SIAM J. CONTROL AND OPTIMIZATION
Vol. 32, No. 3, pp. 673–689, May 1994

005

SOLUTION OF SOME TRANSPORTATION PROBLEMS WITH RELAXED OR ADDITIONAL CONSTRAINTS*

S. T. RACHEV† AND L. RÜSCHENDORF‡

Abstract. The authors consider some modifications of the usual transportation problem by allowing bounds for the admissible supply—respectively, demand—distributions. In particular, the case that the marginal distribution function of the supply is bounded below by a *df* F_1, while the marginal *df* of the demand is bounded above by a *df* is considered. For the case that the difference of the marginals is fixed—this is an extension of the well-known Kantorovich–Rubinstein problem—the authors obtain new and general explicit results and bounds, even without the assumption that the cost function is of Monge type. The multivariate case is also treated. In the last section, the authors study Monge–Kantorovich problems with constraints of a local type, that is, on the densities of the marginals. In particular, the classical Dobrushin theorem on optimal couplings is extended with respect to total variation.

Key words. marginal problem, Monge function, marginal constraint, transportation problem

AMS subject classifications. 60E15, 49A36

1. Introduction. For distribution functions F_1, F_2 let $\mathcal{F}(F_1, F_2)$ denote the set of all *df*'s on $\mathbb{R}^2$ with marginals F_1, F_2 (i.e., $F(x,\infty) = F_1(x), F(\infty, y) = F_2(y)$). Then the transportation problem with cost function $c \geq 0$ is to

$$\text{minimize} \int_{\mathbb{R}^2} c(x,y)dF(x,y) \quad \text{over all } F \in \mathcal{F}(F_1, F_2). \tag{1.1}$$

F_1 may be viewed as the supply distribution and F_2 as the demand distribution. Clearly, (1.1) is an infinite dimensional analogue of the discrete transportation problem: given $a_i \geq 0, b_j \geq 0, \sum_{i=1}^m a_i = \sum_{j=1}^n b_j$,

$$\begin{gathered}\text{minimize} \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij}, \quad \text{subject to the conditions:}\\ \sum_{j=1}^{n} x_{ij} = a_i, \quad 1 \leq i \leq m, \qquad \sum_{i=1}^{m} x_{ij} = b_j, \quad j = 1, \ldots, n, \quad x_{ij} \geq 0, \quad \forall i,j.\end{gathered} \tag{1.2}$$

If $c(x,y)$ (respectively (c_{ij})) satisfies the "Monge" conditions, i.e., c is right continuous and

$$c(x',y') - c(x,y') - c(x',y) + c(x,y) \leq 0 \quad \text{for all } x' \geq x, y' \geq y, \tag{1.3}$$

respectively

$$c_{ij} + c_{i+1,j+1} - c_{i,j+1} - c_{i+1,j} \leq 0, \quad \forall 1 \leq i < m, 1 \leq j < n, \tag{1.4}$$

then the solution of (1.1), (1.2) is well known and based on the "North-West corner rule," which leads to a greedy algorithm. For (1.1) the solution is given by the *df* F^*

$$F^*(x,y) = \min\{F_1(x), F_2(y)\}. \tag{1.5}$$

---

* Received by the editors October 15, 1990; accepted (in revised form) December 1, 1992.

† Statistics Department, University of California, Santa Barbara, California 93106. Supported in part by North Atlantic Treaty Organization grant CRG 900798 and by the Ministere de la Recherche et de la Technologie, France.

‡ Institut für Mathematik Statistik, Universität Münster, Einsteinstr. 62, D-4400 Münster, Germany. This author was supported in part by a Deutsche Forschungsgemeinschaft grant.

F^* is the upper Fréchet-bound. The Fréchet-bounds provide the following characterization of $\mathcal{F}(F_1, F_2)$:

(1.6)
$$F \in \mathcal{F}(F_1, F_2) \quad \text{if and only if}$$
$$F_*(x,y) := (F_1(x) + F_2(y) - 1)_+ \leq F(x,y) \leq F^*(x,y) \quad (\text{here } (\cdot)_+ = \max(0, \cdot)).$$

The lower Fréchet bound yields to a solution of the maximization problem corresponding to (1.1) (cf. [4], [5], [11]–[13]).

In terms of random variables an equivalent formulation of the transportation problem is the following:

(1.7) $$\text{minimize } Ec(X,Y), \quad \text{subject to } F_X = F_1, F_Y = F_2,$$

where X, Y are random variables on a rich enough (e.g., atomless) probability space $(\Omega, \mathcal{U}, \mathcal{P})$. The solutions (1.5) respectively (1.6) then can be represented as distributions of rv's X^*, Y^*:

(1.8) $$X^* = F_1^{-1}(U), \qquad Y^* = F_2^{-1}(U) \quad (\text{for } (1.1), (1.5)),$$

respectively

(1.9) $$X^* = F_1^{-1}(U), \qquad Y^* = F_2^{-1}(1-U) \quad (\text{for } F_*),$$

where U is uniformly distributed on (0,1), and $F_1^{-1}(u) = \inf\{y : F_1(y) \geq u\}$ is the generalized inverse of F_1 (cf. [4], [11]–[13]). (Throughout the paper we assume that *df*'s are right continuous.) For a general review on the Monge-Kantorovich transportation problem we refer to [8] and [1].

In this paper we study modifications of the transportation problem (1.1), where we relax or add new constraints. One type of additional side conditions has been studied by Barnes and Hoffman [2], in the discrete transportation problem (1.2); namely, additional capacity constraints $\sum_{r=1}^{i} \sum_{s=1}^{j} x_{rs} \leq \gamma_{ij}, i \leq m-1, j \leq n-1$, and a solution was obtained by a greedy algorithm.

In the first part of this paper we make use of the assumption that the cost function is of Monge type. These conditions seem to be necessary, since already in the simpler discrete case there are no general explicit solutions without conditions of this type. In the second part, under the restrictions of given difference of the marginals, we obtain explicit results without the Monge condition. We study extensions to the multivariate case for cost functions of the type $c_p(x,y) = \|x-y\|_p$, $\| \ \|_p$ the p-norm on $\mathbb{R}^n$ (c_p is not a Monge function for $n \geq 2$, and this problem is unsolved also in the discrete case). In the final section we consider local constraints on the marginals. In particular, we extend the classical Dobrushin result providing a construction of optimal couplings.

As for the proof of our results we use different methods from marginal problems, stochastic ordering, and duality theory. It seems that it is not possible to derive them all in a unified way; e.g., in §2, we construct in Theorems 1 and 2 solutions of the transportation problem with upper and lower bounds on the marginals under different assumptions on the cost functions. The proof of Theorem 1—for symmetric cost functions—is based on marginal problems, while the proof of Theorem 2—for unimodal cost functions—is based on stochastic ordering arguments.

2. Relaxation of the marginal constraints. Consider for *df*'s F_1, F_2 the set

(2.1) $$\mathcal{H}(F_1, F_2) = \{F : F \text{ is a } df \text{ on } \mathbb{R}^2 \text{ with marginal } df\text{'s } \tilde{F}_1 \leq F_1, \tilde{F}_2 \geq F_2\}$$

of all df's F with $\tilde{F}_1(x) = F(x,\infty) \le F_1(x)$, always $x \in \mathbb{R}^1$, and $\tilde{F}_2(y) = F(\infty, y) \ge F_2(y), \forall y \in \mathbb{R}^1$. We study the transportation problem:

(2.2) $$\text{minimize } \int_{\mathbb{R}^2} c(x,y)dF(x,y), \quad \text{subject to } F \in \mathcal{H}(F_1, F_2)$$

or, equivalently,

(2.3) $$\text{minimize } Ec(X,Y), \quad \text{subject to } F_X \le F_1, F_Y \ge F_2.$$

In the discrete case the problem is to minimize $\sum c_{ij}x_{ij}$ where for some "supplies" $s_1, \dots, s_n$, $a_1 \le s_1, a_1 + a_2 \le s_1 + s_2, \dots$, and for some demands $d_1, \dots, d_n, b_1 \ge d_1, b_1 + b_2 \ge d_1 + d_2, \dots$, ($a_i, b_i$ as in (1.2)). This describes production and consumption processes based on priorities (e.g., by time) with capacities $s_1, \dots, s_n$, such that what is remained in stage i of the production (respectively consumption) process can be transferred to some of the next stages $i+1, \dots, n$.

THEOREM 1. *Suppose the cost function $c(x,y)$ is symmetric, $c(x,y)$ satisfies the Monge-condition* (1.3), *and let $c(x,x) = 0, \forall x$. Define*

(2.3) $$H^*(x,y) = \min\{F_1(x), \max\{F_1(y), F_2(y)\}\}, \qquad x, y \in \mathbb{R}.$$

Then

(2.4)
(a) $H^* \in \mathcal{H}(F_1, F_2)$,
(b) H^* *solves the relaxed transportation problem* (2.2),
(c) $$\int_{\mathbb{R}^2} c(x,y)dH^*(x,y) = \int_0^1 c(F_1^{-1}(u), \min(F_1^{-1}(u), F_2^{-1}(u)))du.$$

Remark 1. Setting the df $G_1(y) = \max\{F_1(y), F_2(y)\}$, we see from Theorem 1 that the relaxed transportation problem (2.2) is equivalent to the transportation problem (1.1) with marginals F_1, G_1. In terms of random variables a solution is given by

(2.5) $$X^* = F_1^{-1}(U), \qquad Y^* = G_1^{-1}(U) = \min(F_1^{-1}(U), F_2^{-1}(U)) \qquad \text{(cf. (1.8))}.$$

Proof. From the Monge condition the function $-c(x,y)$ may be viewed as a "distribution function" corresponding to a nonnegative measure μ_c on $\mathbb{R}^2$. Let X, Y be any real rv's and for $x, y \in \mathbb{R}^1$ denote $x \vee y = \max\{x,y\}, x \wedge y = \min\{x,y\}$. Theorem 1 is a consequence of the following two claims.

CLAIM 1 (Cambanis, Simons, and Stout [4], Dall'Aglio [5] for $c(x,y)$)

(2.6) $$2Ec(X,Y) = \int_{\mathbb{R}^2} (P(X < x \wedge y, Y \ge x \vee y) + P(X \ge x \vee y, Y < x \wedge y))\mu_c(dx,dy).$$

For the proof of Claim 1 define the function $f(x,y,w) : \mathbb{R}^2 \times \Omega \to \mathbb{R}$ by

$$f(x,y,w) = \begin{cases} 1 & \text{if } (X(w) < x, y \le Y(w)) \text{ or } (Y(w) < x, y \le X(w)) \\ 0 & \text{otherwise} \end{cases}$$

By Fubini's theorem,

(2.7) $$E_w \int_{\mathbb{R}^2} f(x,y,w)\mu_c(dx,dy) = \int_{\mathbb{R}^2} (E_w f(x,y,w))\mu_c(dx,dy).$$

Next the symmetry of $c(x,y)$ and $c(x,x)=0$ yields

$$\begin{aligned}(2.8)\qquad \int_{\mathbb{R}^2} f(x,y,w)d\mu_c &= -[c(Y(w),Y(w))+c(X(w),X(w))\\ &-c(X(w),Y(w))-c(Y(w),X(w))] = 2c(X(w),Y(w)).\end{aligned}$$

Clearly,

$$(2.9)\qquad E_w f(x,y,w) = P(X<x\wedge y, Y\ge x\vee y)+P(X\ge x\vee y, Y<x\wedge y).$$

Combining (2.7), (2.8), (2.9), we obtain (2.6).

CLAIM 2. *Define* $X^*=F_1^{-1}(U), Y^*=\min(F_1^{-1}(U),F_2^{-1}(U))$; *then*

$$(2.10)\qquad Ec(X^*,Y^*)=\min\{Ec(X,Y); F_X\le F_1, F_Y\ge F_2\}$$

and the value of the expectation in (2.10) *is given by*

$$\begin{aligned}(2.11)\qquad Ec(X^*,Y^*) &= \frac{1}{2}\int_{\mathbb{R}^2}\max\{0, F_2((x\wedge y)-)-F_1((x\vee y)-)\}\mu_c(dx,dy)\\ &= \int_0^1 c(F_1^{-1}(t),\min\{F_1^{-1}(t),F_2^{-1}(t)\})dt.\end{aligned}$$

For the proof of Claim 2 let X, Y be any rv's with df's $F_X\le F_1, F_Y\ge F_2$. Using Claim 1 we obtain

$$\begin{aligned}2Ec(X,Y) &\ge \int_{\mathbb{R}^2} P(X\ge x\vee y, Y<x\wedge y)\mu_c(dx,dy)\\ &= \int_{\mathbb{R}^2}\{P(Y<x\wedge y)-P(X<x\vee y, Y<x\wedge y)\}\mu_c(dx,dy)\\ (2.12)\qquad &\ge \int_{\mathbb{R}^2}\{P(Y<x\wedge y)-\min\{P(X<x\vee y), P(Y<x\wedge y)\}\mu_c(dx,dy)\\ &= \int_{\mathbb{R}^2}(P(Y<x\wedge y)-P(X<x\vee y))_+\mu_c(dx,dy)\\ &\ge \int_{\mathbb{R}^2}(F_2((x\wedge y)-)-F_1((x\vee y)-))_+\mu_c(dx,dy).\end{aligned}$$

Next we check that the lower bound we get in (2.12) is attained for $X^*=F_1^{-1}(U), Y^*=\min(F_1^{-1}(U),F_2^{-1}(U))$. In fact, by Claim 1 using $X^*\ge Y^*$ and $\{U<F_2(z)\}=\{F_2^{-1}(U)<z\}$ almost surely we obtain

$$\begin{aligned}&2Ec(X^*,Y^*)\\ &= \int_{\mathbb{R}^2}\{P(X^*\ge x\vee y, Y^*<x\wedge y)+P(X^*<x\wedge y, Y^*\ge x\vee y)\}\mu_c(dx,dy)\\ &= \int_{\mathbb{R}^2}P(X^*\ge x\vee y, Y^*<x\wedge y)\mu_c(dx,dy)\\ (2.13)\qquad &= \int_{\mathbb{R}^2}P(F_1^{-1}(U)\ge x\vee y, \min(F_1^{-1}(U),F_2^{-1}(U))<x\wedge y)\mu_c(dx,dy)\\ &= \int_{\mathbb{R}^2}P(F_1^{-1}(U)\ge x\vee y, F_2^{-1}(U)<x\wedge y)\mu_c(dx,dy)\\ &= \int_{\mathbb{R}^2}P(U\ge F_1(x\vee y), U<F_2(x\wedge y))_+\mu_c(dx,dy)\\ &= \int_{\mathbb{R}^2}(F_2((x\wedge y)-)-F_1((x\vee y)-))_+\mu_c(dx,dy).\end{aligned}$$

Obviously, $F_{(X^*,Y^*)} = H^* \in \mathcal{H}(F_1, F_2)$ and the proof of Theorem 1 is completed. □

Remark 2. Equation (2.5) suggests the following "greedy" algorithm for solving the finite discrete transportation problem with relaxed side conditions:

(2.14)
$$\begin{aligned} &\text{minimize} && \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij} \\ &\text{subject to:} && x_{ij} \geq 0 \\ &&& \sum_{s=1}^{j}\sum_{r=1}^{n} x_{rs} \geq \sum_{s=1}^{j} b_s =: G_j, \qquad 1 \leq j \leq n \\ &&& \sum_{r=1}^{i}\sum_{s=1}^{n} x_{rs} \leq \sum_{r=1}^{i} a_r =: F_i, \qquad 1 \leq i \leq n, \end{aligned}$$

where the sum of the "demands" $\sum_{s=1}^{n} b_s$ equals the sum of the "supplies" $\sum_{r=1}^{n} a_r$, assuming that (c_{ij}) are symmetric, $c_{ii} = 0$ and c satisfying the Monge condition (1.4). Denote

(2.15)
$$\begin{aligned} &H_i = \max(F_i, G_i), \qquad 1 \leq i \leq n, \quad \text{and} \\ &\delta_1 = H_1, \delta_{i+1} = H_{i+1} - H_i, \qquad 1 \leq i \leq n-1; \end{aligned}$$

(2.14) is equivalent to the standard transportation problem (1.2) with side conditions (a_i), (δ_i). In the following example we compare the solution of problem (2.14) with inequality constraints with the "greedy" solution of the standard transportation problem with equality constraints (1.2). For the problem with inequality constraints we first calculate the new artificial demands δ_j as in (2.15) and then apply the North-West corner rule.

Example.

$\boxed{y_{ij}}$ x_{ij}							supply a_i	$F_i = \sum_{r=1}^{i} a_r$
	$\boxed{20}$ 10	10					20	20
							0	20
		$\boxed{20}$ 20	$\boxed{20}$ 10	10			40	60
				$\boxed{20}$ 20			20	80
				$\boxed{10}$ 10			10	90
						$\boxed{10}$ 10	10	100
demand b_j	10	30	10	40	0	10		
$G_j = \sum_{s=1}^{j} b_s$	10	40	50	90	90	100		
$H_j = F_j \vee G_j$	20	40	60	90	90	100		
$\delta_1 = H_1, \delta_{j+1} = H_{j+1} - H_j$	20	20	20	30	0	10	"artificial" demands	

$x_{ij} =$ solution of the standard transportation problem (1.2), using the classical North-West corner

$y_{ij} =$ solution of the transportation problem with relaxed side conditions.

We next extend the solution to the non-symmetric case. We assume instead of symmetry the following unimodality condition, saying that for any x, y the functions $c(x,\cdot), c(\cdot, y)$ are unimodal; more precisely,

$$\begin{aligned} c(x,y_1) &\le c(x,y_2) \quad \text{if } x \le y_1 \le y_2 \text{ or } y_2 \le y_1 \le x, \quad \text{and} \\ c(x_1,y) &\le c(x_2,y) \quad \text{if } x_2 \le x_1 \le y \text{ or } y \le x_1 \le x_2. \end{aligned} \tag{2.16}$$

For the proof of this unimodal case we basically make use of stochastic ordering arguments.

THEOREM 2. *If $c(x,x) = 0$ for all x, and c satisfies the Monge condition* (1.3) *and the unimodality condition* (2.16), *then the relaxed transportation problem,*

$$\text{minimize } Ec(X,Y) \text{ subject to: } F_X \ge F_1, F_Y \le F_2, \tag{2.17}$$

has the solution

$$\begin{aligned} &X^* = F_1^{-1}(U), \qquad Y^* = \max(F_1^{-1}(U), F_2^{-1}(U)), \quad \text{so} \\ &F_{X^*,Y^*}(x,y) = \min(F_1(x), \min(F_1(y), F_2(y)) \quad \text{and} \\ &Ec(X^*,Y^*) = \int_0^1 c(F_1^{-1}(u), \max(F_1^{-1}(u), F_2^{-1}(u))du. \end{aligned} \tag{2.18}$$

Proof. Let X, Y be rv's with $F_X \ge F_1, F_Y \le F_2$; then by (1.8)

$$Ec(X,Y) \ge Ec(F_X^{-1}(U), F_Y^{-1}(U)). \tag{2.19}$$

Let $G(y) = \min(F_X(y), F_Y(y))$; then $F_X^{-1} \le F_1^{-1}, F_Y^{-1} \ge F_2^{-1}$ and $G^{-1} = \max(F_X^{-1}, F_Y^{-1})$. We now state the following.

CLAIM 1.

$$\int_0^1 c(F_X^{-1}(u), F_Y^{-1}(u))du \ge \int_0^1 c(F_X^{-1}(u), G^{-1}(u))du. \tag{2.20}$$

To show Claim 1 let for fixed $u \in (0,1), x = F_X^{-1}(u), y_1 = F_X^{-1}(u) \vee F_Y^{-1}(u) = G^{-1}(u)$, $y_2 = F_Y^{-1}(u)$.

Case 1. $x < y_2$. In this case, $x \le y_1 \le y_2$, and, therefore, the unimodality condition (2.18) implies $c(x,y_2) \ge c(x,y_1)$.

Case 2. $y_2 \le x$. In this case, $y_1 = x$ and therefore, $y_2 \le y_1 = x$. Again by the unimodality condition $c(x,y_2) \ge c(x,y_1)$. So Claim 1 holds.

CLAIM 2.

$$\int_0^1 c(F_X^{-1}(u), F_Y^{-1}(u) \vee F_X^{-1}(u))du \ge \int_0^1 c(F_1^{-1}(u), F_2^{-1}(u) \vee F_1^{-1}(u))du. \tag{2.21}$$

For the proof, define $\tilde{x}_1 = F_X^{-1}(u), \tilde{x}_2 = F_Y^{-1}(u), x_1 = F_1^{-1}(u), x_2 = F_2^{-1}(u)$ for fixed u. Then $\tilde{x}_1 \le x_1, x_2 \le \tilde{x}_2$.

$$\begin{aligned} &\text{If} \quad \tilde{x}_1 < \tilde{x}_2, \quad \text{then } \tilde{x}_1 \le \tilde{x}_2 \vee x_2 \le \tilde{x}_2, \\ &\text{if} \quad \tilde{x}_1 \ge \tilde{x}_2, \quad \text{then } \tilde{x}_1 = \tilde{x}_1 \vee x_2 \ge \tilde{x}_2. \end{aligned} \tag{2.22}$$

From (2.22) we obtain the following claim.

CLAIM 3.

$$c(\tilde{x}_1, \tilde{x}_1 \vee x_2) \ge c(x_1, x_1 \vee x_2). \tag{2.23}$$

For the proof of Claim 3 we use the relation $x_1 \geq \tilde{x}_1$. By (2.22) we have two cases.

Case 1. $x_2 > x_1 > \tilde{x}_1$. Then $c(\tilde{x}_1, x_2) = c(\tilde{x}_1, \tilde{x}_1 \vee x_2) \geq c(x_1, x_2) = c(x_1, x_1 \vee x_2)$ by the unimodality condition.

Case 2. (a) $x_1 \geq x_2 \geq \tilde{x}_1$. Then, trivially, $c(\tilde{x}_1, x_2) = c(\tilde{x}_1, x_2 \vee \tilde{x}_1) \geq c(x_1, x_1 \vee x_2) = c(x_1, x_1) = 0$.

(b) $x_1 \geq \tilde{x}_1 \geq x_2$. Then again $c(\tilde{x}_1, \tilde{x}_1) = c(\tilde{x}_1, \tilde{x}_1 \vee x_2) \geq c(x_1, x_1 \vee x_2) = c(x_1, x_1) = 0$, trivially.

Claims 1, 2, and 3 imply (2.18).

Remark 3.

(a) The unimodality assumption (2.16) is natural from the application point of view. Note that the transportation problem in Theorem 2 is the same as in Theorem 1 (where only the indices 1 and 2 have been changed). We used this change to demonstrate that the optimal solution F^* is not unique, but there is a large range of solutions. As a consequence observe that in order to achieve an optimal solution for the transportation problem with side conditions, either the demands can be adjusted by transports on or below the diagonal, or alternatively, the supplies can be adjusted in a similar way. Without the symmetry, respectively the unimodality condition, the solution may change extremely. Consider for any right continuous function $f = f(y) \geq 0$ the cost function $c(x, y) = f(y)$. Then c satisfies the Monge-condition, and so (2.17) is equivalent to the problem,

(2.24)
$$\text{minimize} \int f(y)dF_Y(y) \quad \text{subject to } F_Y \leq F_2,$$

i.e., we are looking for a $df\, \tilde{F}_2 \leq F_2$, such that the distribution of f with respect to $\tilde{F}_2$ has a minimal first moment. Obviously, the solution (2.20) of Theorem 2 is not a solution of (2.24).

(b) For the proof of Theorem 2 the assumption $c(x, x) = 0$ can be replaced by the weaker one,

(2.25)
$$c(x, x) \leq c(x, y) \wedge c(y, x), \quad \forall x, y.$$

3. Given sum of the marginals. Consider a flow in a network with n-nodes $i = 1, \dots, n$, and let x_{ij} be the flow from node i to node j. Assume that for all nodes k the value of $\sum_i x_{ik} + \sum_j x_{kj}$ is fixed to be h_k. For a motivation of this problem let $a_i = \sum_{k=1}^n x_{ik}$, $b_i = \sum_{k=1}^n x_{ki}$ be the amount of labor corresponding to the outflow respectively to the inflow in node i. Assume that the total labor capacity in node i is given by h_i (in a certain time unit); then an admissible flow (x_{ij}) should satisfy the condition

(3.1)
$$h_i = a_i + b_i, \qquad 1 \leq i \leq n.$$

Let $F_1(k) = \sum_{i=1}^k a_i$, $F_2(k) = \sum_{i=1}^k b_i$, $H(k) = \sum_{i=1}^k h_k$; then $h_k = F_1(k) + F_2(k) - (F_1(k-1) + F_2(k-1))$ and (3.1) is equivalent to

(3.2)
$$H(k) = F_1(k) + F_2(k), \qquad 1 \leq k \leq n.$$

Let c_{ij} denote the cost of transporting a unit from node i to node j; then the problem is to minimize the total cost $\sum c_{ij} x_{ij}$ subject to condition (3.2) and $x_{ij} \geq 0$.

The general formulation of this problem is the following. For two df's F_1, F_2 define $G(x) := \frac{1}{2}(F_1(x) + F_2(x))$. For a cost function $c(x, y)$ consider the problem,

(3.3)
$$\text{minimize} \int_{\mathbb{R}^2} c(x, y)dF(x, y) \quad \text{subject to } F \in \mathcal{F}_G,$$

where $\mathcal{F}_G$ is the set of all *df*'s $F(x,y)$ with marginal *df*'s $\tilde{F}_1, \tilde{F}_2$ satisfying $\tilde{F}_1(x) + \tilde{F}_2(x) = 2G$.

In the special case $c(x,y) = |x-y|$, let X, Y be real *rv*'s. Then by the triangle inequality

$$E|X-Y| \leq \inf_{a\in\mathbb{R}^1}(E|X-a| + E|Y-a|), \tag{3.4}$$

(3.4) is the optimal bound if one knows only $E|X-a|, a \in \mathbb{R}^1$. Note that $E|X-a| + E|Y-a| = \int |x-a| d(F_X+F_Y)(x)$ only depends on the sum of the marginals. Equation (3.3) is the best possible improvement of (3.4) provided $F_X + F_Y$ is known.

It was shown in [9] that

$$\sup\{E|X-Y|^p; F_X+F_Y = 2G\} = \int_0^1 |G^{-1}(t) - G^{-1}(1-t)|^p dt, \qquad p \geq 1. \tag{3.5}$$

PROPOSITION 3. *If $c \geq 0$ is symmetric and satisfies the Monge condition* (1.3), *then*

$$\inf\left\{\int c(x,y)dF(x,y); F \in \mathcal{F}_G\right\} = \int_0^1 c(G^{-1}(u), G^{-1}(u))du, \tag{3.6}$$

$$\sup\left\{\int c(x,y)dF(x,y); F \in \mathcal{F}_G\right\} = \int_0^1 c(G^{-1}(u), G^{-1}(1-u))du. \tag{3.7}$$

Optimal pairs of rv's are given by $(G^{-1}(U), G^{-1}(U))$ respectively $(G^{-1}(U), G^{-1}(1-U))$.

Proof. Since c is symmetric, we obtain for any $F \in \mathcal{F}_G$, $\int c(x,y)dF(x,y) = \int \frac{1}{2}(c(x,y) + c(y,x))dF(x,y) = \int c(x,y)d\{[F(x,y)+F(y,x)]/2\}$. But $F_s(x,y) = [F(x,y)+F(y,x)]/2 \in \mathcal{F}(G,G)$, so we obtain (3.6), (3.7) by application of (1.8), (1.9). □

For non-symmetric cost functions we have the following.

PROPOSITION 4. *If $c(x,y)$ satisfies the Monge condition and furthermore $x_1 \leq y \leq x_2$ implies that $c(x_1,x_2) \geq c(y,y)$, then*

$$\inf\left\{\int c(x,y)dF(x,y); F \in \mathcal{F}_G\right\} = \int_0^1 c(G^{-1}(u), G^{-1}(u))du. \tag{3.8}$$

Proof. For *rv*'s X, Y with $F_{X,Y} \in \mathcal{F}_{A+B}$, by the Monge condition $Ec(X,Y) \geq Ec(F_X^{-1}(U), F_Y^{-1}(U))$. Since $F_X(x) + F_Y(x) = 2G(x)$, it follows that $F_X \wedge F_Y \leq G \leq F_X \vee F_Y$, and therefore, $F_X^{-1} \wedge F_Y^{-1} \leq G^{-1} \leq F_X^{-1} \vee F_Y^{-1}$. It follows that $c(F_X^{-1}(U), F_Y^{-1}(U)) \geq c(G^{-1}(U), G^{-1}(U))$ implying (3.8). □

Remark 4. The set of marginals in the class $\mathcal{F}_G$ has a smallest and a largest element, namely

$$F_1^*(x) = \begin{cases} 2G(x), & x < x_0 \\ 1 & x \geq x_0 \end{cases} \quad \text{and} \quad F_2^*(x) = \begin{cases} 2G(x)-1, & x \geq x_0 \\ 0 & x < x_0 \end{cases},$$

where $x_0 = \inf\{y; 2G(y) \geq 1\}$. There is no smallest *df* in $\mathcal{F}_G$. For the proof let $H_1(x), H_2(x)$ be the marginal *df*'s of a smallest element $H \in \mathcal{F}_G$ and let G_1, G_2 be *df*'s such that $G_1(x) + G_2(x) = 2G(x)$. If the lower Fréchet bounds satisfy $(H_1(x) + H_2(y) - 1)_+ \leq (G_1(x) + G_2(y) - 1)_+$, then $H_1 \leq G_1$ and $H_2 \leq G_2$, which amounts to $H_1 = G_1, H_2 = G_2$. In particular, this implies that $(G^{-1}(U), G^{-1}(1-U))$ is in the general non-symmetric case no longer a solution to the problem to maximize $\int c(x,y)dF(x,y)$ in the class $\mathcal{F}_G$. Let e.g., G be the *df* of $\frac{1}{4}\sum_{i=1}^4 \varepsilon_{\{i\}}$; then $P_1 = P^{(G^{-1}(U), G^{-1}(1-U))} = \frac{1}{4}(\varepsilon_{(1,4)} + \varepsilon_{(2,3)} + \varepsilon_{(3,2)} + \varepsilon_{(4,1)})$, while $P_2 = P^{((F_1^*)^{-1}(U), (F_2^*)^{-1}(1-U))} = \frac{1}{2}(\varepsilon_{(1,4)} + \varepsilon_{(2,3)})$. For $c_1(x,y) = 1_{(-\infty,(3,2)]}(x,y)$, we have $E_{P_1}c_1 = \frac{1}{4}, E_{P_2}c_1 = 0$, while for $c_2 = 1_{[(2,3),\infty)}, E_{P_1}c_2 = \frac{1}{4}, E_{P_2}c_2 = \frac{1}{2}$. Note that both functions, $-c_1, -c_2$, are Monge functions (but are not unimodal).

4. Given difference of the marginals. We next consider the case where in the network example we fix the total outflow minus the inflow of each node. This problem is known in the literature as minimal network flow problem (cf. e.g., [3, §9], or [1]). Similarly to §3 the outflow minus the inflow of each node is fixed; i.e., the following Kirchhoff equations hold: $\sum_k x_{ik} - \sum_k x_{ki} = a_i - b_i = h_i$ for all i, or, equivalently, with $F_1(k) = \sum_{j=1}^k a_j, F_2(k) = \sum_{j=1}^k b_j, H(k) = \sum_{j=1}^k h_j, H(k) = F_1(k) - F_2(k), 1 \le k \le n$. Let more generally F_1, F_2 be distribution functions and let $\mathcal{F}_H$ be the set of all "df's" of finite measures on $\mathbb{R}^2$ with marginals $\tilde{F}_1, \tilde{F}_2$ satisfying $\tilde{F}_1 - \tilde{F}_2 = F_1 - F_2 =: H$. We consider the following transportation problem:

$$\text{minimize} \int c(x,y)dF(x,y) \quad \text{subject to } F \in \mathcal{F}_H. \tag{4.1}$$

$c(x,y)$ is symmetric, nonnegative and continuous, but does not need to satisfy the Monge conditions. For the solution we shall make use of the following dual representation (cf. Rachev and Shortt [10]):

$$\begin{aligned} &\inf\left\{\int c(x,y)dF(x,y); F \in \mathcal{F}_H\right\} \\ &= \sup\left\{\int f dH(x); f(x) - f(y) \le c(x,y), \forall x, y\right\}. \end{aligned} \tag{4.2}$$

We first consider a special type of cost functions.

PROPOSITION 5. *Let* $c(x,y) = |x-y| \max(1, h(|x-a|), h(|y-a|))$, *where* h *is monotonically nondecreasing. Then*

$$\inf\left\{\int c(x,y)dF(x,y); F \in \mathcal{F}_H\right\} = \int \max(1, h(|x-a|))|H|(x)dx, \tag{4.3}$$

provided $h(|x-a|)$ *is locally integrable.*

Proof. For the cost function c we observe that $f(x) - f(y) \le c(x,y)$, for all x, y, if and only if f is absolutely continuous with $|f'(x)| \le \max(1, h(|x-a|))$ almost surely. By the dual representation (4.2) and partial integration we obtain

$$\begin{aligned} &\inf\left\{\int c(x,y)dF(x,y); F \in \mathcal{F}_H\right\} \\ &= \sup\left\{\int f d(H)(x); |f'(x)| \le \max(1, h(|x-a|)), \forall x\right\} \\ &= \sup\left\{\int f'(x)(H)(x)dx; |f'(x)| \le \max(1, h(|x-a|)), \forall x\right\} \\ &= \int \max(1, h(|x-a|))|H|(x)dx. \qquad \square \end{aligned}$$

On the basis of the idea of this proof, we next consider more generally

$$c(x,y) = |x-y|\zeta(x,y) \quad \left(\text{i.e. } \zeta(x,y) = \frac{c(x,y)}{|x-y|}\right). \tag{4.4}$$

THEOREM 6 (Generalized Kantorovich–Rubinstein problem). *Assume that for any* $x < t < y, \zeta(t,t) \le \zeta(x,y), \zeta(x,y)$ *symmetric and continuous on the diagonal and also that* $t \to \zeta(t,t)$ *is locally bounded; then*

$$\inf\left\{\int c(x,y)dF(x,y); F \in \mathcal{F}_H\right\} = \int \zeta(t,t)|H|(t)dt. \tag{4.5}$$

Proof. Let $\mathcal{F} = \{f : f(x) - f(y) \le c(x,y)$, for all $x, y\}$ and let $\mathcal{F}^* = \{f$ absolutely continuous and $|f'(t)| \le \zeta(t,t)$, for all $t\}$; then $\mathcal{F} \subset \mathcal{F}^*$ as for $f \in \mathcal{F}$, we have $[f(x) - f(y)]/|x-y| \le \zeta(x,y)$ and, therefore, $\overline{\lim}_{y\to x} [f(x)-f(y)]/|x-y| \le \zeta(x,x)$. Also $\underline{\lim}_{y\to x} [f(x)-f(y)]/|x-y| = -\overline{\lim}\, [f(y)-f(x)]/|x-y| \ge -\overline{\lim}\, \zeta(y,x) = -\zeta(x,x)$. As $\zeta(t,t)$ is locally bounded, f is locally Lipschitz, absolutely continuous, and the inequalities above imply that $|f'(t)| \le \zeta(t,t)$ almost surely. If, conversely, $f \in \mathcal{F}^*$, then $f(x) - f(y) = \int_x^y f'(t)dt$, and therefore, $|f(x)-f(y)| \le \int_x^y |f'(t)|dt \le \int_x^y \zeta(t,t)dt \le |x-y|\zeta(x,y) = c(x,y)$. The dual representation (4.2) again implies (4.3) as in the proof of Proposition 5. □

It is very interesting to observe that restrictions on the difference of the marginals allow this general explicit result without "special" assumptions on c. Note that the solution only depends on the behavior of c at the diagonal, a property that is observed in the minimal network flow problems. Note that from Theorem 6 one obtains the remarkable consequence that

$$\inf\left\{\int |x-y|^p dF(x,y); F \in \mathcal{F}_H\right\} = 0 \tag{4.6}$$

for all $p > 1$, which confirms that cost functions as in Theorem 5 are of the right order.

We next consider an extension to the multivariate case $\mathbb{R}^n$ with the class of cost functions

$$c_p(x,y) = \|x-y\|_p = \left(\sum_{i=1}^n |x_i - y_i|^p\right)^{1/p}, \qquad 1 \le p < \infty.$$

Let F_1, F_2 be n-dimensional distribution functions and let for $H := F_1 - F_2$; $\mathcal{F}_H$ denotes the class of all $2n$-dimensional (joint) distribution functions F with n-dimensional marginals $\tilde{F}_1, \tilde{F}_2$ such that $\tilde{F}_1 - \tilde{F}_2 = H$. Denote

$$A_p(H) := \inf\left\{\int_{\mathbb{R}^{2n}} \|x-y\|_p dF(x,y); F \in \mathcal{F}_H\right\},$$

the value of the optimal multivariate transportation costs. Let $1/q + 1/p = 1$ and assume that F_1, F_2 have densities f_1, f_2 with respect to the Lebesgue measure $h := f_1 - f_2$.

THEOREM 7. (Multivariate transportation problem). (a) *For the value of the optimal transportation costs we have the upper bound*

$$A_p(H) \le B_p(H) := \int_{\mathbb{R}^n} \|y\|_p |J_H(y)| dy, \tag{4.8}$$

where $J_H(y) := \int_0^1 t^{-(n+1)} h(y/t)dt$.

(b) *If there exists a continuous function* $g : \mathbb{R}^n \to \mathbb{R}^1$, *almost everywhere differentiable and satisfying for* $p = 1$

$$\nabla g(y) = (\operatorname{sgn}(y_i J_H(y))) \quad \text{a.e.}, \tag{4.9}$$

respectively for $p > 1$,

$$\nabla g(y) = (\operatorname{sgn}(y_i J_H(y))) \left(\frac{|y_i|}{\|y\|_q}\right)^{q/p}, \tag{4.10}$$

then equality in (4.8) *holds.*

Proof. (a) From the duality theorem in Rachev and Shortt [10]

$$A_p(H) = \sup\left\{\left|\int_{\mathbb{R}^n} f dH\right|; |f(x) - f(y)| \leq \|x - y\|_p\right\}.$$

From the Radermacher theorem we infer that any Lipschitz function f is almost everywhere differentiable, and as $\sup\{\langle\nabla f(y), a\rangle; \|a\|_p = 1\} = \|\nabla f(y)\|_q$, we obtain from the Lipschitz condition that $\|\nabla f(y)\|_q \leq 1$ almost everywhere. Using a Taylor expansion

$$f(y) = f(0) + \int_0^1 \langle\nabla f(ty), y\rangle\, dt,$$

we conclude that

$$\begin{aligned} A_p(H) &\leq \sup\left\{\left|\int_{\mathbb{R}^n}\int_0^1 \langle\nabla f(ty), y\rangle\, dt\, h(y)\, dy\right|; \|\nabla f(y)\|_q \leq 1 \text{ a.e.}\right\} \\ &= \sup\left\{\left|\int_{\mathbb{R}^n}\int_0^1 \langle\nabla f(y), y\rangle \frac{1}{t^{n+1}} h\left(\frac{y}{t}\right) dt\, dy\right|; \|\nabla f(y)\|_q \leq 1 \text{ a.e.}\right\} \\ &\leq \sup\left\{\int_{\mathbb{R}^n} \|y\|_p\, |J_H(y)|\, \|\nabla f(y)\|_q\, dy\; \|\nabla f(y)\|_q \leq 1 \text{ a.e.}\right\} \\ &\leq \int_{\mathbb{R}^n} \|y\|_p\, |J_H(y)|\, dy. \end{aligned} \tag{4.11}$$

(b) In the inequalities

$$|\langle x, y\rangle| \leq \sum |x_i y_i| \leq \|x\|_p\, \|y\|_q, \qquad \|x\|_p = 1,$$

equality is attained for $p > 1$ if and only if

$$x_i = \operatorname{sgn} y_i \frac{|y_i|^{q/p}}{\|y\|_q^{q/p}} = y_i \frac{|y_i|^{q/p-1}}{\|y\|_q^{q/p}},$$

while for $p = 1$ equality holds if and only if $\operatorname{sgn} x_i = \operatorname{sgn} y_i$. This implies part (b) of the Theorem. □

Remark 5. Conditions (4.9), (4.10) are fulfilled in dimension 1 so that the bound (4.8) is sharp and coincides with (4.3). A simple sufficient for $p = 1$ for (4.9) is given by

$$J_H \geq 0 \quad \text{a.e.,} \tag{4.12}$$

which is a stochastic ordering condition. More generally we can allow a "simple" structure of the set $\{J_H \geq 0\}$. We remark that the optimal multivariate transportation problem is a longstanding open problem also in the discrete case.

5. Upper bounds on the total transport mass. Let $\Gamma(x, y)$ be a "distribution function" of a finite measure and define for two fixed distribution functions F_1, F_2 on $\mathbb{R}^1$ the transportation problem:

$$H_\Gamma(x, y) := \sup\{F(x, y); F(x_i, y_i) \leq \Gamma(x_i, y_i), i \in I, F \in \mathcal{F}(F_1, F_2)\}, \tag{5.1}$$

where $(x_i, y_i)_{i\in I} \subset \mathbb{R}^2$ may be finite or not. From the Fréchet-bounds in (1.6), we have the following conditions ensuring the nontriviality of the problem:

$$\Gamma(x_i, y_i) \geq (F_1(x_i) + F_2(y_i) - 1)_+, \quad \forall\, i \in I. \tag{5.2}$$

Problem (5.1) is an extension of a problem treated by Barnes and Hoffman [2] in the finite discrete case and by Olkin and Rachev [7] in the general case. In these papers it was assumed that $F(x,y) \le \Gamma(x,y)$ for all (x,y). Problem (5.2) is motivated by capacitated transportation problems with linearly ordered supply and demand nodes (cf. [2]). Several examples of this problem and extensions to further restrictions on the support of solutions ("staircase supports") are discussed in Hoffman and Veinott [6]. An application to a graph partitioning problem is given in Barnes and Hoffman [2].

THEOREM 8. *Let assumption* (5.2) *be fulfilled and define*

$$F^*(x,y) := \inf_{\substack{x_i \le x \\ y_i \le y}} \{\Gamma(x_i,y_i) + (F_1(x) - F_1(x_i)) + (F_2(y) - F_2(y_i))\} \wedge \min\{F_1(x), F_2(y)\} \tag{5.3}$$

(*with the convention that the infimum is zero, if there do not exist* $x_i \le x, y_i \le y$).

(a) $H_\Gamma(x,y) \le F^*(x,y), \forall x,y.$

(b) *If* F^* *is a df, then*

$$H_\Gamma(x,y) = F^*(x,y) \quad \textit{and} \quad F^* \textit{ is a solution of } (5.1). \tag{5.4}$$

(c) (cf. [2], [7]). *If* $\{(x_i,y_i), i \in I\} = \mathbb{R}^2$, *then* F^* *is a df.*

Proof. (a) For $x_i \le x, y_i \le y$, we have for any admissible F using *rv*'s X, Y with $F_{X,Y} = F$, $F(x,y) = P(X \le x_i, Y \le y) + P(x_i < X \le x, Y \le y) = P(X \le x_i, Y \le y_i) + P(X \le x_i, y_i < Y \le y) + P(x_i < X \le x, Y \le y) \le \Gamma(x_i,y_i) + F_1(x) - F_1(x_i) + F_2(y) - F_2(y_i)$. Furthermore, by the Fréchet bounds (1.6), $F(x,y) \le \min\{F_1(x), F_2(y)\}$. Therefore, $F(x,y) \le F^*(x,y)$.

(b) If F^* is a *df*, then $F^* \in \mathcal{F}(F_1,F_2)$. For the proof observe that for $(x_i,y_i) \le (x,y)$ by (5.2), $\Gamma(x_i,y_i) + F_1(x) - F_1(x_i) + F_2(y) - F_2(y_i) \ge (F_1(x) + F_2(y) - 1)_+$ and so by definition of F^*, $(F_1(x) + F_2(y) - 1)_+ \le F^*(x,y) \le \min\{F_1(x), F_2(y)\}$, which implies by (1.6) that $F^* \in \mathcal{F}(F_1,F_2)$. Since $F^*(x_i,y_i) \le \Gamma(x_i,y_i)$, F^* is an admissible *df*, and, therefore, by (a) a solution of (5.1).

(c) For the proof of (c) we refer to [7]. □

Remark 6. (a) Parts (a) and (b) of Theorem 7 remain valid for any function $\Gamma(x,y) \ge 0$. The difficult part to verify is that F^* is a *df*. But it seems to be clear from the proof that, even in case when F^* is not a *df*, part (a) gives a good upper bound. An indication for this conclusion is part (c) of the theorem.

(b) From (5.4) one obtains for Monge functions c with the regularity condition

$$\int c(x,y_0)F_1(dx) + \int c(x_0,y)F_2(dy) < \infty$$

for some $x_0, y_0 \in \mathbb{R}$ that

$$\inf\left\{\int c(x,y)dF(x,y); F(x,y) \le \Gamma(x,y), \forall x,y, F \in \mathcal{F}(F_1,F_2)\right\} = \int c(x,y)dF^*(x,y). \tag{5.5}$$

(c) In the discrete case the solution F^* of (5.5) can be determined by the Barnes-Hoffman greedy algorithm (see [2], [6], [7]). In fact, if $a_i = F_1(x_i) - F_1(x_i-), i \in M =$

$\{1,\dots,m\}, j \in N = \{1,\dots,n\}, b_i = F_2(y_j) - F_2(y_j-), j \in N = \{1,\dots,n\}, \sum_{i\in M} a_i = \sum_{j\in N} b_j = 1, \sigma_{ij} = \Gamma(x_i, y_j)$, then

$$F^*(x_i, y_j) = \sum_{r=1}^{i}\sum_{s=1}^{j} p_{rs},$$

where p_{rs} are recursively defined:

$$p_{11} = \min(a_1, b_1, \sigma_{11});$$

$$p_{ij} = \min\left\{a_i - \sum_{s=1}^{j-1} p_{is}, b_j - \sum_{r=1}^{i-1} p_{rj}, \sigma_{ij} - \sum_{\substack{r\le i\\(r,s)\ne(i,j)}}\sum_{s\le j} p_{rs}\right\}$$

if p_{rs} is determined for $r \le i < m$ and $s \le j < n$; and

$$p_{ij} = \min\left\{a_i - \sum_{s=1}^{j-1} p_{is}, b_j - \sum_{r=1}^{i-1} p_{rj}\right\}$$

if $i = m$ or $j = n$.

(d) $F(x,y)$ can be viewed as the analogue of the upper Fréchet bound in the set $\mathcal{F}(F_1, F_2)$ under the side constraint $F^*(x,y) \le \Gamma(x,y)$. To obtain a similar analogue for the lower Fréchet bound, consider

$$\max\{G(x,y) : G(x_i, y_i) \le \Delta(x_i, y_i), i \in I, G \in \mathcal{G}(F_1, F_2)\},$$

where $\mathcal{G}(F_1, F_2)$ is the set of all probabilities

$$G(x,y) = G_\mu(x,y) = \mu((-\infty, x] \times [y, \infty)),$$

$x, y \in \mathbb{R}$ of probability measures μ having marginal *df*'s F_1 and F_2, and where Δ determines a positive measure δ by $G_\delta = \Delta$. Then the above maximum is attained at

$$\tilde{G}(x,y) = \inf_{\substack{x_i\le x\\ y_i\ge y}} \{\Delta(x_i, y_i) + F_1(x) - F_1(x_i) + F_2(y_i-) - F_2(y-)\} \wedge F_1(x) \wedge (F_2(y_i-) - F_2(y-)) \tag{5.6}$$

if and only if $\Delta(x_i, y_i) \ge \max(0, F_1(x_i) - F_2(y_i-)), i \in I$ and $\tilde{G}$ generates a measure. If $\{(x_i, y_i), i \in I\} = \mathbb{R}^2$, then $\tilde{G}$ defines an optimal measure $\tilde{\mu}$ by $G_{\tilde{\mu}} = \tilde{G}$. Moreover under the same regularity conditions as in (b)

$$\sup\left\{\int c(x,y)\mu(dx,dy); G_\mu \in \mathcal{G}, G_\mu(x,y) \le \Delta(x,y), x, y \in \mathbb{R}\right\} = \int c(x,y)\tilde{\mu}(dx,dy),$$

(cf. [7] and Theorem 8).

(e) Consider the discrete version of the extremal problem in (d): Find

$$\max\left\{\sum_{i\in M}\cdot\sum_{j\in N} c_{ij}p_{ij}, \text{ subject to } \sum_{j\in N} p_{ij} = a_i \sum_{i\in M} p_{ij} = b_j \text{ and } \sum_{r\le i}\sum_{s\ge j} p_{rs} \le \Delta(x_i, y_j), i \in M, j \in N\right\},$$

where

$$\sum_{j\in N} b_j = \sum_{i\in M} a_i = 1.$$

Then the solution is determined by

$$\tilde{G}(x_i, y_j) = \sum_{r=1}^{i}\sum_{s=j}^{n} p_{rs}$$

$$= \min_{\substack{1\le r\le i\\ j\le s\le n}} \{\Delta(x_i, y_j) + (a_{r+1} + \cdots + a_i) + (b_j + \cdots + b_{s-1})\} \wedge \sum_{r=1}^{i} a_r \wedge \sum_{s=j}^{n} b_s,$$

or in other words by the following greedy algorithm:

$$p_{1n} = \min\{a_i, b_n, \Delta(x_1, y_n)\},$$

$$p_{ij} = \min\left\{a_i - \sum_{s=j+1}^{n} p_{is}, b_j - \sum_{r=1}^{i-1} p_{rj}, \Delta(x_i, y_j) - \sum_{\substack{r\le i\\ (r,s)\neq(i,j)}}\sum_{s\ge j} p_{rs}\right\},$$

if p_{rs} is determined for $r \le i \le m-1$ and $s \ge j > 1$; and

$$p_{ij} = \min\left\{a_i - \sum_{s=j+1}^{n} p_{is}, b_j - \sum_{r=1}^{i-1} p_{rj}\right\}$$

if $i = m$ or $j = 1$ (cf. [7]). □

Consider more generally a finite measure μ on $(\mathbb{R}^2, \mathcal{B}^2)$ and define for two probability measures P_1, P_2 on $(\mathbb{R}^1, \mathcal{B}^1)$ and $A_i \times B_i \in \mathcal{B}^1 \otimes \mathcal{B}^1, i \in I$,

(5.7) $$M^\mu(P_1, P_2) = \{P \in M^1(P_1, P_2); P(A_i \times B_i) \le \mu(A_i \times B_i), i \in I\},$$

where $M^1(P_1, P_2)$ denotes the set of all probability measures P on $\mathbb{R}^2$ with marginals P_1, P_2. As in (5.2), we assume

(5.8) $$\mu(A_i \times B_i) \ge (P_1(A_i) + P_2(B_i) - 1)_+.$$

THEOREM 9. *Under assumption* (5.8) *define*

(5.9) $$P^*(A \times B) = \inf_{\substack{A_i\subset A\\ B_i\subset B}} \{\mu(A_i \times B_i) + (P_1(A) - P_1(A_i)) + (P_2(B) - P_2(B_i))\} \wedge \min(P_1(A), P_2(B), A, B \in \mathcal{B}^1.$$

Then

(5.10) $$h_\mu(A \times B) := \sup\{P(A \times B); P \in M^\mu(P_1, P_2)\} \le P^*(A \times B).$$

If P^* defines a measure, then

(5.11) $$h_\mu(A \times B) = P^*(A \times B) \quad \textit{and } P^* \textit{ is a solution of } (5.9).$$

The proof of Theorem 9 is similar to that of Theorem 8. In contrast to Theorem 8 it allows us to consider "local" bounds in the transportation problem. Observe that in the finite discrete case bounds of the type

(5.12) $$x_{ij} \le \mu_{ij} \quad \text{for some } (i, j)$$

are of this "local" type. So far in the literature there are no results concerning the solution of problem (5.6) respectively (5.12) with local bounds.

6. Local bounds for the transportation plans. While in the preceding sections the additional constraints were formulated mainly in terms of the *df*'s we now consider local constraints formulated for the densities. These restrictions of the local type of course are in some respect much stronger than those in §2 and generally they are much more difficult to handle.

Our first result deals with a transportation problem with the cost function

$$c(x,y) = I(x \neq y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{if } x = y; \end{cases} \tag{6.1}$$

i.e., the cost of transportation is one for any unit that has to be moved and zero otherwise. c does not satisfy a Monge-type condition. We formulate this problem in a general measure space $(S, \mathcal{U})$ only assuming that

$$\{(x,y) : x \neq y\} \in \mathcal{U} \otimes \mathcal{U}. \tag{6.2}$$

Let $M_f(S), M_f(S \times S)$ denote the set of all finite measures on $(S, \mathcal{U})$ respectively. $(S \times S, \mathcal{U} \otimes \mathcal{U})$ and let for $\mu \in M_f(S \times S), \pi_i \mu, i = 1, 2$, denote the marginals of μ. This transportation problem leads to an extension of Dobrushins result on optimal couplings.

THEOREM 10. (Optimal couplings with local restrictions). *Assume that* (6.2) *holds and let* $\mu_1, \mu_2 \in M_f(S)$ *with* $\mu_1(S) \leq \mu_2(S)$. *Then*

(a)

$$\begin{aligned} &\inf\{\mu((x,y); x \neq y); \mu \in M_f(S \times S), \pi_1\mu \geq \mu_1, \pi_2\mu \leq \mu_2\} \\ &= \lambda^-(S) := \sup_{C \in \mathcal{U}} (\mu_1(C) - \mu_2(C)). \end{aligned} \tag{6.3}$$

(b) *The infimum in* (6.3) *is attained for*

$$\mu^*(A \times B) = \gamma(A \cap B) + \frac{\lambda^-(A)\lambda^+(B)}{\lambda^+(S)}, \tag{6.4}$$

where $\lambda^+(A) = \sup_{C \subset A}(\mu_2 - \mu_1)(C), \lambda^-(A) = \sup_{C \subset A}(\mu_1 - \mu_2)(C)$ *and* $\gamma(A) = \mu_2(A) - \lambda^+(A) = \mu_1(A) - \lambda^-(A)$.

Proof. For any $\mu \in M_f(S \times S), \mu(x \neq y) \geq \sup_C \mu(C \times (S \backslash C)) = \sup_C\{\mu(C \times S) - \mu(C \times C)\} \geq \sup_C\{\mu(C \times S) - \mu(S \times C)\} \geq \sup_C\{\mu_1(C) - \mu_2(C)\} = \sup_C\{\lambda^-(C) - \lambda^+(C)\} = \lambda^-(\operatorname{supp} \lambda^-) = \lambda^-(S)$. On the other hand, $\mu^*(A \times S) = \gamma(A) + \lambda^-(A)\lambda^+(S)/\lambda^+(S) = \mu_1(A)$ and $\mu^*(S \times B) = \gamma(B) + \lambda^-(S)\lambda^+(B)/\lambda^+(S) \leq \gamma(B) + \lambda^+(B) = \mu_2(B)$ and $\mu^*(x \neq y) = \int I(x \neq y)(\gamma(dx, dy) + \lambda^-(dx)\lambda^+(dy)/\lambda^+(S)) = \int I(x \neq y) \lambda^-(dx)\lambda^+(dy)/\lambda^+(S) = \lambda^-(S)\lambda^+(S)/\lambda^+(S) = \lambda^-(S)$. □

Consider next finite measures μ_1, μ_2 on $\mathbb{R}$ with densities h_1, h_2 with respect to a dominating measure μ on $\mathbb{R}^1$. Define

$$\mathcal{P}_{\mu_1}^{\mu_2} := \{P \in M^1(\mathbb{R}^2, \mathcal{B}^2); \pi_1 P \geq \mu_1, \pi_2 P \leq \mu_2\}. \tag{6.5}$$

Any $P \in \mathcal{P}_{\mu_1}^{\mu_2}$ has marginals $P_1 = \pi_1 P, P_2 = \pi_2 P$ with densities $f_1 \geq h_1$ and $f_2 \leq h_2$ with respect to μ. We assume first that $1 = \mu_1(\mathbb{R}^1) \leq \mu_2(\mathbb{R}^1)$, i.e., μ_1 is a probability measure and so $f_1 = h_1$.

PROPOSITION 11. *Define* $z_0 = \inf\{y : \int_{(y,\infty)} h_2 d\mu \leq 1\}$,

$$f_2^*(y) = \begin{cases} h_2(y) & \text{if } y > z_0 \\ \dfrac{1 - \int_{(z_0,\infty)} h_2(u)du}{\mu(z_0)} & \text{if } y = z_0 \text{ and } \mu\{z_0\} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.6}$$

and P^* the corresponding probability measure with μ-density f_2^*.

(a) $\sup\{\bar F_P(x,y); P \in \mathcal{P}_{\mu_1}^{\mu_2}\} = 1 - \max(F_{\mu_1}(x), F_{P^*}(y))$, *for all* x, y, *where* $\bar F_P(x,y) = P([x,\infty) \times [y,\infty))$ *is the survival function.*

(b) *The sup in* (a) *is attained for the distribution* $F^* = F_{X^*,Y^*}$, *where* $X^* = F_{\mu_1}^{-1}(U)$, $Y^* = F_{P^*}^{-1}(U)$.

(c) *If c is a cost function, which is componentwise antitone and satisfies the Monge condition, then*

$$\inf\left\{\int c(x,y)dF_P(x,y); P \in \mathcal{P}_{\mu_1}^{\mu_2}\right\} = \int c(x,y)dF^*(x,y). \tag{6.7}$$

Proof. (a), (b) For $P \in \mathcal{P}_{\mu_1}^{\mu_2}$ with marginals F_{μ_1}, G_2 we know that $\bar F_P(x,y) \le P(F_{\mu_1}^{-1}(U) \ge x, G_2^{-1}(U) \ge y) = P(U \ge \max(F_{\mu_1}(x), G_2(y))) = 1 - \max(F_{\mu_1}(x), G_2(y))$. By our construction of P^* we see that $F_{P^*}(y) \le G_2(y)$, for all y, and therefore, $\bar F_P(x,y) \le 1 - \max(F_{\mu_1}(x), F_{P^*}(y))$.

(c) The conditions on the cost function c were considered in [11]. In that terminology $-c$ is a Δ-monotone function. Therefore, (c) follows from (a), (b), and [11]. □

The "antitone" assumption in (c) of Proposition 11 does not have a good interpretation in terms of costs. Under some additional assumptions on the bounding measures we can construct solutions for more natural cost functions. Let again μ_1 have densities h_i with respect to μ, $1 = \mu_1(\mathbb{R}^1) \le \mu_2(\mathbb{R}^1)$.

THEOREM 12. *Assume that for some* $y_0 \in \mathbb{R}_1$ *we have*

$$h_1(u) \le h_2(u) \quad \text{for } u < y_0 \text{ and } h_1(u) \ge h_2(u) \text{ for } u \ge y_0. \tag{6.8}$$

Define $x_0 = \inf\{y : \int_{(y,\infty)} h_1(u)d\mu(u) \ge \int_{(y,\infty)} h_2(u)d\mu(u)\}$ *and define*

$$f_2(u) := \begin{cases} h_2(u) & \text{if } u > x_0 \\ \dfrac{\int_{[x,\infty)} h_1(u)d\mu(u) - \int_{(x_0,\infty)} h_2(u)d\mu(u)}{\mu(x_0)} & \text{if } u = x_0 \text{ and } \mu\{x_0\} > 0, \\ h_1(u) & \text{if } u < x_0. \end{cases} \tag{6.9}$$

Then for any cost function c satisfying the Monge condition (1.3) *and the unimodality condition* (2.16) *holds:*

$$\inf\left\{\int c(x,y)dF_P(x,y); P \in \mathcal{P}_{\mu_1}^{\mu_2}\right\} = \int_0^1 c(F_{\mu_1}^{-1}(u), F_2^{-1}(u))du, \tag{6.10}$$

where F_2 is the df of the measure with density f_2 with respect to μ. The optimal distribution is induced by the rv's $X^* = F_{\mu_1}^{-1}(U), Y^* = F_2^{-1}(U)$.

Proof. For any $P \in \mathcal{P}_{\mu_1}^{\mu_2}$ with marginals F_{μ_1}, G_2, we have by the Monge condition: $\int c(x,y)dF_P(x,y) \ge \int_0^1 c(F_{\mu_1}^{-1}(u), G_2^{-1}(u))du$. By our construction of F_2 we find that

$$\begin{aligned} &G_2(y) \ge F_2(y) \ge F_{\mu_1}(y) \quad \text{for all } y \ge x_0 \quad \text{and} \\ &F_2(y) = F_{\mu_1}(y) \quad \text{for all } y \le x_0; \end{aligned} \tag{6.11}$$

(6.11) implies that $F_{\mu_1}^{-1}(u) \ge F_2^{-1}(u) \ge G_2^{-1}(u)$ for $u > F_2(x_0)$ and $F_2^{-1}(u) = F_{\mu_1}^{-1}(u)$ for $u \le F_2(x_0)$. Our assumptions on c imply that $c(F_{\mu_1}^{-1}(u), G_2^{-1}(u)) \ge c(F_{\mu_1}^{-1}(u), F_2^{-1}(u))$ for all u. □

Remark 7. It is not difficult to extend the solution of Theorem 12 to the case $\mu_1(\mathbb{R}^1) < 1$ and the conditions $f_1 \geq h_1$, $f_2 \leq h_2$, for the densities of an admissible plan P, if we still have assumption (6.8). Again choose x_0 as in (6.9) and define

$$(6.12) \qquad f_2(x) = \begin{cases} h_2(x), x > z_0, \\ \dfrac{1 - \int_{(z_0,\infty)} h_2(x)d\mu(x)}{\mu(z_0)} & \text{if } x = z_0 \text{ and } \mu(z_0) > 0, \\ 0 & \text{otherwise,} \end{cases}$$

where $z_0 = \inf\{y : \int_{(y,\infty)} h_2(x)d\mu(x) \leq 1\}$. Define $y_0 = \inf\{y : \int_{(y,\infty)} h_2(x)d\mu(x) \leq \int_{(y,\infty)} h_1(x)d\mu(x)\}$,

$$(6.13) \qquad f_1(x) = \begin{cases} h_1(x) & \text{if } x > y_0, \\ f_2(x) & \text{if } x < y_0, \\ \dfrac{\int_{[y_0,\infty)} (h_2(x) - h_1(x))d\mu(x)}{\mu(y_0)} & \text{if } \mu(y_0) > 0. \end{cases}$$

Then we have for c as in Theorem 12

$$(6.14) \quad \inf\left\{\int c(x,y)dF_P(x,y); \pi_1 P \geq \mu_1, \pi_2 P \leq \mu_2\right\} = \int_0^1 c(F_1^{-1}(u), F_2^{-1}(u))du,$$

where F_i have densities f_i with respect to μ, $i = 1, 2$. □

Acknowledgments. We are indebted to the referees for several valuable suggestions, among them the relevance of the transportation problem with fixed sum of marginals as considered in §3 for the "fractional b-matching problem" in the context of undirected graphs (cf. Neuhauser, G. L. and Wolsey, L. A. (1988), *Integer and Combinatorial Optimization*, Wiley). The authors are grateful to Michel Balinski for many helpful discussions.

REFERENCES

[1] E. J. ANDERSON AND P. NASH, *Linear Programming in Infinite Dimensional Spaces. Theory and Applications*, Wiley, New York, 1987.

[2] E. R. BARNES AND A. J. HOFFMAN, *On transportation problems with upper bounds on leading rectangles*, SIAM J. Discr. Meth., 6 (1985), pp. 487–496.

[3] M. S. BAZARAA AND J. J. JAVIS, *Linear Programming and Network Flows*, Wiley, New York, 1977.

[4] S. CAMBANIS, G. SIMONS, AND W. STOUT, *Inequalities for $Ek(X,Y)$ when the marginals are fixed*, Z. Wahrsch. verw. Gebiete 36 (1976), pp. 285–294.

[5] G. DALL'AGLIO, *Fréchet classes and compatibility of distribution functions*. Sympos. Math., 9 (1972), pp. 131–150.

[6] A. HOFFMAN AND A. VEINOTT, JR., *Staircase transportation problems with superadditive rewards and cumulative capacities*, Math. Programming, 1992, to appear.

[7] I. OLKIN AND S. T. RACHEV, *Marginal problems with additional constraints*, Tech. report 270, Dept. of Statistics, Stanford Univ., Stanford, CA, 1990.

[8] S. T. RACHEV, *The Monge-Kantorovich mass transference problem and its stochastic applications*, Theory Probab. Appl., 29 (1984), pp. 647–676.

[9] ———, *Extreme functionals in the space of probability measures*, LNM 1155, 320–348, Springer-Verlag, 1985.

[10] S. T. RACHEV AND R. M. SHORTT, *Duality theorems for Kantorovich-Rubinstein and Wasserstein functionals*, Dissertationes Math. CCXCIX (1990).

[11] L. RÜSCHENDORF, *Inequalities for the expectation of Δ-monotone functions*, Z. Wahrsch. verw. Gebiete, 54 (1980), pp. 341–349.

[12] A. H. TCHEN, *Inequalities for distributions with given marginals*, Ann. Probab., 8 (1980), pp. 814–827.

[13] W. WHITT, *Bivariate distributions with given marginals*, Ann. Statist., 4 (1976), pp. 1280–1289.