Stochastic Orders and Decision under Risk
IMS Lecture Notes – Monograph Series (1991)

BOUNDS FOR DISTRIBUTIONS WITH MULTIVARIATE MARGINALS

BY LUDGER RÜSCHENDORF

University of Münster

Some explicit results and bounds are derived for integrals of functions on product spaces assuming the knowledge of certain multivariate marginals of the underlying distribution. One class of bounds is obtained by using a general reduction principle and the relation to Bonferroni type bounds. A further method is to reduce the problem by conditioning to a problem with simpler marginal constraints. It is proved that one can reduce general decomposable marginal systems to the case of series and star-like systems. For this reason, series and star-like systems are given special consideration. For some nonregular systems, one can derive good bounds by considering all regular subsystems. This method implies, in particular, a characterization or the marginal problem for a circle system of marginals. The special queston of construction of optimal couplings for random vectors w.r.t. the L_p-distance is discussed. This question is related to the investigation of sharp inequalities of the type $f(x) + g(y) \leq |x - y|^p$. Finally, a combinatorial application is given to the support of multidimensional permutation matrices.

1. Introduction. The formal definition of the model with multivariate marginals is the following. Let $S = S_1 \times \cdots \times S_n$ be the product of n Borel spaces with σ-algebra $\mathcal{B} = \bigotimes_{i=1}^n \mathcal{B}_i$, $\mathcal{B}_i$ the Borel σ-algebras on S_i. Let $\mathcal{E} \subset \mathcal{P}\{1, \cdots, n\}$, the system of all subsets of $\{1, \cdots, n\}$, with $\bigcup_{J \in \mathcal{E}} J = \{1, \cdots, n\}$ and let for $J \in \mathcal{E}$, $P_J \in M^1\left(\prod_{j \in J} S_j\right)$ be a consistent system of multivariate distributions on $\pi_J(S) = \prod_{j \in J} S_j =: S_J$, π_J being the J-projection from S to S_J and $M^1(S_J)$ denoting the set of all probability measures on S_J. Consistency means that $J_1, J_2 \in \mathcal{E}$, $J_1 \cap J_2 \neq \emptyset$ implies that $\pi_{J_1 \cap J_2} P_{J_1} = \pi_{J_1 \cap J_2} P_{J_2}$. Define

$$M_{\mathcal{E}} := M(P_J, J \in \mathcal{E}) \tag{1}$$

to be the set of all probability measures on S with marginals P_J, $J \in \mathcal{E}$.

AMS 1980 Subject Classification: 60E15

Key words and phrases: Multivariate marginals, reduction principle, Fréchet bounds, Bonferroni inequality, marginal problem, blocking clutter

Models of this type arise in applications in the following situation. Suppose that concerning a system of n components there are several studies, each study investigating the common behavior of a part J of the components. If the number of measurements in each study is large enough, one obtains in this way the joint distribution P_J of the components in J (at least approximatively) for each J in the system $\mathcal{E}$ of studies available. To make this more concrete, assume that in order to classify a patient into two possible states I, II of illness there are n diagnosis methods, each method i yielding in case I the distribution P_i, in case II the distribution Q_i, $1 \leq i \leq n$. If one does know the joint distributions of the different methods, the relevant model is $M_{\mathcal{E}}$, with $\mathcal{E} = \{\{1\}, \cdots, \{n\}\}$, the system of simple marginals. An interesting question in this situation is whether it is possible to combine the diagnosis methods in order to obtain a better diagnosis than the best one of each of the individual methods. Note that the underlying spaces S_i of the observations may be very different. In this example, e.g. $x_1 \in S_1$ might be a vector in $\mathbb{R}^k$ (a fever curve), x_2 a random set (distribution of particles in blood serum), x_3 a continuous time process (EKG) etc., so that some of the classical methods to determine dependence properties (like correlation, regression, etc.) are not applicable and explaining, therefore, the assumptions made above partially. A main tool for investigating this problem are bounds of the type considered in this paper. For a detailed discussion of this and other statistical and probabilistic applications we refer to the review paper [31].

A more theoretical problem is the question whether $M_{\mathcal{E}} \neq \emptyset$ (the marginal problem). For "decomposable" or "regular" systems $\mathcal{E}$ consistency of $\{P_J; J \in \mathcal{E}\}$ implies that $M_{\mathcal{E}} \neq \emptyset$ (cf. [36], [14], [32]). The simplest indecomposable (nonregular) system is given when $n = 3$ and $\mathcal{E} = \{\{1,2\},\{1,3\},\{2,3\}\}$. Indecomposability of systems is a consequence of the existence of "cycles". In [32] decomposable systems are called regular (simplicial) complexes.

The aim of this paper is to investigate bounds for

$$\begin{aligned} M_{\mathcal{E}}(\varphi) &= \sup\{\int \varphi dP; P \in M_{\mathcal{E}}\} \text{ resp.} \\ m_{\mathcal{E}}(\varphi) &= \inf\{\int \varphi dP; P \in M_{\mathcal{E}}\} = -M_{\mathcal{E}}(-\varphi) \end{aligned} \tag{2}$$

for measurable funtions $\varphi : S \to \mathbb{R}^1$. For $A \in \mathcal{B}_1 \otimes \cdots \otimes \mathcal{B}_n$ we define

$M_{\mathcal{E}}(A) = M_{\mathcal{E}}(1_A)$, $m_{\mathcal{E}}(A) = m_{\mathcal{E}}(1_A)$. Trivially we have

$$M_{\mathcal{E}}(\varphi) \leq U(\varphi) := \inf\left\{\sum_{J\in\mathcal{E}}\int f_J dP_J; \sum_{j\in\mathcal{E}} f_J \circ \pi_J \geq \varphi\right\} \tag{3}$$

and

$$m_{\mathcal{E}}(\varphi) \geq I(\varphi) := \sup\left\{\sum_{J\in\mathcal{E}}\int f_J dP_J; \Sigma f_J \circ \pi_J \leq \varphi\right\}. \tag{4}$$

From duality theory it is know that, under some additional assumptions on φ, equality holds in (3), (4) (cf. [16] and [27], Th. 3, the proof there being valid for bounded continuous φ).

For the case of given one dimensional marginals in $S_i = \mathbb{R}^1$, several bounds and sharp results have already appeared in the literature (cf. [38], [3], [22–25], [27], [35], [8]). In this paper we will restrict mainly to the case of multivariate marginals, for which only few results are known. This problem has a well established applied motivation and an obvious relation to the stochastic ordering of multivariate distributions. But there seem to be no general results on the stochastic ordering w.r.t. larger function classes as e.g. convex functions etc. applicable to this kind of question.

To mention some applications of this type of results, let $\mathcal{E} = \{\{1\},\{2\}\}$ and $S_1 = S_2 = Y$ be a partially ordered space with partial order $\leq$. Let $A = \{(x,y) \in Y \times Y : x \leq y\}$ be closed and $P_1, P_2 \in M^1(S)$, then (cf. Strassen [33]): $M_{\mathcal{E}}(A) = 1$ if and only if $P_1 \leq_{\text{st}} P_2$, where $\leq_{\text{st}}$ denotes the stochastic order. This a.s. representation result has been very influential in applied problems on the ordering of queues, etc. One can determine $M_{\mathcal{E}}(A)$ explicitly in terms of P_1, P_2 generally (cf. (29)). A different type of applications concerns the construction of minimal metrics which are of importance for approximation problems. Two famous examples are the description of the Levy-Prohorov metric as the minimal Ky-Fan metric, due to Strassen [33] and the representation of the minimal L^1-metric (the Kantorovic-Rubinstein theorem) saying that for $\varphi(x,y) = d(x,y)$, $x, y \in Y$, a metric space with metric d, $m_{\mathcal{E}}(\varphi) = \sup\{\int f d(P_1 - P_2); f \in \text{Lip}(Y), f \in \mathcal{L}^1(P_i)\}$, where $\text{Lip}(Y) := \{f : Y \to \mathbb{R}^1; |f(x) - f(y)| \leq d(x,y), \forall x, y\}$ is the set of all Lipschitz-functions. This result is of importance for the transportation problem (cf. [19], [28]). For $Y = \mathbb{R}^1$, $m_{\mathcal{E}}(\varphi)$ is noted Gini's measure of discrepancy. Its properties were studied at the beginning of this century by Gini. It was later shown by Salvemini (1943) and Dall'Aglio (1956) that $m_{\mathcal{E}}(\varphi) = \int |F_1(x) - F_2(x)| dx$, F_i the distribution functions of P_i.

For the notions of orders and related functions (like Schur-convexity, superadditive functions, etc.) we generally refer to the book of Marshall and Olkin; the basic notions of convex analysis can be found in the book of Rockafellar.

In Section 2 we formulate a general reduction principle. As a very special case of this reduction principle the construction of the copula function is obtained. Section 3 includes a very natural explanation of the identity of Fréchet- and Bonferroni-type bounds based on this result. For some cases one can improve the Bonferroni-type bounds by conditional bounds which are discussed in Section 4. The idea of conditional bounds is to reduce a model with marginal constraints by conditioning to a model with simpler constraints. Based on this idea we in particular obtain a characterization of the marginal problem in the simplest indecomposable case. In Section 5 we discuss inequalities of the type $f(x)+g(y) \leq |x-y|^p$. This problem is essential for solving the dual problem for $\varphi(x,y) = |x-y|^p$ (cf. [3]). Finally, in Section 6 we discuss an application to a combinatorial optimization problem.

2. A Reduction Principle. Let $h_i : S_i \to W_i$, $1 \leq i \leq n$, where W_i are also Borel spaces and define $W = \prod_{i=1}^n W_i$,

$$\begin{aligned} h &= (h_1, \cdots, h_n) : S \to W, \\ h(x) &= (h_1(x_1), \cdots, h_n(x_n)), \\ h_J &= (h_j)_{j\in J} : S_J \to W_J. \end{aligned} \tag{5}$$

Then we can formulate the following reduction principle, which will be applied in the following sections to derive bounds for functions of the type $\varphi \circ h$ as e.g. $\Pi h_i(x_i)$ or $\max\{h_i(x_i)\}$.

PROPOSITION 1. With $M_{\mathcal{E}}^h = \{P^h; P \in M_{\mathcal{E}}\}$, P^h the distributon of h w.r.t. P:

(a) $M_{\mathcal{E}}^h = M(P_J, J \in \mathcal{E})^h \subset M(P_J^{h_J}, J \in \mathcal{E})$.

(b) For $\varphi \in \mathcal{L}^1(M_{\mathcal{E}}^h) = \bigcap_{Q\in M_{\mathcal{E}}^h} \mathcal{L}^1(Q)$ we have:

$$\begin{aligned} M_{\mathcal{E}}(\varphi\circ h) &\leq \sup\{\int \varphi dQ; Q \in M(P_J^{h_J}, J \in \mathcal{E})\} \\ m_{\mathcal{E}}(\varphi\circ h) &\geq \inf\{\int \varphi dQ; Q \in M(P_J^{h_J}, J \in \mathcal{E})\}. \end{aligned} \tag{6}$$

(c) If $\mathcal{E} = \{\{1\}, \cdots, \{n\}\}$, then

$$M_{\mathcal{E}}^h = M(P_1^{h_1}, \cdots, P_n^{h_n}) \text{ and equality in (6) holds.} \tag{7}$$

PROOF. The proof of (a) and (b) is obvious. For the proof of (c) let $\mu \in M(P_1^{h_1}, \cdots, P_n^{h_n})$, let $(\Omega, \mathcal{A}, R)$ be a nonatomic probability space and $U : (\Omega, \mathcal{A}, R) \to (W, \sigma(W))$ with $R^U = \mu$. Then $P_j^{h_j} = R^{U_j}$ and by Proposition 1 of [20] there exists a random variable $X_j : (\Omega, \mathcal{A}) \to (S_j, \mathcal{B}_j)$ such that $h_j \circ X_j = U_j[R]$, $1 \le j \le n$, and $R^{X_j} = P_j$. Therefore, $X = (X_1, \cdots, X_n) : (\Omega, \mathcal{A}) \to (S, \mathcal{B})$, $R^X \in M(P_1, \cdots, P_n) = M_{\mathcal{E}}$ and $h(X) = U[R]$. This implies that $\mu = (R^X)^h$ and, therefore, $M(P_1^{h_1}, \cdots, P_n^{h_n}) \subset M(P_1, \cdots, P_n)^h$.

We give a second proof of (c) based on duality theory.

Second PROOF. Considering the set of finite measures supplied with the topology of weak convergence, (7) is equivalent to $M_{\mathcal{E}}(\varphi \circ h) = \sup\{\int \varphi dQ : Q \in M(P_1^{h_1}, \cdots, P_n^{h_n})\}$ for all bounded continuous functions $\varphi : S \to \mathbb{R}^1$, as follows from an application of the Hahn-Banach separation theorem. From duality theory we have

$$M_{\mathcal{E}}(\varphi \circ h) = \inf\left\{\sum_{i=1}^n \int f_i dP_i; f_i \in \mathcal{L}^1(P_i), 1 \le i \le n, \sum_{i=1}^n f_i \circ \pi_i \ge \varphi \circ h\right\}$$
$$= \inf\left\{\sum_{i=1}^n \int g_i \circ h_i dP_i; g_i \circ h_i \in \mathcal{L}^1(P_i), \sum_{i=1}^n g_i \circ \pi_i \ge \varphi\right\}$$

(define $g_i(y) = \inf\{f_i(x); h_i(x) = y\}$ if there exists an x with $h_i(x) = y$ and arbitrary otherwise) $= \inf\left\{\Sigma \int g_i dP_i^{h_i} : \sum_{i=1}^n g_i \circ \pi_i \ge \varphi\right\} = \sup\{\int \varphi dQ;$ $Q \in M(P_1^{h_1}, \cdots, P_n^{h_n})\}$. ∎

REMARK 1. In general one only has an inequality in (a). Let e.g. $S_i = \mathbb{R}^1$, $\mathcal{E} = \{\{1,2\}, \{2,3\}, \{1,3\}\}$ and $P_{12} = P_{13} = P_{23} = R^{(U,1-U)}$, where R^U is the uniform distribution on $(0,1)$ and where $R^{(U,1-U)}$ is the distribution of the pair $(U, 1-U)$ which is uniform on the diagonal $\{(u, 1-u); u \in [0,1]\} \subset [0,1]^2$, then $M(P_{12}, P_{13}, P_{23}) = \emptyset$. For $h_1 = h_2 = \mathrm{id}_{\mathbb{R}^1}$, $h_3 \equiv c \in (0,1)$ holds $P_{12}^{h_{12}} = P_{12}$, $P_{13}^{h_{13}} = R(0,1) \otimes \epsilon_c$, $P_{23}^{h_{23}} = R(0,1) \otimes \epsilon_c$, ϵ_c the one point measure in c, and therefore, $M(P_{12}^{h_{12}}, P_{13}^{h_{13}}, P_{23}^{h_{23}}) \ne \emptyset$. ∎

A constructive proof of the reduction principle can also be given in the following case of a series marginal configuration.

PROPOSITION 2. *Let* $\mathcal{E} = \{\{i, i+1\}, 1 \le i \le n-1\}$, *then*

$$M_{\mathcal{E}}^h = M(P_{ii+1}^{h_{ii+1}}, 1 \le i \le n-1). \tag{8}$$

PROOF. For the proof it is enough to consider the case $n = 1$; the general case follows by induction. Also we may assume w.l.g. that $S_i = W_i = \mathbb{R}^1$,

$1 \le i \le n$. For $\mu \in M(P_{12}^{h_{12}}, P_{23}^{h_{23}})$ holds

$$\mu = \int \mu^{(\pi_1,\pi_2,\pi_3)|\pi_2=\hat{h}_2} dP_2^{h_2}(\hat{h}_2). \tag{9}$$

Since

$$\mu^{(\pi_1,\pi_2,\pi_3)|\pi_2=\hat{h}_2} = \mu^{(\pi_1,\pi_3)|\pi_2=\tilde{h}_2} \otimes \epsilon_{\hat{h}_2}$$

and

$$\mu^{(\pi_1,\pi_3)|\pi_2=\tilde{h}_2} \in M(P_{1|2}^{h_1|h_2=\tilde{h}_2}, P_{3|2}^{h_3|h_2=\tilde{h}_2})$$

we obtain by (7) (cf. also the proof of Theorem 3 in [27]) a Markov kernel $\tilde{\lambda}$ from $(\mathbb{R}^1, \hat{\mathcal{B}}^1)$ to $(\mathbb{R}^2, \mathcal{B}^2)$, $\hat{\mathcal{B}}^1$ the universal completion of $\mathcal{B}^1$ (the intersection of all completions of $\mathcal{B}^1$ w.r.t. all $P \in M^1(\mathbb{R}^1, \mathcal{B}^1)$) such that $\tilde{\lambda}(\tilde{h}_2, \cdot) \in M(P_{1|2}^{\pi_1|h_2=\tilde{h}_2}, P_{3|2}^{\pi_3|h_2=\tilde{h}_2})$ and the image $\tilde{\lambda}(\tilde{h}_2, \cdot)^{(h_1,h_3)} = \mu^{(\pi_1,\pi_3)|\pi_2=\tilde{h}_2}$. Define

$$v := \int \tilde{\lambda}(h_2(x_2), \cdot) \otimes \epsilon_{x_2} dP_2(x_2) \tag{10}$$

then by construction $v \in M(P_{12}, P_{23})$ and by (9) $v^h = \mu$.

Second PROOF. As in Proposition 1 we give a proof by duality theory in the case $n = 2$. By duality

$$\begin{aligned} M_{\mathcal{E}}(\varphi \circ h) = \inf\{\int f_{12} dP_{12} + \int f_{23} dP_{23}; f_{12} \in \mathcal{L}^1(P_{12}) \\ f_{23} \in \mathcal{L}^1(P_{23}), f_{12}(x_1, x_2) + f_{23}(x_2, x_3) \ge \varphi \circ h(x)\}. \end{aligned} \tag{11}$$

Defining $\tilde{f}_{12}(\tilde{h}_1, x_2) = \inf\{f_{12}(x_1, x_2) : h_1(x_1) = \tilde{h}_1\}$, $\tilde{f}_{23}(x_2, \tilde{h}_3) = \inf\{f_{23}(x_2, x_3) : h_3(x_3) = \tilde{h}_3\}$ we see that we can restrict to functions of this type. For $\mu \in M(P_{12}^{(h_1,\pi_2)}, P_{23}^{(\pi_2,h_3)})$ define

$$\begin{aligned} g_{12}(\tilde{h}_1, h_2(x_2)) &= E_\mu(\tilde{f}_{12}(\tilde{h}_1, x_2) \mid h_2(x_2) = \tilde{h}_2) \\ g_{23}(h_2(x_2), \tilde{h}_3) &= E_\mu(\tilde{f}_{23}(x_2, \tilde{h}_3) \mid h_2(x_2) = \tilde{h}_2). \end{aligned} \tag{12}$$

Then

$$\begin{aligned} \int g_{12} dP_{12}^{(h_1,\pi_2)} &= \int E_\mu(\tilde{f}_{12}(\tilde{h}_1, x_2) \mid h_2(x_2) = \tilde{h}_2) dP_{12}^{(h_1,\pi_2)}(\tilde{h}_1, x_2) \\ &= \int E_{P_{12}^{(h_1,\pi_2)}}(\tilde{f}_{12}(\tilde{h}_1, x_2) \mid h_2(x_2) = \tilde{h}_2) dP_{12}^{(h_1,\pi_2)}(\tilde{h}_1, x_2) \\ &= \int \tilde{f}_{12}(\tilde{h}_1, x_2) dP_{12}^{(h_1,\pi_2)}(\tilde{h}_1, x_2) \end{aligned}$$

and, similarly,

$$\int g_{23} dP_{23}^{(\pi_2,h_3)} = \int \tilde{f}_{23}(x_2,\tilde{h}_3) dP_{23}^{(\pi_2,h_3)}(x_2,\tilde{h}_3).$$

Furthermore, g_{12}, g_{23} are admissible since

$$\begin{aligned} g_{12}(\tilde{h}_1,\tilde{h}_2) + g_{23}(\tilde{h}_2,\tilde{h}_3) &= E_\mu(\tilde{f}_{12}(\tilde{h}_1,x_2) + \tilde{h}_{23}(x_2,\tilde{h}_3) \mid h_2(x_2) = \tilde{h}_2) \\ &\geq E_\mu(\varphi(\tilde{h}_1,\tilde{h}_2)) = \varphi(\tilde{h}_1,\tilde{h}_2). \end{aligned}$$

So we have

$$M_{\mathcal{E}}(\varphi \circ h) = \inf\{\int g_{12} dP_{12}^{h_{12}} + \int g_{23} dP_{23}^{h_{23}}; g_{12}\circ\pi_{12} + g_{23}\circ\pi_{23} \geq \varphi\}.$$ ∎

In the next step we consider a star-like configuration.

PROPOSITION 3. *Let $\mathcal{E} = \{(1,j), 2 \leq j \leq n\}$, then*

$$M_{\mathcal{E}}^h = M(P_{1j}^{h_{1j}}, 2 \leq j \leq n). \tag{13}$$

PROOF. As in the first proof of Proposition 2 for $\mu \in M(P_{1j}^{h_{12}}, 2 \leq j \leq n)$, we write

$$\mu = \int \mu^{(\pi_1,\cdots,\pi_n)|\pi_1=\tilde{h}_1} dP_1^{h_1}(\tilde{h}_1)$$

and have

$$\mu^{(\pi_2,\cdots,\pi_n)|\pi_1=\tilde{h}_1} \in M(P_{2|1}^{\pi_2|\pi_1=\tilde{h}_1}, P_{3|1}^{\pi_3|\pi_1=\tilde{h}_1}, \ldots, P_{n|1}^{\pi_n|\pi_1=\tilde{h}_1}).$$

We can then proceed as in the proof of Proposition 2 using Proposition 1 for $n-1$. A duality proof of (13) similar to that of Proposition 2 can also be given. ∎

We now can prove the reduction principle in its general form.

THEOREM 4. (Reduction principle). *If $\mathcal{E}$ is a decomposable (regular) system, then*

$$M_{\mathcal{E}}^h = M(P_J^{h_J}, J \in \mathcal{E}). \tag{14}$$

PROOF. We shall use the terminology on simplical complexes from the paper of Shortt [31]. Since $\mathcal{E}$ is regular, there exists a normal sequence $\mathcal{E} = \mathcal{E}_0 \supset \mathcal{E}_1 \supset \cdots \supset \mathcal{E}_r = \emptyset$ such that $\mathcal{E}_{j+1}$ is normal in $\mathcal{E}_j$, i.e. there exists an extremal simplex $T_j \in \mathcal{E}_j$ such that $\mathcal{E}_{j+1} = \{T' \in \mathcal{E}_j; T' \cap p(|\mathcal{E}_j|, T_j) = \emptyset\}$, where $|\mathcal{E}_j|$ is the vertex set of $\mathcal{E}_j$ and $p(|\mathcal{E}_j|, T_j)$ are the proper vertices of T_j. Let R_j denote the maximal intersecton of T_j with maximal simplices in $\mathcal{E}_{j+1}$

(which is the independent of these simplices since T_j is extremal). We call R_j the splitting set of T_j. Now our proof goes by induction on $r = \ell(\mathcal{E})$ the length of $\mathcal{E}$.

If $r = 1$, then $\mathcal{E}$ is of the type $\mathcal{E} = \langle\{T_1, \cdots, T_k\}\rangle$, where $T_1, \cdots, T_k$ are maximal, pairwise disjoint and $\langle\ \rangle$ denotes the generated (simplicial) complex. So in this case Theorem 4 follows from Proposition 1.

$r \to r+1$. Let $\mathcal{E} = \mathcal{E}_0 \supset \mathcal{E}_1 \supset \cdots \supset \mathcal{E}_{r+1} = \emptyset$ be a complex of length $r+1$. Let $T_0 \in \mathcal{E}_0$ be extremal in $\mathcal{E}_0$ such that $\mathcal{E}_1$ is normal in $\mathcal{E}_0$ w.r.t. T_0 with splitting set R_0. So $\mathcal{E}_1 = \{T' \in \mathcal{E}_0 = \mathcal{E}; T' \cap p(T_0, \mathcal{E}) = \emptyset\}$ is a regular subcomplex of $\mathcal{E}_0$ of length r. If $\mu \in M(P_J^{h_J}, J \in \mathcal{E})$, then by the assumption of induction there exists an element $v \in M(P_J, J \in \mathcal{E}_1)$ such that with $h' = (h_j)_{j \in |\mathcal{E}_1|}$, $v^{h'} = \mu^{\pi|\mathcal{E}_1|}$. With the notation $a = (j)_{j \in p(T_0, \mathcal{E})}$, $b = (j)_{j \in R_0}$ and $c = (j)_{j \in |\mathcal{E}_1| \setminus R_0}$ we have that $\mu \in M(P_{ab}^{h_{ab}}, P_{bc}^{h_{bc}})$ with $P_{bc} = v$, $P_{ab} = P_{T_0}$. Therefore, by Proposition 2 there exists an element $\tau \in M(P_{T_0}, v)$ with $\tau^h = \mu$. So τ solves the problem. ∎

EXAMPLE.

(a) Let $\mathcal{E} = \mathcal{E}_0 = \{345, 234, 278, 12, 19\}$; then we obtain the following normal series, with extremal simplices T_i and splitting sets R_i. $T_0 = \{345\}$, $R_0 = \{34\}$, $\mathcal{E}_1 = \{234, 278, 12, 19\}$, $T_1 = \{234\}$, $R_1 = \{2\}$, $\mathcal{E}_2 = \{278, 12, 19\}$, $T_2 = \{278\}$, $R_2 = \{2\}$, $\mathcal{E}_3 = \{12, 19\}$, $T_3 = \{12\}$, $R_3 = \{1\}$, $\mathcal{E}_4 = \{19\}$, $\mathcal{E}_5 = \emptyset$. So the length of $\mathcal{E}$ is $\ell(\mathcal{E}) = 5$.

(b) The system $\mathcal{E} = \{123, 35, 456, 57, 678\}$ is not regular since it contains a "cycle" $\{456\}$, $\{57\}\{678\}$. ∎

3. Bonferroni-Type Bounds. An interesting special case of the reduction principle in Prop. 1.a) and (6) arises, when $h_i = 1_{A_i}$, $A_i \in \mathcal{B}_i$, $1 \le i \le n$. With $B_i := S_1 \times \cdots \times A_i \times \cdots \times S_n$ and assuming that $\mathcal{E}$ is a complex (i.e. $J_1 \in \mathcal{E}$, $J_2 \subset J_1$ implies that $J_2 \in \mathcal{E}$), it can be seen that for all $P \in M(P_J, J \in \mathcal{E}) = M_{\mathcal{E}}$, we have

$$P(A_1 \times \cdots \times A_n) = P\left(\bigcap_{j=1}^{n} B_j\right), \tag{15}$$

while the information that $Q \in M(P_J^{h_J}, J \in \mathcal{E})$ is, under the validity of the reduction principle, equivalent to the knowledge of

$$p_J = P\left(\bigcap_{j \in J} B_j\right), \quad \text{where } p_J := P_J(A_J),\ A_J = \prod_{j \in J} A_j. \tag{16}$$

So one consequence of Prop. 1.a) and (6) is

$$M_{\mathcal{E}}(A_1 \times \cdots \times A_n) \leq \tilde{M}_{\mathcal{E}}(p_J, J \in \mathcal{E}) := \sup\left\{ P\left(\bigcap_{i=1}^n B_i\right); \right. \\ \left. p_J := P_J(A_J) = P\left(\bigcap_{j \in J} B_j\right), J \in \mathcal{E} \right\}, \tag{17}$$

with equality if the reduction principle is valid. This means that the upper bounds in Prop. 1.a) and (6) are exactly the Bonferroni bounds (of higher order) for sets (B_j). For this reason we call the bounds in (6) Bonferroni-type bounds. But the duality theory, the transition from $M_{\mathcal{E}}(\varphi)$ to $\tilde{M}_{\mathcal{E}}$ corresponds, for $\varphi = \varphi(1_{A_1}, \cdots, 1_{A_n})$, to the transition

$$U(\varphi) \leq \inf \left\{ \sum_{J \in \mathcal{E}} \alpha_J p_J; \sum_{J \in \mathcal{E}} \alpha_J 1_{A_J} \circ \pi_J \geq \varphi, A_J \in \mathcal{B}_J, p_J = P_J(A_J) \right\} \\ =: \tilde{U}(\varphi), \tag{18}$$

which for the special case $\varphi = 1_{A_1 \times \cdots \times A_n}$ reduces to

$$\tilde{U}(\varphi) = \inf \left\{ \sum_{J \in \mathcal{E}} \alpha_J p_J; \sum_{J \in \mathcal{E}} \alpha_J \geq 1, \sum_{\substack{J \subset E \\ J \in \mathcal{E}}} \alpha_J \geq 0, \ \ \forall E \subset \{1, \cdots, n\} \right\}, \tag{19}$$

(cf. Hailperin [11]).

As a consequence of the reduction principle, we obtain an explanation for the identity of the Fréchet-bounds and the Bonferroni bounds of first order proved in [23]. Using Proposition 1 we obtain

PROPOSITION 5. *Assume that $\mathcal{E} = (\{1\}, \cdots, \{n\})$, then:*

(a) *(cf. [23]) If $A_i \in \mathcal{B}_i$, $1 \leq i \leq n$, and $p_i = P_i(A_i)$, then*

$$\sup\{P(A_1 \times \cdots \times A_n); P \in M(P_1, \cdots, P_n)\} = \min\{p_i\}, \\ \inf\{P(A_1 \times \cdots \times A_n); P \in M(P_1, \cdots, P_n)\} = (\Sigma p_i - (n-1))_+, \tag{20}$$

with $x_+ = \max(x, 0)$.

(b) (cf. [22], [35], [25]) If $S_i = W_i = \mathbb{R}^1$ and φ is Δ-monotone or Δ-monotone in pairs, and uniformly integrable w.r.t. $M(P_1, \cdots, P_n)$, then

$$\sup\{\int \varphi \circ h \, dP; P \in M(P_1, \cdots, P_n)\} = \int_0^1 \varphi(F_{h_1}^{-1}(u), \cdots, F_{h_n}^{-1}(u)) du, \tag{21}$$

where F_{h_i} is the df of $P_i^{h_i}$.

(c) (cf. [20]) Define

$$L_k := \bigcup_{\substack{J\subset\{1,\cdots,n\}\\ |J|=k}} \bigcap_{j\in J}\{\pi_j \in A_j\},$$

the event that $\{\pi_j \in A_j\}$ for at least k-components, then

$$\begin{aligned} &\sup\{P(L_k); P \in M(P_1,\cdots,P_n)\} = b_k \\ &\inf\{P(L_k); P \in M(P_1,\cdots,P_n)\} = a_k, \end{aligned} \tag{22}$$

where

$$\begin{aligned} b_k &:= \min\left(1, \frac{1}{k-r}\sum_{i=1}^{n-r} p_{(i)}\right), \\ a_k &= \max\left(0, \frac{\sum_{i=r+1}^{n} p_{(i)} - (k-1)}{n-r-(k-1)}\right), p_{(1)} \le \cdots \le p_{(n)}, \end{aligned}$$

the ordered vector of $(p_1,\cdots,p_n)$. ∎

Δ-monotone functions are essentially measure generating functions (for definition cf. [22]). "Δ-monotone in pairs" means that a function is Δ-monotone when fixing all up to two components. An equivalent notation is "L-superadditive function" (cf. [25]).

We next list some consequences of the reduction principle in connecton with Bonferroni-bounds of higher order. For reference on Bonferroni-bounds, we refer to the books of Galambos [9] and Tong [34].

PROPOSITION 6.

(a) For $A_i \in \mathcal{B}_i$, we have

$$M_{\mathcal{E}}(A_1 \times \cdots \times A_n) \le \min_{J\in\mathcal{B}} P_J(A_J). \tag{23}$$

(b) If $\mathcal{E} = J_k^n = \{T \subset \{1,\cdots,n\}, |T| = k\}$, then $M_{\mathcal{E}}(A_1\times\cdots\times A_n) \le 1 + \sum_{s=1}^{k}(-1)^s q_s$ for $k \in 2\mathbb{N}$,

$$m_{\mathcal{E}}(A_1\times\cdots\times A_n) \ge 1 + \sum_{s=1}^{k}(-1)^s q_s \;\; \textit{for} \;\; k \in 2\mathbb{N} - 1 \tag{24}$$

where $q_s := \sum_{|T|=s} P_T\left(\bigcap_{j\in T} A_j^c\right)$.

(c)

$$m_{\mathcal{E}}(A_1\times\cdots\times A_n)\geq\begin{cases}\sum_{s=1}^{n-1}(-1)^{s+1}r_s-1 & \textit{for } n\in 2\mathbb{N}\\ \sum_{s=1}^{n-1}(-1)^{s+1}r_s-\min_{|T|=n-1}P_T(A_T) & \textit{if } n\in 2\mathbb{N}-1\end{cases}$$

where $r_s:=\sum_{|T|=s}P_T(A_T)$.

(d) If $\mathcal{E}=J_2^n$, $q_i:=P_i(`A_i^c)$, $q_{ij}:=P_{ij}(A_i^c\times A_j^c)$, *then*

$$\begin{aligned}M_{\mathcal{E}}(A_1\times\cdots\times A_n)&\leq 1-\sum_{i=1}^n q_i+\sum_{i<j}q_{ij}\\ m_{\mathcal{E}}(A_1\times\cdots\times A_n)&\geq 1-\sum_{i=1}^n q_i+\sup_{\mathcal{T}}\sum_{(i,j)\in\mathcal{T}}q_{ij},\end{aligned}\tag{25}$$

where the supremum is on the set of all spanning trees of the complete graph with n nodes.

Proof. (a) is immediate from (18); (b) are the usual Bonferroni-bounds. (c) the inductive bounds in (c) are a modification of Poincaré's Theorem (c.f. Warmuth [37]). (d) follows from (18), since for $P\in M_{\mathcal{E}}$ and with $\varphi:=\max_{1\leq i\leq n}1_{\{\pi_i\in A_i^c\}}=1_{\cup\{\pi_i\in A_i^c\}}$ we have $P(A_1\times\cdots\times A_n)=1-P(\cup\{\pi_i\in A_i^c\})=1-\int\varphi dP$ and

$$\begin{aligned}M_{\mathcal{E}}(\varphi)&\leq\tilde{M}_{\mathcal{E}}(\varphi)=\tilde{U}(\varphi)\\ &=\inf\Big\{\Sigma\alpha_ip_i-\Sigma\alpha_{ij}p_{ij};\forall I\subset\{1,\cdots,n\}:\sum_{i\in J}\alpha_i-\sum_{\substack{i<j\\ i,j\in J}}\alpha_{ij}\geq 1\Big\}\\ &\leq\inf\Big\{\Sigma\alpha_ip_i-\Sigma\alpha_{ij}p_{ij};\alpha_{ij}\geq 0,\sum_{i\in J}\alpha_i-\sum_{\substack{i<j\\ i,j\in J}}\alpha_{ij}\geq 1,\forall J\Big\}\\ &=\Sigma p_i-\sup_{\mathcal{T}}\sum_{(i,j)\in\mathcal{T}}p_{ij}\end{aligned}\tag{26}$$

c.f. [18], Prop. 1, and [12]). An even simpler and more direct proof of (d) is mmediate from the inequality

$$1_{\bigcup_{i=1}^n B_i}\leq\sum_{i=1}^n 1_{B_i}-\sum_{(i,j)\in\mathcal{T}}1_{B_i\cap B_j},\tag{27}$$

holding for all trees τ, where a spanning tree is a system of $n-1$ edges (i,j) containing each of the n nodes of a graph. ∎

REMARK 2.

(a) For $S_j = \mathbb{R}^1$, $A_j = (-\infty, x_j]$, $P \in M_{\mathcal{E}}$ follows from (23):

$$F_P(x) \le \min_{J\in\mathcal{E}} F_J(x_J), F_J := F_{P_J}. \tag{28}$$

For the case $\mathcal{E} = I_k^n = \{T \subset \{1,\cdots,n\}, |T| = k\}$ (28) was established by Warmuth [37]. The right hand side of (28) does not define a df with marginals P_J (in contrast to the statement in Theorem 2.3 in [37]) and, therefore, (28) cannot be claimed to be a sharp bound as the following example shows.

EXAMPLE. Let $n = 3$, $F^u(x,y,z) = \min(xy, xz, yz)$, $x,y,z \in (0,1)$, i.e. we consider the case $P_{12} = P_{13} = P_{23} = R(0,1) \otimes R(0,1)$. Suppose that F^u is a df; then let X, Y, Z be rv's with df F^u. Since

$$F^u(x,y,z) = \begin{cases} xy & \text{if } x,y \le z \\ xz & \text{if } x,z \le y \\ yz & \text{if } y,z \le x \end{cases}$$

we obtain for all $x,y \subset [0,1]$, $0 = F^u(x,y,z) - F^u(x,y,\max(x,y)) = P(X \le x, Y \le y, Z > \max(x,y))$ implying, that $\max(X,Y) \ge Z$ a.s. Similarly, $\max(X,Z) \ge Y$ a.s. and $\max(Y,Z) \ge X$ a.s. implying $X = Y = Z$ a.s. and, therefore, $F_{(X,Y,Z)}(x,y,z) = \min(x,y,z)$ a contradiction.

(b) Warmuth [36] applied Prop. 6.c) to derive inductive lower bounds for the case $\mathcal{E} = I_k^n$. But again the resulting lower bounds do not define df's and, therefore, the bounds cannot be claimed to be sharp. Similarly, one can also derive inductive upper bounds corresponding to Prop. 6.c). Various other bounds, such as Chung-Erdös bounds, can be applied in specific applications.

(c) Similarly to (d) if $\mathcal{E} = \{(i,i+1); 1 \le i \le n-1\}$, $m_{\mathcal{E}}(A_1 \times \cdots \times A_n) \ge 1 - \sum_{i=1}^n q_i + \sum_{i=1}^{n-1} q_{ii+1}$. More generally for any complex $\mathcal{E}$ (i.e. $J \in \mathcal{E}$, $T \subset J \Rightarrow T \in \mathcal{E}$) in (18) and (19) one can restrict to $\alpha_J \in \{-1,0,1\}$. The minimal admissible $\alpha = (\alpha_J)_{J\in\mathcal{E}}$ are related to some interesting combinatorial configurations. ∎

4. The Method of Conditioning. In Section 2, we have seen that the general decomposable (regular) case of marginal constraints can be reduced to some basic simple configurations namely $M(P_1, \cdots, P_n)$ (which was treated in Section 2), the sequence structure $M(P_{12}, P_{23}, \cdots, P_{n-1,n})$ and the star-like structure $M(P_{1j}, 2 \le j \le n)$. In this section, we will derive (sharp) bounds for these basic structures by the method of conditioning and then apply this method also to some (nonregular) circle structures like $M(P_{12}, P_{23}, P_{13})$.

For $B \in \mathcal{B}_1 \otimes \mathcal{B}_2$ one has the "explicit" result

$$\begin{aligned} U_{12}(B) :&= \sup\{P(B); P \in M(P_1, P_2)\} \\ &= \inf\{P_1(B_1) + P_2(B_2); B \subset B_1 \times S_2 \cup S_1 \times B_2\} \end{aligned} \tag{29}$$

(cf. Strassen [32] and Kellerer [15], Prop. 3.3). Similarly, the dual result holds

$$\begin{aligned} L_{12}(B) :&= \inf\{P(B); P \in M(P_1, P_2\} \\ &= \sup\{P_1(B_1) + P_2(B_2) - 1; B \supset B_1 \times B_2\}. \end{aligned} \tag{30}$$

These bounds imply for any integrable function φ

$$\begin{aligned} U_{12}(\varphi) &= \sup\{\int \varphi dP; P \in M(P_1, P_2)\} \\ &= \sup\left\{\int_0^\infty P(\varphi \ge t)dt - \int_{-\infty}^0 P(\varphi \le t)dt; P \in M(P_1, P_2)\right\} \\ &\le \int_0^\infty U_{12}(\varphi \ge t)dt - \int_{-\infty}^0 L_{12}(\varphi \le t)dt. \end{aligned} \tag{31}$$

For some cases with known sharp bounds cf. Remark 1.b). Now let $\mathcal{E} = \{\{1,2\}, \{2,3\}\}$ be the basic sequence structure and define

$$P_{1|x_2} := P_{12}^{\pi_1|\pi_2 = x_2} \text{ the conditional distribution.} \tag{32}$$

Define for $x_2 \in S_2$, $B \in \mathcal{B}_1 \otimes \mathcal{B}_3$

$$U_{13|x_2}(B) := \inf\{P_{1|x_2}(B_1) + P_{3|x_2}(B_3); B \subset B_1 \times S_2 \cup S_1 \times B_3\} \tag{33}$$

and

$$L_{13|x_2}(B) := \sup\{P_{1|x_2}(B_1) + P_{3|x_2}(B_3); B \supset B_1 \times B_3\}.$$

PROPOSITION 7. *For $B \in \mathcal{B}_1 \otimes \mathcal{B}_2 \otimes \mathcal{B}_3$, $\mathcal{E} = \{\{1,2\},\{2,3\}\}$ we have*

$$\begin{aligned} M_{\mathcal{E}}(B) &= \int U_{13|x_2}(B_{x_2}) dP_2(x_2) \\ m_{\mathcal{E}}(B) &= \int L_{13|x_2}(B_{x_2}) dP_2(x_2) \end{aligned}, \tag{34}$$

B_{x_2} being the x_2-section of B.

Proof. For any $P \in M_{\mathcal{E}}$ we have $P(B) = \int P_{13|x_2}(B_{x_2})dP_2(x_2)$, where

$$P_{13|x_2} = P^{(\pi_1,\pi_3)|\pi_2=x_2} \in M(P_{1|x_2}, P_{3|x_2}).$$

Therefore, $M_{\mathcal{E}}(B) \leq \int U_{13|x_2}(B_{x_2})dP_2(x_2)$ and $m_{\mathcal{E}}(B) \geq \int L_{13|x_2}(B_{x_2}) dP_2(x_2)$. Let, on the other hand, $x_2 \to P^*_{13|x_2}(B_{x_2})$ be a maximum point of $x_2 \to \sup\{Q(B_{x_2});\ Q \in M(P_{1|x_2}, P_{3|x_2})\}$, which can be chosen as a Markov kernel w.r.t. the completed σ-algebras as in [27]. Then $P^* := P^*_{13|x_2} \times P_1(dx_2) \in M_{\mathcal{E}}$ and, therefore, $M_{\mathcal{E}}(B) = \int P^*_{13|x_2}(B_{x_2})dP_2(x_2) = P^*(B)$. ∎

Remark 3.

(a) Similarly to (34) one obtains for $\mathcal{E} = \{\{1,2\},\{2,3\}\}$, $\varphi = \varphi(x_1, x_2, x_3)$

$$M_{\mathcal{E}}(\varphi) = \int U_{13|x_2}\varphi_{x_2}dP_2(x_2) \tag{35}$$

and

$$m_{\mathcal{E}}(\varphi) = \int L_{13|x_2}\varphi_{x_2}dP_2(x_2),$$

φ_{x_2} denoting the section at x_2.

(b) For the case $M(P_{12}, P_{23}, P_{34})$, we now can use the representation

$$M(P_{12}P_{23}, P_{34}) = \bigcup_{Q_{(12)3} \in M(P_{12},P_{23})} M(Q_{(12)3}, P_{34}), \tag{36}$$

and, therefore,

$$M_{\mathcal{E}}(\varphi) = \sup\{\psi(Q_{(12)3}); Q_{(12)3} \in M(P_{12}, P_{23})\} \tag{37}$$

with

$$\psi(Q_{(12)3}) := \sup\{\int \varphi dQ; P \in M(Q_{(12)3}, P_{34})\}.$$

This can be generalized by induction to $P_{12}, \cdots, P_{n-1,n}$, but generally will be difficult to apply. ∎

Proposition 8. *For $\mathcal{E} = \{\{1,j\}, 2 \leq j \leq n\}$ we have with*

$$U_{2,\cdots,n|x_1}\varphi = \sup\{\int \varphi_{x_1}dP; P \in M(P_{2|x_1}, \cdots, P_{n|x_1})\}$$

and

$$L_{2,\cdots,n|x_1}\varphi = \inf\{\int \varphi_{x_1}dP; P \in M(P_{2|x_1}, \cdots, P_{n|x_1})\}$$

$$\begin{aligned} M_{\mathcal{E}} &= \int U_{2,\cdots,n|x_1}\varphi dP_1(x_1) \\ m_{\mathcal{E}} &= \int L_{2,\cdots,n|x_1}\varphi dP_1(x_1). \end{aligned} \tag{38}$$

PROOF. The proof is analogous to that of Proposition 7. ∎

We now turn to the simplest nonregular structure, a circle of length 3. As a consequence of (34) and (35) we obtain

PROPOSITION 9. *If* $\mathcal{E} = \{\{1,2\},\{2,3\},\{1,3\}\}$ *and* $\varphi = \varphi(x_1,x_2,x_3)$, *then*

$$\begin{aligned} M_{\mathcal{E}}(\varphi) \le \min\{&\int U_{23|x_1}(\varphi_{x_1})dP_1(x_1), \\ &\int U_{13|x_2}(\varphi_{x_2})dP_2(x_2), \int U_{12|x_3}(\varphi_{x_3})dP_3(x_3)\} =: U(\varphi) \\ m_{\mathcal{E}}(\varphi) \ge \max\{&\int L_{23|x_1}(\varphi_{x_1})dP_1(x_1), \\ &\int L_{13|x_2}(\varphi_{x_2})dP_2(x_2), \int L_{12|x_3}(\varphi_{x_3})dP_3(x_3)\} = L(\varphi). \end{aligned} \tag{39}$$

PROOF. (39) follows from the relation $M(P_{12},P_{13},P_{23}) = M(P_{12},P_{23}) \cap M(P_{13},P_{23}) \cap M(P_{12},P_{13})$ and Proposition 7. ∎

REMARK 4. For $A = A_1 \times A_2 \times A_3$, $x_1 \in A_1$ holds $U_{23|x_1}(A_{x_1}) = U_{23|x_1}(A_2 \times A_3) = \min(P_{2|x_1}(A_2), P_{3|x_1}(A_3))$ and, therefore,

$$U(A) \le \min\{P_{12}(A_1 \times A_2), P_{13}(A_1 \times A_3), P_{23}(A_2 \times A_3)\} \tag{40}$$

is an improvement of the Bonferroni-bounds in (23). ∎

We now consider the so called marginal problem for a cycle, i.e. the question whether $M(P_{12},P_{23},P_{13}) \neq \emptyset$. Let

$$C(P_{12},P_{23}) = \{P_{13} \in M^1(S_1 \times S_3); M(P_{12},P_{23},P_{13}) \neq \emptyset\} \tag{41}$$

denote the compatibility set of P_{12}, P_{23}. Dall'Aglio [5], [6] has shown that (in the case $S_i = \mathbb{R}^1$) the set of df's F_{13} corresponding to elements $P_{13} \in C(P_{12},P_{23})$ is convex and has a minimum and maximum element

$$\begin{aligned} F_{13}^{\ell}(x_1,x_3) &:= \int \max\{F_{1|x_2}(x_1) + F_{3|x_2}(x_3) - 1, 0\}dP_2(x_2) \\ &\le F_{13}(x_1,x_3) \le F_{13}^{u}(x_1,x_3) := \int \min\{F_{1|x_2}(x_1), F_{3|x_2}(x_3)\}dF_2(x_2). \end{aligned} \tag{42}$$

On the other hand, an example in [5] shows that the inequality (42) does *not* imply that $P_{13} \in C(P_{12}, P_{23})$.

REMARK 5. For general $\mathcal{E}$ it is well known that $M_{\mathcal{E}} \neq \emptyset$ if and only if for bounded, measurable $f_J : S_J \to \mathbb{R}^1$, $J \in \mathcal{E}$

$$\sum_{J \in \mathcal{E}} f_J \circ \pi_J \geq 0 \text{ implies } \sum_{J \in \mathcal{E}} \int f_J dP_J \geq 0 \tag{43}$$

(cf. [14]). But (43) is difficult to verify. On the other hand in some cases a construction is easy. Let e.g. $\mathcal{E} = \{\{1,2\},\{2,3\},\{1,3\}\}$ $P_{12} = P_1 \otimes P_2$, $P_{13} = P_1 \otimes P_3$, then $C(P_{12}, P_{13}) = M^1(S_2 \times S_3)$, since for $P_{23} \in M^1(S_2 \times S_3)$, $\tilde{P}(A) := \int P_{23}(A_{x_1}) dP_1(x_1)$ defines an element from $M(P_{12}, P_{13}, P_{23})$. (For some general constructions cf. [29]). ∎

For $\varphi = \varphi(x_1, x_3)$ define

$$\begin{aligned} U_{13|2}(\varphi) : &= \int U_{13|x_2}(\varphi) dP_2(x_2), \text{ with} \\ U_{13|x_2}(\varphi) : &= \sup\{\int \varphi dP_{x_2}; P_{x_2} \in M(P_{1|x_2}, P_{3|x_2})\}, \end{aligned} \tag{44}$$

and, similarly, $L_{13|2}(\varphi)$. The bounds in Proposition 9 allow to prove a characterization of the marginal problem for a cycle.

THEOREM 10. $P_{13} \in C(P_{12}, P_{23}) \Leftrightarrow \forall \varphi = \varphi(x_1, x_3) \geq 0$ *measurable, bounded we have*

$$L_{13|2}(\varphi) \leq P_{13}(\varphi) \leq U_{13|2}(\varphi) \tag{45}$$

$\Leftrightarrow$ *(45) is satisfied for continuous, bounded* $\varphi \geq 0$.

PROOF. "$\Rightarrow$" follows from Proposition 8. "$\Leftarrow$" $U_{13|x_2}$, $L_{13|x_2}$ are subadditive resp. superadditive functionals. Define

$$\tau_{x_2}(f) := \inf_{g \in B(S_1 \times S_3)} (U_{13|x_2}(f+g) + L_{13|x_2}(-g)),$$

$$\tau(f) := \int \tau_{x_2}(f) dP_2(x_2)$$

for f bounded, measurable. Then the condition (45) is equivalent to $P_{13}(f) \leq \tau(f) = \int \tau_{x_2}(f) dP_2(x_2)$, $\forall f \in B(S_1 \times S_3)$. By Strassen's disintegration theorem (cf. [32]) P_{13} has a representation $P_{13}(\varphi) = \int \tilde{P}_{13|x_2}(\varphi) dP_2(x_2)$ with a Markov kernel $\tilde{P}_{13|x_2}(\varphi) \leq \tau_{x_2}(\varphi)$, $\forall x_2$, $\forall \varphi$ or, equivalently, $L_{13|x_2}(\varphi) \leq$

$\tilde{P}_{13|x_2}(\varphi) \leq U_{13|x_2}(\varphi)$ for $\varphi \geq 0$ bounded, measurable. Define for $A = A_1 \times A_2 \times A_3$

$$P(A) := \int \tilde{P}_{13|x_2}(A_{x_2})dP_2(x_2), \tag{46}$$

then $P_{13|x_2} = \tilde{P}_{13|x_2}$ and for $A = A_1 \times A_2 \times A_3$, $L_{13|x_2}(A_{x_2}) = U_{13|x_2}(A_{x_2} = P_{1|x_2}(A_1)$ for $x_2 \in A_2$ i.e. $P(A_1 \times A_2 \times S_3) = P_{12}(A_1 \times A_2)$. Similarly, $P(S_1 \times B_2 \times B_3) = P_{23}(B_2 \times B_3)$ and by definition $P(B_1 \times S_2 \times B_3) = \int \tilde{P}_{13|x_2}(B_1 \times B_3)dP_2(x_2) = P_{13}(B_1 \times B_3)$ i.e. $P \in M(P_{12}, P_{13}, P_{23})$. ∎

REMARK 6.

(a) In the example of Dall'Aglio [5] it is easy to see that condition (45) is not fulfilled for certain $\varphi = 1_A$. For $\varphi = 1_A$ the upper and lower bounds in (45) were determined explicitly in (32), (33). It is enough to postulate (45) for functions $f_n = \Sigma \alpha_k 1_{A_1^k \times A_2^k}$, $\alpha_k \geq 0$, using approximation properties of $L_{13|2}$, $U_{13|2}$. It is not enough to consider indicator functions $\varphi = 1_B$ in (45) only.

(b) For more general nonregular cases, it seems in analogy to Proposition 9 to be a good strategy to consider good or optimal bounds for the maximal regular subconfigurations of $\mathcal{E}$. ∎

5. Inequalities of the Type $f(x)+g(y) \leq |x-y|^p$ and Minimal L_p-Distances. In this section we consider the problem of determining solutions of

$$\sigma_p(P_1, P_2) = \inf\{Ed^p(X,Y); X \sim P_1, Y \sim P_2\}, \tag{47}$$

where P_1, P_2 are probability mesures on a separable metric space (S,d) and $p \geq 1$. Random variables X, Y, which solve (47) are called optimal couplings w.r.t. the L_p-distance. From the point of view of stochastic orderings, a more ambitious aim would be (in the case $S = \mathbb{R}^k$) to describe the pairs of random vectors (X,Y) such that $X \sim P_1$, $Y \sim P_2$ and $X - Y$ is minimal w.r.t. Schur order (assuming w.l.g. that first moments of P_1, P_2 are identical). It is easy to see that there is not one smallest pair for $k > 1$, while for $k = 1$ the quantile transforms $X = F_X^{-1}(U), Y = F_Y^{-1}(U)$ are solutions. Therefore, considerations on the Schur-ordering do not imply immediately the construction of optimal couplings (for some special results cf. [26]).

The Kantorovich dual representation (cf. [19], [15]) gives

$$\begin{aligned}\sigma_p(P_1,P_2) = \sup\{\int f dP_1 + \int g dP_2; f(x) + g(y) \\ \leq d^p(x,y), f \in \mathcal{L}^1(P_1), g \in \mathcal{L}^1(P_2)\}\end{aligned} \tag{48}$$

and by Theorem 2.21 of [15], there exist solutions f, g of (48) if $\int d^p(x,a)dP_1(x) < \infty$, $\int d^p(y,a)dP_2(y) < \infty$. Define the conjugate f^* and the p-subdifferential $\partial_p f$ of f by

$$\begin{aligned} f^*(y) &:= \inf_x(|x-y|^p - f(x)), \\ \partial_p f(x) &= \{y : f(x) + f^*(y) = |x-y|^p\} \\ \partial_p f^*(y) &= \{x : f(x) + f^*(y) = |x-y|^p\}. \end{aligned} \tag{49}$$

These notions are defined in analogy to the corresponding notions in convex analysis, where $|x-y|^p$ is replaced by $\langle x, y\rangle$. Then one gets

Proposition 11. *Let $\int d^p(x,a)dP_1(x) < \infty$, $\int d^p(y,a)dP_2(y) < \infty$ for some $a \in S$. Then there exists a solution X, Y of (47). (X,Y) is a solution of (47) if and only if $Y \in \partial_p f(X)$ a.s. for some $f \in \mathcal{L}^1(P_1)$ or, equivalently, if and only if $X \in \partial_p g(Y)$ a.s. for some $g \in \mathcal{L}^1(P_2)$.*

Proof. If $Y \in \partial_p f(X)$ for some $f \in \mathcal{L}^1(P_1)$, then for any $\tilde{X} \sim P_1, \tilde{Y} \sim P_2$ we have $Ed^p(\tilde{X},\tilde{Y}) \geq Ef(\tilde{X}) + Ef^*(\tilde{Y}) = E(f(X) + f^*(Y)) = Ed^p(X,Y)$, i.e. (X,Y) is a solution. The converse follows from the existence of solutions of the dual problem from general duality theory. ∎

Therefore, the problem of determining the minimal L_p-metrics is reduced to the determination of the p-subdifferentials or, equivalently, to the discussion of sharp inequalities of the type $f(x) + g(y) \leq |x-y|^p$. For $p = 2$ and $S = H$ a Hilbert space one gets this way a complete characterization of optimal couplings by means of the conjugate duality theory of Rockafellar (cf. [30], [17]).

For general $p \geq 1$ we start the discussion of inequalities as above in the case $S = \mathbb{R}^1$. For monotone, bijective, differentiable functions $\phi : \mathbb{R}^1 \to \mathbb{R}^1$ the question is to prove the existence of functions, $f, g : \mathbb{R}^1 \to \mathbb{R}^1$ such that

$$f(x) + g(y) \leq |x-y|^p \text{ and } f(x) + g(\phi(x)) = |x - \phi(x)|^p, \quad x, y \in \mathbb{R}^1. \tag{50}$$

From (50) we obtain that

$$H(x,y) := |x-y|^p - f(x) - g(y) \geq 0 \tag{51}$$

has for x fixed a minimum point in $y = \phi(x)$, and, therefore,

$$0 = \frac{\partial}{\partial y} H(x,y) = p|x-y|^{p-1} s(y-x) - g'(y)\ |_{y=\phi(x)},$$

where

$$s(x) := \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0 \end{cases}$$

is the sign function and where g is assumed to be differentiable, i.e. $g'(\phi(x)) = p|x - \phi(x)|^{p-1}s(\phi(x) - x)$. This implies with $y = \phi(x)$ $g'(y) = |\phi^{-1}(y) - y|^{p-1}s(y - \phi^{-1}(y))$ or, equivalently,

$$g(y) = c_1 + \int_{\phi(o)}^{y} p|\phi^{-1}(u) - u|^{p-1}s(u - \phi^{-1}(u))du. \tag{52}$$

Similarly, $H(x, y)$ has for fixed y a minimum in $x = \phi^{-1}(y)$, so we get the necessary condition

$$f(x) = c_2 + \int_{o}^{x} p|u - \phi(u)|^{p-1}s(u - \phi(u))du. \tag{53}$$

We choose

$$c_1 + c_2 = |\phi(o)|^p$$

and can state

PROPOSITION 12. *For $p \geq 1$ and ϕ as above there exists exactly (up to constants) one pair of differentiable functions f, g satisfying (50). f, g are given by (52), (53) and (54).*

PROOF. The uniqueness of solutions follows from the discussion above. By the substitution $v = \phi^{-1}(u)$ i.e. $du = \phi'(v)dv$ we obtain

$$g(y) = c_1 - \int_{o}^{\phi^{-1}(y)} p|v - \phi(v)|^{p-1}s(v - \phi(v))\phi'(v)dv. \tag{55}$$

So we obtain

$$\begin{aligned} f(x) + g(\phi(x)) &= \int_{o}^{x} p|u - \phi(u)|^{p-1}s(u - \phi(u))(1 - \phi'(u))du + |\phi(o)|^p \\ &= |\phi(o)|^p + \int_{o}^{x} d(|u - \phi(u)|^p) = |x - \phi(x)|^p. \end{aligned} \tag{56}$$

To prove the inequality

$$H(x, y) = |x - y|^p - f(x) - g(y) \geq 0. \tag{57}$$

it is enough to show that $H(x, \cdot)$ has a local minimum in $y = \phi(x)$, since $y = \phi(x)$ is the only critical point of $H(x, \cdot)$. First consider the case $y < \phi(x)$.

Then with

$$D(y) = \frac{\partial}{\partial y} H(x,y) = p|x-y|^{p-1}s(y-x) - p|\phi^{-1}(y) - y|^{p-1}s(y - \phi^{-1}(y)) \quad (58)$$

we have for $x < y < \phi(x)$, $\phi^{-1}(y) < x < y$, so $D(y) = p|x-y|^{p-1} - p|\phi^{-1}(y) - y|^{p-1} < 0$. If $x > \phi(x) > y$, then $\phi^{-1}(y) < x < \phi^{-1}(x)$, so $D(y) = -p|x - y|^{p-1} - p|\phi^{-1}(y) - y|^{p-1}s(y - \phi^{-1}(y)) < 0$, since for $y \leq \phi^{-1}(y)$ we have $|x - y| = (x - y) > \phi^{-1}(y) - y = |\phi^{-1}(y) - y$, while for $\phi^{-1}(y) < x$ both summands have a negative sign.

In the second case: $y > \phi(x)$ first consider the case $x > y > \phi(x)$. Then $\phi^{-1}(x) > \phi^{-1}(y) > x$ and so $D(y) = -p|x - y|^{p-1} + p|\phi^{-1}(y) - y| > 0$. In the case: $x < \phi(x) < y$ holds: $\phi^{-1}(y) > x > \phi^{-1}(x)$ and so $D(y) = p|x - y|^{p-1} - p|\phi^{-1}(y) - y|^{p-1}s(y - \phi^{-1}(y))$. If $y \leq \phi^{-1}(y)$, then $D(y) > 0$; if $y > \phi^{-1}(y)$, then $|x - y| > |y - \phi^{-1}(y)|$ and, therefore, $D(y) > 0$. Altogether, we found that $H(x, \cdot)$ has a local minimum in $y = \phi(x)$. ∎

REMARK 7. From Proposition 5 we obtain the (well known) consequence that a pair $(X, \phi(X))$ is an optimal coupling w.r.t. σ_p distance for ϕ monotone, bijective, differentiable. If the image of ϕ is any interval $I \subset \mathbb{R}^1$ we can modify the above proof by defining $g(y)$ for $y \in I$ as in (52) and $g(y) = \inf_x(|x - y|^p - f(x))$ otherwise. ∎

Consider now the case $S = \mathbb{R}^k$ and $|x-y| = (\Sigma(x_i-y_i)^2)^{1/2}$ the euclidean distance and let $\phi : \mathbb{R}^k \to \mathbb{R}^k$ be bijective differentiable and satisfy (50) with $x, y \in \mathbb{R}^k$ and f, g differentiable. Then as in (51) – (53) we obtain unique differentiable functions f, g by

$$f(x) = c_1 + \int_{\delta(o\to x)} p|u - \phi(u)|^{p-1}s(u - \phi(u)) \cdot du \quad (59)$$

where $\delta(o \to x)$ is a continuous path from o to x and $s(u - \phi(u)) = (s(u_i - \phi(u)_i))$. Similarly,

$$g(y) = c_1 + \int_{\gamma(\phi(0)\to y)} p|\phi^{-1}(v) - v^{p-1}s(v - \phi^{-1}(v)) \cdot dv \quad (60)$$

$\gamma(\phi(o) \to y)$ a continuous path from $\phi(o)$ to y.

LEMMA 13. *If* $c_1 + c_2 = |\phi(o)|^p$, *then*

$$f(x) + g(\phi(x)) = |x - \phi(x)|^p. \quad (61)$$

PROOF. With the substitution $v = \phi^{-1}(u)$, i.e. $u = \phi(v)$, $du = D\phi(v) \cdot dv$

$$\begin{aligned} g(y) &= c_2 + \int_{\gamma(o \to \phi^{-1}(y))} p|v - \phi(v)|^{p-1} D\phi(v) s(\phi(v) - v) \cdot dv \\ &= c_2 - \int_{\gamma(o \to \phi^{-1}(v))} p|v - \phi(v)|^{p-1} D\phi(v) s(v - \phi(v)) \cdot dv. \end{aligned}$$

Therefore,

$$\begin{aligned} f(x) + g(\phi(x)) &= c_1 + c_2 + \int_{\delta(o \to x)} p|v - \phi(v)|^{p-1} s(v - \phi(v))(1 - D\phi(v)) \cdot dv \\ &= |\phi(o)|^p + \int_{\delta(o \to x)} d(|v - \phi(v)|^p) \\ &= |\phi(o)|^p + |x - \phi(x)|^p - |\phi(o)|^p = |x - \phi(x)|^p. \end{aligned}$$

∎

Using $|x - y|^p - |\phi^{-1}(y) - y|^p = \int_{\delta(\phi^{-1}(y) \to x)} d_u(|u - y|^p) = \int_{\delta(\phi^{-1}(y) \to x)} p|u - y|^{p-1} s(u - y) du$, the inequality

$$\begin{aligned} f(x) + g(y) &= f(x) - f(\phi^{-1}(y)) + f(\phi^{-1}(y)) + g(y) \\ &= |\phi^{-1}(y) - y|^p + \int_{\delta(\phi^{-1}(y) \to x)} p|u - \phi(u)|^{p-1} s(u - \phi(u)) du \le |x - y|^p, \end{aligned} \tag{62}$$

is equivalent to the following conditions on ϕ:

$$f(x) - f(\phi^{-1}(y)) \le |x - y|^p - |\phi^{-1}(y) - y|^p, \quad \forall x, y \tag{63}$$

and also to:

$$\int_{\delta(\phi^{-1}(y) \to x)} [p|u - y|^{p-1} s(u - y) - p|u - \phi(u)|^{p-1} s(u - \phi(u))] du \ge 0, \quad \forall x, y. \tag{64}$$

For $p = 2$ this is satisfied if $\phi = \nabla f$ is the gradient of a convex function $f : \mathbb{R}^k \to \mathbb{R}^1$ (cf. Knott and Smith [17], Th. 2.1). For general p we do not have a corresponding simple condition and so state this as an open problem.

PROBLEM. Find simple conditions on ϕ such that any of the inequalities (62), (63) and (64) is satisfied. Is the condition valid for $p = 2$ also sufficient for any other p?

6. A Combinatorial Application. Let γ_n be the set of permutations of $\{1, \cdots, n\}$ and define for $\pi_1, \cdots, \pi_m \in \gamma_n$

$$S(\pi_1, \cdots, \pi_m) = \{(i, \pi_1(i), \cdots, \pi_m(i));\ 1 \le i \le n\} \tag{65}$$

and

$$\mathcal{R} = \{S(\pi_1, \cdots, \pi_m); \pi_i \in \gamma_n, \ 1 \le i \le m\}, \tag{66}$$

the "clutter" of supports of m-dimensional permutation matrices (the notation clutter meaning that no member of $\mathcal{R}$ is contained in any other member). Finally, let $\gamma = b(\mathcal{R})$ denote the blocking clutter of $\mathcal{R}$ consisting of the minimal subsets of $\{1, \cdots, n\}^{m+1}$ that have nonempty intersection with each element of $\mathcal{R}$. The following theorem extends a theorem of Gross [10] who considered the case $m = 1$. Our proof is related to the Fréchet-bounds. For further results on the support of m-dimensional stochastic matrices cf. [1], [2], [4], [13].

THEOREM 14. *The blocking clutter of $\mathcal{R}$ is given by*

$$\gamma = \left\{ A_1 \times \cdots \times A_{m+1} \subset \{1, \cdots, n\}^{m+1}; \sum_{i=1}^{m+1} |A_i| = mn + 1 \right\}. \tag{67}$$

PROOF. For the proof we proceed by a series of some lemmas.

LEMMA 1. *For $A_i \subset \{1, \cdots, n\}$, $1 \le i \le n + 1$, holds*

a) $\min_{\pi_1, \cdots, \pi_m \in \gamma_n} |S(\pi_1, \cdots, \pi_n) \cap A_1 \times \cdots \times A_{m+1}| = \max\{0, \sum_{i=1}^{m+1} |A_i| - mn\}$;

b) $\max_{\pi_1, \cdots, \pi_m \in \gamma_n} |S(\pi_1, \cdots, \pi_m) \cap A_1 \times \cdots \times A_{m+1}| = \min\{|A_i| : 1 \le i \le m + 1\}$.

PROOF. For $\pi_1, \cdots, \pi_m \in \gamma_n$ holds

$$|S(\pi_1, \cdots, \pi_m) \cap A_1 \times \cdots \times A_{m+1}| = n - |S(\pi_1, \cdots, \pi_m) \cap (A_1 \times \cdots \times A_{m+1})^c|.$$

Since

$$(A_1 \times \cdots \times A_{m+1})^c \subset \sum_{i=1}^{m+1} M \times \cdots \times A_i^c \times \cdots \times M,$$

$M := \{1, \cdots, n\}$, it follows that

$$|S(\pi_1, \cdots, \pi_m) \cap A_1 \times \cdots \times A_{m+1}| \ge n - \sum_{i=1}^{m+1} |A_i^c| = \sum_{i=1}^{m+1} |A_i| - mn.$$

So the right hand side in a) is a lower bound. If, conversely, $\Sigma |A_i| \ge mn$, then it is easy to construct permutations $\pi_1^*, \cdots, \pi_m^*$, such that $|S(\pi_1^*, \cdots, \pi_m^*) \cap A_1 \times \cdots \times A_{m+1}| = \max\{0, \Sigma|A_i| - mn\}$ observing that one can find $n - |A_i|$ points in M^{m+1} with ith components $|A_i| + 1, \cdots, n$ and jth components smaller than or equal to $|A_j|$, $j \ne i$. The proof of b) is similar. ∎

Note that Lemma 1 can also be deduced from the Fréchet-bounds (cf. (20), with $p_i = |A_i|/n$).

Next define for $1 \leq r \leq m+1$, $E_{r,k}$ to be the set of all elements of $\{1,\cdots,n\}^{m+1}$ with fixed rth coordinate k, $1 \leq k \leq n$ (i.e. $E_{r,k}$ is the r-hyperplane in point k). The proof of the following $m+1$-dimensional version of Hall's theorem on systems of distinct representatives is obvious from Theorem 5.4 of Jurkat/Ryser [13], who considered the case $m = 2$.

LEMMA 2. *Let $A \subset \{1,\cdots,n\}^{m+1}$. There exist $\pi_1,\cdots,\pi_m \in \gamma_n$ such that $S(\pi_1,\cdots,\pi_m) \subset A$ if and only if for any $1 \leq r \neq r' \leq m+1$ and $1 \leq k_1,\cdots,k_\ell \leq n$, there are $1 \leq k'_1,\cdots,k'_\ell \leq n$ and $x_1,\cdots,x_\ell \in A$ with $x_i \in E_{r,k_1} \cap E_{r',k'_i}$, $1 \leq i \leq \ell$.* ∎

LEMMA 3. *If $A \subset \{1,\cdots,n\}^{m+1}$ and $A \cap S(\pi_1,\cdots,\pi_m) \neq \emptyset$, $\forall \pi_i \in \gamma_n$, then there exist $A_i \subset \{1,\cdots,n\}$, $1 \leq i \leq m+1$ with $A_1 \times \cdots \times A_{m+1} \subset A$* and $A_1 \times \cdots \times A_{m+1} \cap S(\pi_1,\cdots,\pi_m) \neq \emptyset$ for $\pi_i \in \gamma_n$.

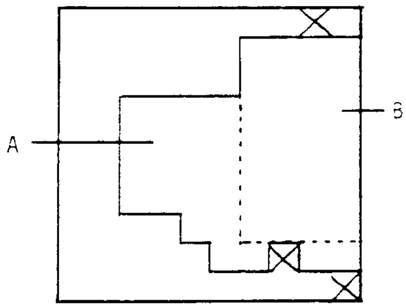
PROOF. By assumption there is no $\pi_1,\cdots,\pi_m \in \gamma_n$ with $S(\pi_1,\cdots,\pi_m) \subset A^c$. Therefore, by Lemma 2 there are $r \neq r'$ and hyperplanes $E_{r,k_1},\cdots,E_{r,k_\ell}$ but no ℓ r'-hyperplanes as in Lemma 2 (with A^c now). Let ℓ be minimal with this property. Then there are $x_1,\cdots,x_{\ell-1} \in A^c$ and $k'_1,\cdots,k'_{\ell-1}$ with $x_i \in E_{r,k_i} \cap E_{r',k'_i}$, $1 \leq i \leq \ell-1$. Define $B := A \cap \bigcup_{i=1}^{\ell}(E_{r,k_i} \cap \bigcup_{j \neq k'_i} E_{r',j})$; then B is a product set since otherwise one could find $k'_\ell \neq k_i$, $i \leq \ell-1$ and $x_\ell \in E_{r,k_\ell} \cap E_{r',k'_\ell}$ in contradiction to the construction of ℓ. Clearly, for B and $k_1,\cdots,k_\ell$ as introduced above there are also no $k'_1,\cdots,k'_\ell$ and $x_1,\cdots,x_\ell \in B^c$ such that $x_i \in E_{r,k_i} \cap E_{r,k'_i}$, $1 \leq i \leq \ell$, and, therefore, by Lemma 2, B^c does not contain the support of an m-dimensional permutation. ∎

As an illustration of the method of proof of Lemmas 2 and 3 consider the following example with $n = 10$

A the region in the boundary, $B = \{2,\cdots,8\} \times \{7,8,9,10\}$, $k_1 = 10$, $k_2 = 9$, $k_3 = 8$, $k_4 = 7$, $k'_1 = 10$, $k'_2 = 1$, $k'_3 = 9$.

Combining Lemma 1 and Lemma 3, we obtain that minimal blocking sets are product sets $A_1 \times \cdots \times A_{m+1}$ with $\sum_{i=1}^{m+1} |A_i| = mn+1$, i.e. the proof of Theorem 14.

As a corollary to Theorem 14 and to the combinatorial duality Theorem of Edmonds and Fulkerson [7] we obtain



Corollary 15. *Let* $a_{i_1,\cdots,i_{m+1}}$, $1 \leq i_j \leq n$, *be a real array, then*

$$\max_{\pi_1,\cdots,\pi_m \in \gamma_n} \min_{1 \leq i \leq n} a_{i,\pi_1(i),\cdots,\pi_m(i)} = \min_{\substack{A_1,\cdots,A_{m+1} \subset \{1,\cdots,n\} \\ \Sigma|A_i| = mn+1}} \max_{(i_1,\cdots,i_{m+1}) \in \prod_{i=1}^{m+1} A_i} a_{i_1,\cdots,i_{m+1}}. \tag{68}$$ ∎

Remark 8. Gross [10] has also given an algorithm to solve the bottleneck problem for $m = 1$. This algorithm can be generalized in an obvious way to the general case $m \geq 1$. ∎

REFERENCES

[1] Brualdi, R. A. and Csima, J. (1975). Stochastic patterns. *J. Combin. Theory* **19**, 1–12.

[2] Brualdi, R. A. and Csima, J. (1975). Extremal plane stochastic matrices of dimension three. *Linear Algebra Appl.* **11**, 105–133.

[3] Cambinis, S., Simons, G. and Stout, W. (1976). Inequalities for $Ek(X,Y)$ when the marginals are fixed. *Z. Wahrsch. Verw. Gebiete* **36**, 285–294.

[4] Csima, J. (1970). Multidimensional stochastic matrices and patterns. *J. Algebra* **14**, 194–202.

[5] Dall'Aglio, G. (1959). Sulla compatibilità delle funzioni di ripartizione doppia. *Rend. Mat.* **18**, 385–413.

[6] Dall'Aglio, G. (1972). Fréchet classes and compatibility of distribution functions. *Sympos. Math.* **9**, 131–150.

[7] Edmonds, J. and Fulkerson, D. R. (1970). Bottleneck extrema. *J. Combin. Theory* **8**, 299–306.

[8] GAFFKE, N. and RÜSCHENDORF, L. (1981). On a class of extremal problems in statistics. *Math. Operationsforsch. Statist., Ser. Optim.* **12**, 123–135.

[9] GALAMBOS, J. (1978). *The Asymptotic Theory of Extreme Order Statistics.* Wiley, New York.

[10] GROSS, O. (1959). The Bottleneck Assignment Problem. Rand Corporation P-1630.

[11] HAILPERIN, T. (1965). Best possible inequalities for the probability of a logical function of events. *Amer. Math. Monthly* **72**, 343–359.

[12] HUNTER, D. (1976). An upper bound for the probabilty of a union. *J. Appl. Probab.* **13**, 597–603.

[13] JURKAT, W. B. and RYSER, H. J. (1968). Extremal configurations and decompositon theorems. I. *J. Algebra* **8**, 194–222.

[14] KELLERER, H. G. (1964). Verteilungsfunktionen mit gegebenen Marginalerteilungen. *Z. Wahrsch. Verw. Gebiete* **3**, 247–270.

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

[16] KELLERER, H. G. (1987). Measure theoretic versions of linear programming. Preprint.

[17] KNOTT, M. and SMITH, C. S. (1984). On the optimal mapping of distributions. *J. Optim. Theory Appl.* **43**, 39–49.

[18] MAURER, W. (1980). Bivalent trees and forests. Preprint.

[19] RACHEV, S. T. (1984). The Monge Kantorovich mass transference problem and its stochastic applications. *Theory Probab. Appl.* **24**, 647–671.

[20] RACHEV, S. T. and RÜSCHENDORF, L. (1990). A transformation property of minimal metrics. *Theory Probab. Appl.* **35**, 131–137.

[21] RÜGER, B. (1979). Scharfe untere und obere Schranken für die Wahrscheinlichkeit der Realisation von k unter n Ereignissen. *Metrika* **26**, 71–77.

[22] RÜSCHENDORF, L. (1980). Inequalities for the expectation of Δ-monotone functions. *Z. Wahrsch. Verw. Gebiete* **54**, 341–34.

[23] RÜSCHENDORF, L. (1981). Sharpness of Fréchet bounds. *Z. Wahrsch. Verw. Gebiete* **57**, 293–302.

[24] RÜSCHENDORF, L. (1982). Random variables with maximum sums. *Adv. in Appl. Probab.* **14**, 623–632.

[25] Rüschendorf, L. (1983). Solution of a statistical optimization problem by rearrangement methods. *Metrika* **30**, 55–62.

[26] Rüschendorf, L. (1983). On the multidimensional assignment problem. *Methods Oper. Res.* **47**, 107–113.

[27] Rüschendorf, L. (1984). On the minimum discrimination information theorem. *Statist. Decisions*, Suppl. Issue No. 1, 263–283.

[28] Rüschendorf, L. (1985). The Wasserstein distance and approximation theorems. *Z. Wahrsch. Verw. Gebiete* **70**, 117–129.

[29] Rüschendorf, L. (1985). Construction of multivariate distributions with given marginals. *Ann. Inst. Statist. Math.* **37**, 225–233.

[30] Rüschendorf, L. and Rachev, S. T. (1990). A characterization of random variables with minimum L^2-distance. *J. Multivariate Anal.* **32**, 48–54.

[31] Rüschendorf, L. (1990). Fréchet bounds and their applications. Preprint.

[32] Shortt, R. M. (1983). Combinatorial methods in the study of marginal problems over separable spaces. *J. Math. Anal. Appl.* **97**, 462–479.

[33] Strassen, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Statist.* **36**, 423–439.

[34] Tong, Y. L. (1980). *Probability Inequalities in Multivariate Distributions.* Academic Press.

[35] Tchen, A. (1980). Inequalities for distributions with given marginals. *Ann. Probab.* **8**, 814–827.

[36] Vorobev, N. N. (1962). Consistent families of measures and their extensions. *Theory Probab. Appl.* **7** 147–163.

[37] Warmuth, W. (1988). Fréchet-bounds for multidimensional distributions functions. *Statistics* **19**, 283–294.

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

Institut für Mathematische Statistik
Einsteinstrasse 62
D-4400 Münster, Germany