Sci. Paris, 334 (2002), pp. 793–795.

[12] H. G. Kellerer, *Duality theorems for marginal problems*, Z. Wahrsch. Verw. Gebiete, 67 (1984), pp. 399–432.

[13] Y.-H. Kim and R. McCann, *Continuity, curvature, and the general covariance of optimal transportation*, J. Eur. Math. Soc. (JEMS), 12 (2010), pp. 1009–1040.

[14] M. Knott and C. Smith, *On a generalization of cyclic monotonicity and distances among random vectors*, Linear Algebra Appl., 199 (1994), pp. 363–371.

[15] V. Levin, *Abstract cyclical monotonicity and Monge solutions for the general Monge-Kantorovich problem*, Set-Valued Anal., 7 (1999), pp. 7–32.

[16] G. Loeper, *On the regularity of maps solutions of optimal transportation problems*, Acta Math., 202 (2009), pp. 241–283.

[17] X.-N. Ma, N. Trudinger, and X.-J. Wang, *Regularity of potential functions of the optimal transportation problem*, Arch. Ration. Mech. Anal., 177 (2005), pp. 151–183.

[18] R. McCann, *Polar factorization of maps on Riemannian manifolds*, Geom. Funct. Anal., 11 (2001), pp. 589–608.

[19] I. Olkin and S. T. Rachev, *Maximum submatrix traces for positive definite matrices*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 390–397.

[20] B. Pass, *On the local structure of optimal measures in the multi-marginal optimal transportation problem*, Calc. Var. Partial Differential Equations, to appear. (Published online March 30, 2011. DOI: 10.1007/s00526-011-0421-z.)

[21] B. Pass, *Structural Results on Optimal Transportation Plans*, Ph.D. thesis, University of Toronto, Toronto, Ontario, 2011.

[22] L. Rüschendorf and L. Uckelmann, *On optimal multivariate couplings*, in Proceedings of the Prague 1996 Conference on Marginal Problems, V. Benes and I. Stepan, eds., Kluwer, Dordrecht, The Netherlands, 1997, pp. 261–274.

[23] L. Rüschendorf and L. Uckelmann, *On the n-coupling problem*, J. Multivariate Anal., 81 (2002), pp. 242–258.

[24] C. Villani, *Topics in Optimal Transportation*, Grad. Stud. Math. 58, American Mathematical Society, Providence, RI, 2003.

[25] C. Villani, *Optimal Transport: Old and New*, Grundlehren Math. Wiss. 338, Springer-Verlag, Berlin, 2009.