J. Math. Anal. Appl. 506 (2022) 125666

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

The multistochastic Monge–Kantorovich problem

Nikita A. Gladkov [a], Alexander V. Kolesnikov [b], Alexander P. Zimin [b,c,*]

[a] *UCLA Department of Mathematics, Los Angeles, United States of America*
[b] *National Research University Higher School of Economics, Russian Federation*
[c] *Department of Mathematics, Massachusetts Institute of Technology, Cambridge, United States of America*

A R T I C L E I N F O

Article history:
Received 26 December 2020
Available online 17 September 2021
Submitted by G. Carlier

Keywords:
Optimal transportation
Monge-Kantorovich problem
Multistochastic Monge-Kantorovich problem
Multimarginal transportation problem
Dual linear program
Fractals

A B S T R A C T

The multistochastic Monge–Kantorovich problem on the product $X = \prod_{i=1}^n X_i$ of n spaces is a generalization of the multimarginal Monge–Kantorovich problem. For a given integer number $1 \le k < n$ we consider the minimization problem $\int c d\pi \to \inf$ on the space of measures with fixed projections onto every $X_{i_1} \times \cdots \times X_{i_k}$ for arbitrary set of k indices $\{i_1, \dots, i_k\} \subset \{1, \dots, n\}$. In this paper we study basic properties of the multistochastic problem, including well-posedness, existence of a dual solution, boundedness and continuity of a dual solution.

Contents

* Corresponding author at: Usacheva str., 6, 119048, Moscow, Russian Federation.
E-mail addresses: gladkovna@gmail.com (N.A. Gladkov), sascha77@mail.ru (A.V. Kolesnikov), azimin@mit.edu (A.P. Zimin).

https://doi.org/10.1016/j.jmaa.2021.125666

1. Introduction

This paper is a continuation of our previous work [19], where we studied a natural generalization of the transportation or Monge–Kantorovich problem.

Let μ and ν be probability measures on measurable spaces X and Y, and let $c : X \times Y \to \mathbb{R}$ be a measurable function. The classical Kantorovich problem is the minimization problem

$$\int\limits_{X\times Y} c(x, y)\; d\pi \to \inf$$

on the space $\Pi(\mu, \nu)$ of probability measures on $X \times Y$ with fixed marginals μ and ν.

It is well-known that this problem is closely related to another linear programming problem, which is called “dual transportation problem”

$$\int f\; d\mu + \int g\; d\nu \to \sup .$$

The dual transportation problem is considered on the couples of integrable functions (f, g), satisfying $f(x) + g(y) \le c(x, y)$ for all $x \in X$, $y \in Y$.

Nowadays, the Monge–Kantorovich theory attracts growing attention. The reader can find huge amount of information in the following books and surveys papers: [1,6,13,16,22,23,32,33,36,37].

A particular case of the multistochastic problem is the multimarginal transportation problem. In the multimarginal problem one considers the product of $n > 2$ spaces and n independent marginals $\mu_1, \dots, \mu_n$. Some classical results on the multimarginal problem is contained in book [32], in particular, functional-analytical duality theorems, applications to probability etc. Nevertheless, till recent, only the case of two marginals was in focus of research. A revival of interest in the case of many marginals is partially motivated by applications in economics and quantum physics [10,11,14,31]. Our motivation to study the cost function xyz in $\mathbb{R}^3$ is partially related to the multimarginal problem considered in [18].

In [19] we introduce a more general problem, which we call “multistochastic problem”. Compare to the classical (multimarginal) case this new problem is genuinely more difficult. Even its well-posedness depends on the structure of the marginals in a complicated way. The aim of this work is to fill many gaps related to basic properties of the problem.

The paper is organized as follows: the reader can consider Section 2 as an extended introduction, where we present the results of the paper, our previous results, open questions, examples, and discuss relations to other problems. In Section 3 we study sufficient conditions for existence of a feasible measure for the multistochastic problem. In Section 4 we give a proof of a duality theorem which is based on the duality theory for linearly constrained transportation problem. In Section 5 we study sufficient conditions for the existence of a dual solution and construct an example of non-existence. In Section 6 we give explicit uniform bounds for the dual solution under assumption that the cost function is bounded. Then we prove uniqueness of the primal and dual solutions in our main example studied in [19]. Finally, we give an example showing that a dual solution can be discontinuous even for a nice cost function c.

2. The multistochastic Monge–Kantorovich problem. Preliminaries, examples, and open questions

We start with the formulation of the multistochastic problem in the most general setting. Let $X_1, X_2, \dots, X_n$ be measurable spaces equipped with σ-algebras $\mathcal{B}_1, \dots, \mathcal{B}_n$. It will be assumed throughout that X_i are Polish spaces and $\mathcal{B}_i$ are Borel σ-algebras. For arbitrary space X let us denote by $\mathcal{P}(X)$ the space of all probability measures on X.

Definition 2.1. Let p, q be nonnegative integers, $q \le p$. Let us denote by $\mathcal{I}_{p,q}$ or $\mathcal{I}_{pq}$ the family of subsets $\{1, 2, \dots, p\}$ of cardinality q. In addition, the family of all subsets of $\{1, 2, \dots, p\}$ will be denoted by $\mathcal{I}_p = \cup_{q=0}^{p} \mathcal{I}_{pq}$.

Definition 2.2. For all $\alpha \in \mathcal{I}_n$ let us set $X_\alpha = \prod_{i\in\alpha} X_i$. The product of all spaces $X = \prod_{i=1}^{n} X_i$ will be denoted by X. For a fixed $\alpha \in \mathcal{I}_n$ the projection of X onto X_α will be denoted by Pr_α. In addition, for arbitrary $x \in X$ the image of x under projection Pr_α will be denoted by x_α: $x_\alpha = \mathrm{Pr}_\alpha(x)$. We also denote by $\mathrm{Pr}_\alpha(\mu) \in \mathcal{P}(X_\alpha)$ the pushforward of measure $\mu \in \mathcal{P}(X)$ by Pr_α.

Definition 2.3. Assume that for every $\alpha \in \mathcal{I}_{nk}$ we are given a probability measure μ_α on X_α. We say that a measure $\mu \in \mathcal{P}(X)$ is **uniting** if $\mathrm{Pr}_\alpha(\mu) = \mu_\alpha$ for all $\alpha \in \mathcal{I}_{nk}$. The set of all uniting measures will be denoted by $\Pi(X, \{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$. We will omit the explicit mention of the space X in the notation if this space is uniquely determined from the context.

Given the family of probability measures $\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$, we will consider the cost function c defined on the space X with the following property: there exists a collection of integrable functions $\{c_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$, $c_\alpha \in L_1(X_\alpha, \mu_\alpha)$, such that $|c(x)| \le \sum_{\alpha\in\mathcal{I}_{nk}} c_\alpha(x_\alpha)$. Every such a function c is integrable with respect to all uniting measures $\pi \in \Pi(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$. Indeed, one has $\int_X |c|\, d\pi \le \sum_{\alpha\in\mathcal{I}_{nk}} \int_X c_\alpha(x_\alpha)\ d\mu_\alpha$, if $|c(x)| \le \sum_{\alpha\in\mathcal{I}_{nk}} c_\alpha(x_\alpha)$. So, for these cost functions c the following problem is correctly defined:

Problem 2.4 *(Primal (n, k)-Monge–Kantorovich problem). Given Polish spaces $X_1, \dots, X_n$, fix family of measures $\mu_\alpha \in \mathcal{P}(X_\alpha)$, $\alpha \in \mathcal{I}_{nk}$ and a measurable cost function c. Assume that there exist integrable functions $c_\alpha \in L_1(X_\alpha, \mu_\alpha)$, $\alpha \in \mathcal{I}_{nk}$, such that $|c(x)| \le \sum_{\alpha\in\mathcal{I}_{nk}} c_\alpha(x_\alpha)$. Then we are looking for*

$$\inf_{\pi\in\Pi(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})} \int_X c\ d\pi,$$

where infimum is taken among the all uniting measures π.

In what follows, we will additionally require c and $\{c_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$ to be continuous, and this motivates us to introduce the following functional spaces

$$C_L(X_\alpha, \mu_\alpha) = C(X_\alpha) \cap L^1(\mu_\alpha),$$

$$C_L(X, \{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}}) = \left\{ c \in C(X) : |c(x)| \le \sum_{\alpha\in\mathcal{I}_{nk}} c_\alpha(x_\alpha) \text{ for some } \{c_\alpha\}_{\alpha\in\mathcal{I}_{nk}}, c_\alpha \in C_L(X_\alpha, \mu_\alpha) \right\}.$$

The notation C_L means "the function c is continuous and belongs to the space $L^1(\mu)$". In addition, we denote by $C_b(X)$ the space of all bounded continuous functions on X, $C_b(X) \subset C_L(X, \{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ for every family $\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$.

Example 2.5. **($(3,2)$-problem)** Consider a product of three spaces $X = X_1 \times X_2 \times X_3$, probability measures μ_{12}, μ_{23}, μ_{13} on $X_1 \times X_2$, $X_2 \times X_3$, $X_1 \times X_3$ respectively. Then $\mu \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$ if and only if μ is a measure on X such that

$$\mathrm{Pr}_{12}(\mu) = \mu_{12}, \mathrm{Pr}_{13}(\mu) = \mu_{13}, \mathrm{Pr}_{23}(\mu) = \mu_{23}.$$

Throughout this work we use notation like μ_{ij} (meaning the measure on $X_i \times X_j$) or $\mathcal{I}_{nk}$ with two indices without commas in most places. Also note that μ_{ij} is the same as μ_{ji} since indices α are unordered sets.

In this introductory section we briefly describe several aspects of this problem. In particular, we discuss previously known results, examples, open problems, and relation to other research.

2.1. Feasibility of the problem, Latin squares and descriptive geometry

The multistochastic problem is overdetermined and a uniting measure does not always exist. It is clear that a necessary condition for existence of a uniting measure is the following consistency condition:

$$\mathrm{Pr}_{\alpha\cap\beta}(\mu_\alpha) = \mathrm{Pr}_{\alpha\cap\beta}(\mu_\beta) = \mathrm{Pr}_{\alpha\cap\beta}(\mu).$$

This condition is not sufficient (see [19] and other examples below), but we show that this condition is sufficient for existence of a **signed** uniting measure (see Theorem 3.6).

Nevertheless, in certain situations the set of feasible measures is very rich. This happens, for instance, if X_i are finite sets of the same cardinality and all the measures μ_α are uniform. The natural continuous generalization is: $X_i = [0,1]$ and μ_α are the Lebesgue measures on $[0,1]^k$ of the corresponding dimension k. A natural related discrete combinatorial object is a **Latin square**. We recall that a Latin square is an $n \times n$ array filled with symbols from $\{1, \dots, n\}$, each occurring exactly once in each row and exactly once in each column. To see the relation let us consider an $n \times n$ Latin square S containing first n integers. Then the discrete measure

$$\frac{1}{n^2} \sum_{i,j} \delta_{i,j,S(i,j)}$$

on $\{1, \dots, n\}^3$ has uniform projections to discrete xy, xz, yz planes.

More generally, the (n,k)-multistochastic problem is always feasible for the system of measures

$$\mu_\alpha = \prod_{i\in\alpha} \mu_i, \ \alpha \in \mathcal{I}_{nk},$$

where $\mu_1, \dots, \mu_n$ are fixed measures on $X_1, \dots, X_n$.

We believe that this example provides a natural source of applications, this is why a big part of our results is related to this particular case.

An interesting example of the optimal transportation problem was studied in connection with applications to the density functional theory, namely, the Hohenberg–Kohn theory. The Hohenberg–Kohn theory considers a model of N electrons whose arrangement in the space $\mathbb{R}^{3N}$ is determined by the density $\rho_N(x_1, \dots, x_N)$. The energy of pairwise interaction of electrons is specified as the density integral over the Coulomb potential:

$$\mathcal{V}_{ee} = \int \sum_{1 \le i < j \le N} \frac{\rho_N(x_1, \dots, x_N)}{|x_i - x_j|}\, dx_1 \dots dx_N.$$

Due to the symmetry,

$$\mathcal{V}_{ee} = \int \sum_{1 \le i < j \le N} \frac{\rho_2(x, y)}{|x - y|}\, dx dy,$$

where $\rho_2(x, y) = \int \rho_N(x, y, z_3, \cdots, z_N)\, dz_3 \dots dz_N$.

In the Hohenberg-Kohn theory, the ground state is described by a functional that depends only on the density of one electron

$$\rho(x) = \int \rho_N(x, z_2, \cdots, z_N)\, dz_2 \cdots dz_N.$$

For this purpose $\mathcal{V}_{ee}(\rho_2)$ is approximated by the functional $\mathcal{V}_{ee}(\rho)$, depending only on ρ. The correct approximation is the key problem in this theory.

It turns out that the natural approximation is the approximation by the functional

$$F(\rho) = \inf_{\pi \in \Pi(\rho, \rho)} \int \frac{1}{|x - y|}\, \pi(dx, dy).$$

For example, this functional occurs when the so-called "semi-classical limit" is taken. Trivially, the functional F is the Kantorovich functional (for the pair of equal marginals) with the cost function $\frac{1}{|x-y|}$. This cost function is called Coulomb cost function.

In [12] the passage to the limit is made rigorously, and some sufficient conditions for the existence and uniqueness of a solution for the Kantorovich functional are found. For a generalization to a wider class of "repulsive cost functions" see [10]. For further progress in physical applications, see [4]. In [9] transport inequalities and concentration inequalities for the Coulomb cost function are obtained.

In this setting, the multistochastic Monge-Kantorovich problem can be applied if we have additional restrictions on joint distributions of electrons. For example, if we know that every set of k electrons is mutually independent, we can consider the (n, k)-problem with the fixed family of projections $\mu_\alpha = \prod_{i \in \alpha} \mu_i$.

Other source of inspiration might arise from the engineering, in particular, the descriptive geometry. In engineering it is common to depict a three-dimensional body using its two (as originally suggested by father of descriptive geometry, Gaspard Monge, who also gave his family name to Monge–Kantorovich problem) or three orthogonal projections onto orthogonal two-dimensional planes. So an engineer might find themselves reconstructing a three-dimensional body by its top view, front view and side view. When instead of a body one has a measure, that turns into finding a set of uniting measures in $(3, 2)$-problem.

A necessary and sufficient condition for existence of a measure with a given system of marginal distributions in the spirit of linear programming duality was established by H. Kellerer [24]. Assume we are given a system of marginal distributions μ_α, where α belongs to some system A of subsets of $\{1, \dots, n\}$. This system admits a uniting measure if and only if

$$\sum_{\alpha \in A} \int f_\alpha(x_\alpha)\, d\mu_\alpha \ge 0$$

for all bounded continuous family of functions $\{f_\alpha(x_\alpha)\}_{\alpha \in \mathcal{I}_{nk}}$ satisfying $\sum_{\alpha \in A} f_\alpha(x_\alpha) \ge 0$. We give an independent proof of this fact for $A = \mathcal{I}_{nk}$ in Section 3. Note, however, that this criterion does not seem to be very practical. We establish some easy-to-check sufficient conditions for existence of uniting measure in terms of uniform bounds for densities. In particular, we prove the following (see Theorem 3.11):

Theorem 2.6. *For given natural numbers $1 \le k < n$ there exists a constant $\lambda_{nk} > 1$ which admits the following property.*

Assume we are given a consistent (see Definition 3.2) family of probability measures $\mu_\alpha \in \mathcal{P}(X_\alpha)$, $\alpha \in \mathcal{I}_{nk}$, and another family of probability measures $\nu_i \in \mathcal{P}(X_i)$, $1 \le i \le n$. Assume that every measure μ_α, $\alpha \in \mathcal{I}_{nk}$, is absolutely continuous with respect to $\nu_\alpha = \prod_{i \in \alpha} \nu_i$:

$$\mu_\alpha = \rho_\alpha \cdot \nu_\alpha.$$

Finally, assume that there exist constants $0 < m \le M$ such that every density ρ_α satisfies $m \le \rho_\alpha \le M$ ν_α-almost everywhere for all $\alpha \in \mathcal{I}_{nk}$.

Then $\Pi(\{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}})$ is not empty provided $\frac{M}{m} \le \lambda_{nk}$.

We will give precise bounds for the constant λ_{32}.

Remark 2.7. Solvability of the primal problem. As soon as the set of uniting measures is not empty, the proof of existence of a solution to the primal problem for a lower semicontinuous cost is a standard exercise.

Theorem 2.8 *([19]). Assume that the cost function $c \ge 0$ is lower semicontinuous. If $\Pi(\{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}})$ is not empty, then there exists a solution to the multistochastic problem.*

2.2. Examples. Fractal structure versus smooth structure

The main example of an explicit solution to a multistochastic problem was found in [19]. The unexpected beauty of this example was the main motivation for us for subsequent study of the multistochastic problem.

In the following example we consider a $(3, 2)$-problem. Denote by $\Pi(\mu_{xy}, \mu_{yz}, \mu_{xz})$ the set of measures with projections $\mathrm{Pr}_{xy}\pi = \mu_{xy}, \mathrm{Pr}_{xz}\pi = \mu_{xz}, \mathrm{Pr}_{yz}\pi = \mu_{yz}$.

Theorem 2.9 *([19]). Let $\mu_{xy} = \lambda_{xy}, \mu_{xz} = \lambda_{xz}, \mu_{yz} = \lambda_{xz}$ be the two–dimensional Lebesgue measures $[0, 1]^2$ and let $c = xyz$. Then there exists a unique solution to the corresponding $(3, 2)$-problem*

$$\int xyz \, d\pi \to \min, \ \ \pi \in \Pi(\mu_{xy}, \mu_{yz}, \mu_{xz}).$$

It is concentrated on the set

$$\mathfrak{S} = \{(x, y, z) \colon x \oplus y \oplus z = 0\},$$

where $\oplus$ is the bitwise addition (see Definition 6.13). See Fig. 1.

The set $\mathfrak{S}$ is called *Sierpińsky tetrahedron.*

We stress that some fractal solutions to a multimarginal transportation problem were known before our work. See, for instance, [14], where multimarginal problem with the cost function of the type $h(\sum_{i=1}^n x_i)$ and the Lebesgue measure projections was considered. Though we don't see any direct relation between these examples, they have something in common: in both cases the entire construction relies on the dyadic decomposition.

Remark 2.10. The $(3, 2)$-problem can admit not only fractal but also smooth solutions. For instance, consider measurable functions $f(x)$, $g(y)$ and $h(z)$ on $[0, 1]$. Assume that h is injective, the set $\Gamma = \{f(x) + g(y) + h(z) = 0\}$ is not empty, and μ is a probability measure concentrated on Γ: $\mu(\Gamma) = 1$. Set $\mu_{xy} = \mathrm{Pr}_{xy}\mu$, $\mu_{xz} = \mathrm{Pr}_{xz}\mu$, $\mu_{yz} = \mathrm{Pr}_{yz}\mu$. Then μ is the unique element of $\Pi(\mu_{xy}, \mu_{yz}, \mu_{xz})$. Indeed, let $\nu \in \Pi(\mu_{xy}, \mu_{yz}, \mu_{xz})$.

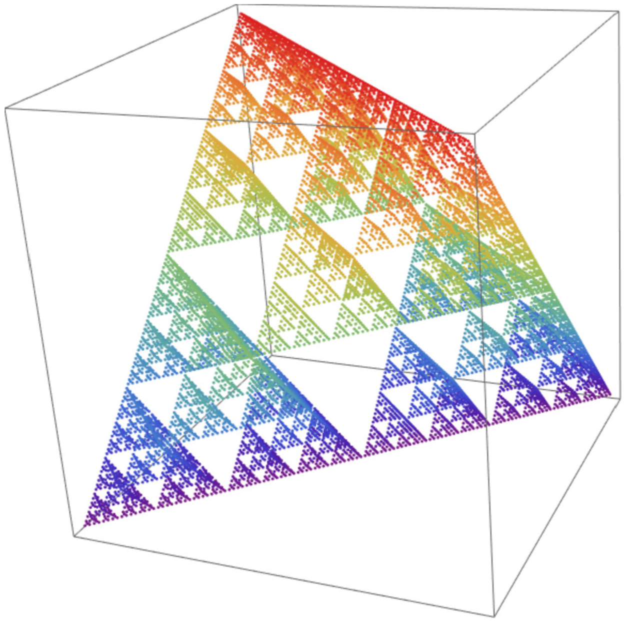

Fig. 1. The solution is supported on Sierpińsky tetrahedron.

Clearly, $\int (f(x) + g(y) + h(z))^2 \, d\nu$ depends solely on the integrals of pairwise products of functions f, g, h with respect to measures $\mu_{xy}, \mu_{yz}, \mu_{xz}$. Hence

$$\int (f(x) + g(y) + h(z))^2 \, d\nu = \int (f(x) + g(y) + h(z))^2 \, d\mu = 0,$$

this implies that ν is concentrated on Γ. Since h is injective, Γ is the graph of the mapping $(x, y) \to h^{-1}(-f(x) - g(y))$, hence ν is uniquely determined by its projection μ_{xy}, thus coincides with μ.

In particular, this observation can be applied to construct an example of a solution concentrated on a smooth set.

Example 2.11. The Lebesgue measure on $[0, 1]^3 \cap \{x_1 + x_2 + x_3 = 1\}$ is a solution to the $(3, 2)$-problem, where marginals are the two-dimensional Lebesgue measures concentrated on the set $\{x_i + x_j \leq 1\} \subset [0, 1]^2$ and arbitrary cost function.

It is clear, that the smoothness of the solution in this example is just a matter of fact that $\Pi(\mu_{xy}, \mu_{yz}, \mu_{xz})$ contains a unique (smooth) element. However, it is natural to expect that the solution may have a fractal/non-regular structure provided uniting measures constitute a sufficiently large set.

The following problem, yet vaguely formulated, seems to be crucial for understanding of the structure of solutions to (n, k)-problem.

Open problem 1. Is it true that solutions to (n, k)-problem have "fractal structure" provided $\Pi(\{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}})$ contains sufficiently "rich" set of measures?

2.3. Duality and the Kantorovich problem with linear constraints

As in the classical case the multistochastic problem admits the corresponding dual problem:

Problem 2.12 *(Dual (n, k)-Monge–Kantorovich problem). Assume we are given Polish spaces $X_1, \ldots, X_n$, a fixed family of measures $\{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}}$, and a cost function $c \in C_L(X, \{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}})$. Find*

$$\sup_{f \le c} \sum_{\alpha \in \mathcal{I}_{nk}} \int\limits_{X_\alpha} f_\alpha \, d\mu_\alpha,$$

where the supremum is taken among the functions f having the form $f(x) = \sum_{\alpha \in \mathcal{I}_{nk}} f_\alpha(x_\alpha)$, where $f_\alpha \in L^1(X_\alpha, \mu_\alpha)$.

Definition 2.13. We say that there is no duality gap for the (n, k)-problem if

$$\min_{\pi \in \Pi(\{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}})} \int c \, d\pi = \sup_{f \le c} \sum_{\alpha \in \mathcal{I}_{nk}} \int\limits_{X_\alpha} f_\alpha \, d\mu_\alpha,$$

where $f_\alpha \in L^1(X_\alpha, \mu_\alpha)$, $f(x) = \sum_{\alpha \in \mathcal{I}_{nk}} f_\alpha(x_\alpha)$.

The absence of duality gap was shown in [19] under assumption of compactness of the spaces X_i. In this work we prove the following result:

Theorem 2.14. *There is no duality gap for (n, k)-problem provided X_i are Polish spaces and $c \in C_L(X, \{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}})$.*

Our approach is based on the result of D. Zaev [38] on duality for the classical Kantorovich problem with **linear constraints**. The transportation problem with linear constraints is the standard Kantorovich problem with additional constraints of the type $l(P) = 0$, where l is a linear functional on the space of measures. The proof of Zaev is based on the general minimax principle.

2.4. Structure of dual solutions. Monge problem

Our main example of a dual solution is given in the following theorem.

Theorem 2.15 *([19]). Let $\mu_{xy} = \lambda_{xy}, \mu_{xz} = \lambda_{xz}, \mu_{yz} = \lambda_{xz}$ be the two dimensional Lebesgue measures on $[0, 1]^2$ and $c = xyz$. Then the triple of functions $(f(x, y), f(x, z), f(y, z))$, where*

$$f(x, y) = \int\limits_0^x \int\limits_0^y t \oplus s \, dtds - \frac{1}{4} \int\limits_0^x \int\limits_0^x t \oplus s \, dtds - \frac{1}{4} \int\limits_0^y \int\limits_0^y t \oplus s \, dtds$$

solves the corresponding dual multistochastic problem.

Remark 2.16. The uniqueness result for this problem under assumption of continuity of the dual solution is proved in the present paper in Theorem 2.30.

The solution to the dual problem given in Theorem 2.15, has the following relation to the solution π to the primal problem (see Theorem 2.9): π is concentrated on the graph of the mapping $(x, y) \mapsto f_{xy}(x, y)$, i.e.

$$z = f_{xy}(x, y) \tag{1}$$

π-almost everywhere.

Let us note that f admits a non-negative mixed derivative f_{xy}, but derivatives f_{xx}, f_{yy} do not exist (at least in the classical sense).

The relation (1) can be derived from the fact that the support $\mathfrak{S}$ of the solution π is a fractal set. Indeed, function $f(x,y)+f(x,z)+f(y,z)-xyz$ is non positive and equals zero π-a.e. Thus for π-almost all points the first order condition

$$f_x(x,y)+f_x(x,z)=yz \tag{2}$$

is satisfied.

Next, it is easy to show that for π-almost every point $M=(x_0,y_0,z_0)\in\mathfrak{S}$ the set $\mathfrak{S}$ contains points of the type $M+t_n v$, where t_n is a sequence tending to zero and vector v belongs to a set V containing three independent vectors. One can prove this using the fractal structure of $\mathfrak{S}$. Consequently, one can differentiate (2) along V and deduce (1) from these relations.

Thus in this particular case the solution admits the following properties.

(a) The solution is concentrated on the graph of a mapping $z=T(x,y)$.
(b) This mapping T has the form $T(x,y)=f_{xy}(x,y)$, where (f,g,h) is a solution to the dual problem. The same holds for g,h.
(c) Function $f(x,y)$ is a cumulative distribution function (up to a term depending on x and a term depending on y) of a positive measure on a plane. Equivalently, $f_{xy}(x,y)\geq 0$ almost everywhere.

These properties together resemble the result by Brenier [7] that in usual Monge–Kantorovich problem for cost function $-x\cdot y$ (or equivalently $\frac{1}{2}||x-y||^2$) one has the optimal transport plan concentrated on the graph of the gradient of some convex function φ such that (φ,φ^*) is a solution to the dual problem. Here we have function $c=xyz$ and optimal transport plan concentrated on the graph of the mixed derivative of some function f with positive mixed derivative which is also a solution to the dual problem.

Definition 2.17. (Optimal mapping.) Let T satisfy (a). Then we say that T is an optimal mapping.

One can ask whether any solution to $(3,2)$-problem (under natural assumptions on the marginals) with the cost function xyz does satisfy properties (a), (b), (c). We show that in fact no one of these properties are satisfies in general.

Example 2.18. The solutions to $(3,2)$-problems are not always concentrated on graphs; (a) fails. Consider the sphere $S=\{x^2+y^2+z^2=1\}$, and consider the quarter sphere $S_1=S\cap\{x\geq 0,y\geq 0\}$, $S_2=S\cap\{x<0,y\geq 0\}$, $S_3=S\cap\{x<0,y<0\}$ and $S_4=S\cap\{x\geq 0,y<0\}$. Let π be the surface measure on the 3/4-part of the sphere $S_1\sqcup S_2\sqcup S_4$, and let $\mu_{xy},\mu_{xz},\mu_{yz}$ be the corresponding two-dimensional projections.

Slightly modifying the arguments of Remark 2.10 we prove that if $\widehat{\pi}$ is a measure with projections μ_{xy}, μ_{xz} and μ_{yz}, then $\widehat{\pi}$ is concentrated on the set $S_1\sqcup S_2\sqcup S_4$. For each point of S_2 there is no other point of $S_1\sqcup S_2\sqcup S_4$ with the same projection onto the coordinate plane Oxz, and therefore the restriction of the measure $\widehat{\pi}$ to S_2 is fully determined by its projection μ_{xz} and coincides with $\pi|_{S_2}$.

Similarly, the restriction of $\widehat{\pi}$ to S_4 is fully determined by its projection μ_{yz} and coincides with $\pi|_{S_4}$. Hence, $\widehat{\pi}|_{S_1}=\widehat{\pi}-\pi|_{S_2}-\pi|_{S_4}$. Thus, the projections of $\widehat{\pi}|_{S_1}$ and $\pi|_{S_1}$ to the coordinate planes are the same, and then $\widehat{\pi}|_{S_1}=\pi|_{S_1}$. So we conclude that π is the only measure with projections $\mu_{xy},\mu_{xz},\mu_{yz}$, and there is no optimal mappings T_{xy}, T_{xz} and T_{yz}.

See also Example 5.10 for a discrete counterexample.

Example 2.19. Example without dual solutions satisfying (1); (b) fails. This example is considered in Theorem 6.33. In this example f_{xy} is either zero or not defined.

Example 2.20. Non-uniqueness for the dual problem; (c) fails. In the problem considered in Example 2.11 there exist many dual solutions. To see this let us note that the following inequality holds for all $(x, y, z) \in [0, 1]^3$ and a fixed constant $A > 0$, equality holds if and only if $x + y + z = 1$:

$$(x + y + z - 1)^2(x + y + z + A) \geq 0.$$

Developing the left-hand side we see that this inequality is equivalent to

$$xyz \geq f_A(x, y) + f_A(x, z) + f_A(y, z),$$

where

$$f_A(x, y) = -\frac{1}{12}(x^3 + y^3) - \frac{1}{2}xy(x + y) - (A - 2)\left(\frac{x^2}{12} + \frac{xy}{3} + \frac{y^2}{12}\right) - \frac{1 - 2A}{12}(x + y) - \frac{A}{18}.$$

Clearly, the triple $(f_A(x, y), f_A(x, z), f_A(y, z))$ solves the dual problem for every $A > 0$. Note that (1) and (c) fails for all $A > 0$.

Thus we see that the particular form (1) of the optimal mapping related to $(3, 2)$-problem with cost function xyz is related to the fractal structure of the solution. Motivated by these observations we state the following problem.

Open problem 2. Assume that π is a solution to a $(3, 2)$-problem with the cost function xyz. Find general sufficient conditions for presentation of π in the form

$$z = f_{xy}(x, y),$$

where $f(x, y)$, $g(x, z)$, $h(y, z)$ solve the corresponding dual multistochastic problem.

It seems quite difficult to describe the general structure of solutions to $(3, 2)$-problem with $c = xyz$, since it is very sensitive to non-local properties of the marginals. Something can be established under very strong "smoothness" assumptions, as presented in the proposition below. But we stress that this situation can not pretend to describe a reasonable model case.

Proposition 2.21. *Consider a triple of twice continuously differentiable functions $f(x, y)$, $g(x, z)$, $h(y, z)$ satisfying $f(x, y) + g(x, z) + h(y, z) \geq xyz$. Assume, in addition, that*

$$\Gamma = \{f(x, y) + g(x, z) + h(y, z) = xyz\}$$

is a two-dimensional smooth surface.

Let $\Gamma_x, \Gamma_y, \Gamma_z$ be sets defined by equations:

$$\Gamma_x = \{x = h_{yz}\},\ \Gamma_y = \{y = g_{xz}\},\ \Gamma_z = \{z = f_{xy}\}.$$

Then for every point $(x_0, y_0, z_0) \in \Gamma$ the following alternative holds:

(A) (x_0, y_0, z_0) *belongs to at least two of sets* $\Gamma_x, \Gamma_y, \Gamma_z$*:* $(x_0, y_0, z_0) \in (\Gamma_x \cap \Gamma_y) \bigcup (\Gamma_x \cap \Gamma_z) \bigcup (\Gamma_y \cap \Gamma_z)$
(B) $(x_0, y_0, z_0) \notin \Gamma_x \cup \Gamma_y \cup \Gamma_z$ *and the vector field*

$$N = \left(\frac{1}{x - h_{yz}}, \frac{1}{y - g_{xz}}, \frac{1}{z - f_{xy}}\right)$$

is orthogonal to Γ at (x_0, y_0, z_0).

Proof. Since every $(x, y, z) \in \Gamma$ is a minimum point of $f(x, y) + g(x, z) + h(y, z) - xyz$, then the functions

$$u = yz - f_x(x, y) - g_x(x, z),\ v = xz - f_y(x, y) - h_y(y, z),\ w = xy - g_z(x, z) - h_z(y, z)$$

vanish on Γ. Hence their gradients

$$\begin{aligned}\nabla u &= (-f_{xx} - g_{xx}, z - f_{xy}, y - g_{xz})\\ \nabla v &= (z - f_{xy}, -f_{yy} - h_{yy}, x - h_{yz})\\ \nabla w &= (y - g_{xz}, x - h_{yz}, -g_{zz} - h_{zz})\end{aligned}$$

are orthogonal to Γ. Then they are collinear, because Γ is two-dimensional.

Assume that (x, y, z) belongs to at least one of the sets $\Gamma_x, \Gamma_y, \Gamma_z$, say to Γ_z. Then $z = f_{xy}$ at this point. This implies that either ∇u is zero or $-f_{yy} - h_{yy} = 0$, $x - h_{yz} = 0$ (because $\nabla v = \lambda \nabla u$, $\nabla w = \mu \nabla u$ for some λ, μ). In the first case $y = g_{xz}$ and $(x, y, z) \in \Gamma_z \cap \Gamma_y$, while in the second case $x = h_{xy}$ and $(x, y, z) \in \Gamma_z \cap \Gamma_x$.

Repeating these arguments with the other derivatives, we see that either $(x, y, z) \in (\Gamma_x \cap \Gamma_y) \bigcup (\Gamma_x \cap \Gamma_z) \bigcup (\Gamma_y \cap \Gamma_z)$ or $(x, y, z) \notin \Gamma_x \cap \Gamma_y \cap \Gamma_z$. In the second case all the 2×2 minors equal zero, hence

$$f_{yy} + h_{yy} = -\frac{(x - h_{yz})(z - f_{xy})}{y - g_{xz}}$$

(similarly for other coordinates). This gives that N is orthogonal to Γ. $\square$

Remark 2.22. Having in mind our main example from Theorem 2.9 with the solution on the tetrahedron, one can expect that (B) never holds. However, (B) can happen, an example is given in Example 2.20. Our belief is that one should expect (A) for fractal solutions and (B) for smooth solutions. In fact, we show in Section 6.2 that (under some additional assumptions) any dual solution in our main example must satisfy alternative (A) and we derive uniqueness from this. However, we are unable so far to make a precise statement saying that (A) / (B) corresponds to fractal/smooth structure.

Remark 2.23. (Vector fields orthogonal to smooth solutions). Assume that π is a solution to a $(3, 2)$-problem concentrated on the surface Γ and alternative (B) holds. Assume, in addition, that π has a density with respect to the two-dimensional Hausdorff measure

$$\pi = p(x, y, z) \cdot \mathcal{H}^2|_\Gamma.$$

Denote by $\rho_{xy}, \rho_{xz}, \rho_{yz}$ the density of the corresponding projections $\mu_{xy}, \mu_{xz}, \mu_{yz}$. Then

$$\rho_{xy}(x, y)|\cos(N, (0, 0, 1))| = p(x, y, z)$$

for every $(x, y, z) \in \Gamma$ and

$$\rho_{xy}(x, y) = p(x, y, z)|z - f_{xy}|\sqrt{\frac{1}{(x - h_{yz})^2} + \frac{1}{(y - g_{xz})^2} + \frac{1}{(z - f_{xy})^2}}.$$

Similarly for the other densities. This easily leads to the following relations: for every $(x, y, z) \in \Gamma$ the vector field

$$\left(\frac{\operatorname{sign}(x - h_{yz})}{\rho_{yz}}, \frac{\operatorname{sign}(y - g_{xz})}{\rho_{xz}}, \frac{\operatorname{sign}(z - f_{xy})}{\rho_{xy}}\right)$$

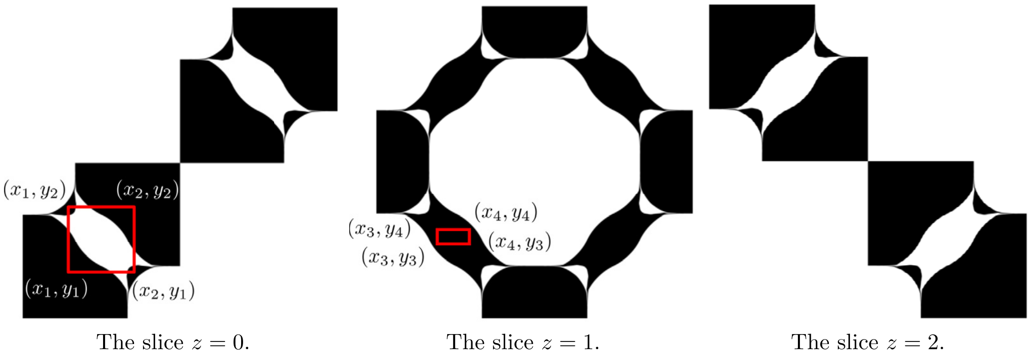

Fig. 2. The visualization of the primal solution to the problem considered in Example 2.24. Each picture shows the restriction of the primal solution to the set $z = 0, 1, 2$. In the white points the density function is equal to 0, and in the black points it is equal to 3. Almost every horizontal and vertical section of the black body has a length 1/3. Compare to Fig. 3.

is orthogonal to Γ and

$$\frac{1}{p^2(x,y,z)} = \frac{1}{\rho_{xy}^2(x,y)} + \frac{1}{\rho_{xz}^2(x,z)} + \frac{1}{\rho_{yz}^2(y,z)}.$$

In particular, we obtain that one of the vector fields

$$\left(\frac{\pm 1}{\rho_{yz}}, \frac{\pm 1}{\rho_{xz}}, \frac{\pm 1}{\rho_{xy}}\right)$$

is (locally) orthogonal to Γ.

Example 2.24. (c) fails; relation to the transportation problem with uniform bound on density. Consider the following $(3,2)$-problem with $X = Y = [0,1]$ and $Z = \{0,1,2\}$. Let $\mu_x = \mu_y$ be the Lebesgue measure on $[0,1]$, and let μ_z be the uniform discrete measure on $\{0,1,2\}$. $c = xyz$ and $\mu_{xy} = \mu_x \otimes \mu_y$, $\mu_{xz} = \mu_x \otimes \mu_z$, $\mu_{yz} = \mu_y \otimes \mu_z$. Then the solution is concentrated on the graph of a function $z = T(x,y)$, where T takes values in $\{0,1,2\}$.

We suspect that there exists a solution (f,g,h) to the dual problem such that f admits a mixed derivative f_{xy} everywhere except for the boundaries of the black regions from Fig. 2, and, wherever it exists, $f_{xy}(x,y) = z = T(x,y)$. That is close to the property (c) of our main example, but whatever happens on boundaries prevents f from being a cumulative distribution function of a positive measure.

Based on Fig. 2, we show that the inequality

$$f(x_1,y_1) + f(x_2,y_2) - f(x_1,y_2) - f(x_2,y_1) \geq 0$$

can not hold for all $x_1 < x_2$, $y_1 < y_2$. Alternatively, f can not be a cumulative distribution function of some non-negative measure ζ, so (c) fails and in particular $z = f_{xy}$ can not hold everywhere.

By the complementary slackness condition, at almost every black point (x,y,z) we have $f(x,y)+g(x,z)+h(y,z) = xyz$. In particular, if we choose the points (x_1,y_1), (x_1,y_2), (x_2,y_1), and (x_2,y_2) forming a rectangle R_1 as on Fig. 2 (slice $z = 0$), we obtain the following equation:

$$\begin{aligned}\zeta(R_1) = f(x_1,y_1) + f(x_2,y_2) - f(x_1,y_2) - f(x_2,y_1)& \\ = f(x_1,y_1) + g(x_1,0) + h(y_1,0) + f(x_2,y_2) + g(x_2,0) + h(y_2,0)& \\ - f(x_1,y_2) - g(x_1,0) - h(y_2,0) - f(x_2,y_1) - g(x_2,0) - h(y_1,0)& \\ = 0.&\end{aligned}$$

Similarly, if we choose the points (x_3, y_3), (x_3, y_4), (x_4, y_3), and (x_4, y_4) forming a rectangle R_2 as on Fig. 2 (slice $z = 1$), then ζ of R_2 is strictly positive by the complementary slackness condition:

$$\begin{aligned}\zeta(R_2) = f(x_3, y_3) + f(x_4, y_4) - f(x_3, y_4) - f(x_4, y_3)&\\ &= f(x_3, y_3) + g(x_3, 1) + h(y_3, 1) + f(x_4, y_4) + g(x_4, 1) + h(y_4, 1)\\ &\qquad - f(x_3, y_4) - g(x_3, 1) - h(y_4, 1) - f(x_4, y_3) - g(x_4, 1) - h(y_3, 1)\\ &= x_3y_3 + x_4y_4 - x_3y_4 - x_4y_3 > 0.\end{aligned}$$

This contradicts the fact that the rectangle R_2 is a subset of R_1.

Remark 2.25. It is worth noting that patterns in sets $\{z = 0\}$, $\{z = 1\}$ (see Fig. 2) appeared in literature before. One can show that these patterns are exactly solutions to an optimal transportation problem with capacity constraints considered in [29,28] which we will mention below. More explanations are given in [39].

Remark 2.26. It worth noting that the condition $f(x_1, y_1) + f(x_2, y_2) - f(x_1, y_2) - f(x_2, y_1) \geq 0$ for all $x_1 < x_2$, $y_1 < y_2$ corresponds to a bit different primal problem, where assumptions on the marginals are replaced by assumptions that the marginals are **first-ordered stochastically dominated** by given measures. That means that the distribution functions of marginals are pointwise not greater than the distribution functions of given measures. We choose this term since it is generalizes the concept of first-order stochastic dominance from decision theory from $\mathbb{R}$ to $\mathbb{R}^2$. But we don't pursue this viewpoint here.

2.5. Solvability of the dual problem

Section 5 is devoted to existence of a solution to the dual problem. We establish a sufficient existence condition for the dual problem in the spirit of a classical result of Kellerer [25] for the multistochastic problem, but with a self-contained independent proof.

The main assumption on the cost function for **solvability of the dual problem** is the following bound:

$$|c(x)| \leq \sum_{\alpha \in \mathcal{I}_{nk}} C_\alpha(x_\alpha), \tag{3}$$

for some integrable functions, $C_\alpha \colon X_\alpha \to \mathbb{R} \cup \{+\infty\}$ This is a generalization of the Kellerer's assumption.

However, as shown in Section 5.3, unlike in multimarginal case, this bound is not enough even for $(3, 2)$-problem. So another assumption, which is specific for (n, k)-problem, should be done on marginals. Namely, we have to assume that the system of measures $\{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}}$ is **reducible**. The latter means that there exists a measure $\mu \in \Pi(\{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}})$ and the system of probability measures $\{\nu_i\}_{i=1}^n$, $\nu_i \in \mathbb{P}(X_i)$ such that for some $0 < c < C$

$$c\nu \leq \mu \leq C\nu, \tag{4}$$

where $\nu = \prod_i \nu_i$. Our main existence/nonexistence result is the following Theorem (see details in Theorem 5.19 and Proposition 5.24):

Theorem 2.27. *If the system $\{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}}$ is reducible, then under assumption (3) there exists a relaxed in a sense of Definition 5.16 solution to the dual multistochastic problem.*

Without assumption of reducibility the dual solution may not exist. More precisely, there exists an example of a probability measure μ on the space $X = \mathbb{N}^3$ and the cost function $c \colon X \to \{0, 1\}$ such that there is no solution to the dual multistochastic problem for the system

$$\mu_{ij} = \mathrm{Pr}_{ij}\mu.$$

Now the question is how to understand if the collection $\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$ is reducible. Luckily, we have Theorem 3.13 extending Theorem 2.6 which gives us a sufficient condition for the collection of measures to be reducible.

2.6. Other properties of dual solutions: boundedness and (dis)continuity

In Section 6 we study basic properties of solutions to the dual $(3,2)$ problem: boundedness and continuity. It is known that for the classical (multimarginal) problem the dual solution is bounded provided $|c|$ is bounded. But this is crucial that in the classical case the dual solution is a sum of functions in non-overlapping variables. This is the reason why it is hard to extend the arguments to the general (n,k)-case. We establish the following result on the boundedness of solutions.

Theorem 2.28. *Let X_1, X_2, X_3 be Polish spaces, $\mu_i \in \mathcal{P}(X_i)$ for $1 \le i \le 3$, and let $\mu_{ij} = \mu_i \otimes \mu_j$ for all $\{i,j\} \in \mathcal{I}_{3,2}$. Let $c\colon X \to \mathbb{R}_+$ be a bounded continuous cost function. If $\{f_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$ is a solution to the related dual problem, then*

$$f_{12}(x_1,x_2) + f_{13}(x_1,x_3) + f_{23}(x_2,x_3) \ge -12\,\|c\|_\infty$$

for $\mu_1\otimes\mu_2\otimes\mu_3$-almost all points $x \in X$.

Moreover, there exists a solution $\{\widehat{f}_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$ to the dual problem such that for every couple $\{i,j\} \in \mathcal{I}_{3,2}$ and for all $x \in X$ the following inequality holds:

$$-26\frac{2}{3}\,\|c\|_\infty \le \widehat{f}_{ij}(x_i,x_j) \le 13\frac{1}{3}\,\|c\|_\infty\,.$$

Another important feature of the classical Monge–Kantorovich problem: for a cost function c with nice geometric/regularity properties the corresponding dual solutions are regular. This happens because the dual functions are related by Legendre transform, which is highly regularizing. We can not expect this for the (n,k)-problem, the following example demonstrates that a solution can be unique and discontinuous even for very simple and nice cost: maximum of two linear functions (see Theorem 6.33).

Example 2.29. Let $X = Y = Z = [0,1]$. Consider the $(3,2)$-problem with the cost function

$$c = \max(0, x + y + 3z - 3),$$

where $\mu_{zy}, \mu_{xz}, \mu_{yz}$ are the Lebesgue measures restricted to $[0,1]^2$. Then the dual problem admits a unique discontinuous solution, given by the following formulas:

$$f_{12}(x_1,x_2) = 0 \text{ for all points } (x_1,x_2) \in [0,1]^2;$$

$$f_{13}(x_1,x_3) = \begin{cases} 0, & \text{if } x_3 < \frac{2}{3}, \\ x_1 + \frac{3}{2}x_3 - \frac{3}{2}, & \text{if } x_3 \ge \frac{2}{3}; \end{cases}$$

$$f_{23}(x_2,x_3) = \begin{cases} 0, & \text{if } x_3 < \frac{2}{3}, \\ x_2 + \frac{3}{2}x_3 - \frac{3}{2}, & \text{if } x_3 \ge \frac{2}{3}. \end{cases}$$

2.7. Uniqueness result for the main example

In Section 6 we establish the following results for our main example: $(3,2)$-problem with the two-dimensional Lebesgue marginals.

Theorem 2.30. *If a triple of functions $\{f_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$ is a solution to the problem from Theorem 2.15 and every f_{ij} is continuous for all $\{i,j\}\in\mathcal{I}_{3,2}$, then there exist continuous functions $f_i\colon [0,1]\to\mathbb{R}$, $1\le i\le 3$, such that*

$$f_{12}(x_1,x_2) = f(x_1,x_2) + f_1(x_1) - f_2(x_2),$$
$$f_{23}(x_2,x_3) = f(x_2,x_3) + f_2(x_2) - f_3(x_3),$$

and

$$f_{13}(x_1,x_3) = f(x_1,x_3) + f_3(x_3) - f_1(x_1),$$

where

$$f(x,y) = \int_0^x\int_0^y s\oplus t\,dsdt - \frac{1}{4}\int_0^x\int_0^x s\oplus t\,dsdt - \frac{1}{4}\int_0^y\int_0^y s\oplus t\,dsdt.$$

Remark 2.31. We believe that this problem admits no other (discontinuous) solutions, but have no proof of this.

2.8. Relation to other problems

We mentioned already that the multistochastic problem is closely related to the Kantorovich problem with linear constraints studied by Zaev in [38]. More precisely, our problem can be reduced to the Kantorovich problem with linear constraints, see explanations in Section 4.

Another related problem is, of course, problem with uniform constraint on the density, sometimes called "the capacity constrained problem" (see [29,28,15]). The solution to the problem from Example 2.24 admits the following structure: there is a partition of the unit square into several parts, each of them is either a homothetic image of the body shown on Fig. 3 or its complement. This set is a solution to a capacity constrained problem and appeared for the first time in [28]: find a function $0\le h\le 3$ on $[0,1]^2$ maximizing integral

$$\int_A xyh(x,y)\;dxdy$$

such that $h(x,y)dxdy$ has the Lebesgue projections onto both axes. Then the solution h takes values in $\{0,3\}$ and $\{h=3\}$ is the body on Fig. 3. The precise construction relating these two problems is fairly tedious and we will not give its description here. It can be found in [39].

It seems to be a highly nontrivial task to give the precise description of Fig. 3. This is especially difficult, because numerical experiments demonstrate that it coincides up to a very small set with a figure, which boundary is piecewise smooth and can be parametrized by piecewise elementary functions (polynomials).

Among the other problems which can be "embedded" into the linearly constrained transportation problem let us mention the martingale transportation problem [23,2], problems with symmetries [17,26,27].

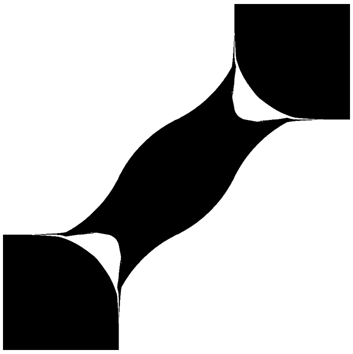

Fig. 3. The support of a solution to a capacity constrained problem (see [29,28]). Compare to Fig. 2.

Finally, there is a connection between the multistochastic problem and the transportation problem with convex constraints, in particular, problems on the space of measures with given ordering. In particular, in the $(3, 2)$-problem with the cost function xyz the natural ordering on the space of measure is ordering by first-order stochastic dominance, i.e. for two measures μ, ν on the plane we say that μ is not less than ν if the distribution function F_μ is not less than F_ν (see Remark 2.26). We plan to study the related modified $(3, 2)$-problem in the subsequent work. Here we just mention that there are many recent paper with very interesting results dealing with convex ordering and optimal transportation, see [21,20].

3. Existence of a uniting measure for (n, k)-problem

3.1. Setting of the problem, basic facts

Unlike the classical Monge–Kantorovich problem, existence of a uniting measure for a (n, k)-problem is a nontrivial task. In the multimarginal Monge–Kantorovich problem, which is a particular case of (n, k)-problem with $k = 1$, the uniting measure always exists: this is $\prod_{i=1}^n \mu_i$. In the case of (n, k)-problem one has the following necessary condition:

Proposition 3.1. *Assume that the set $\Pi(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ is not empty. Let $\mu \in \Pi(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ be arbitrary uniting measure. Then for all $\alpha, \beta \in \mathcal{I}_{nk}$ the following relation holds:*

$$\mathrm{Pr}_{\alpha\cap\beta}(\mu_\alpha) = \mathrm{Pr}_{\alpha\cap\beta}(\mu_\beta) = \mathrm{Pr}_{\alpha\cap\beta}(\mu).$$

Definition 3.2. We say that the collection of measures $\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$ is *consistent*, if it satisfies $\mathrm{Pr}_{\alpha\cap\beta}(\mu_\alpha) = \mathrm{Pr}_{\alpha\cap\beta}(\mu_\beta)$ for all $\alpha, \beta \in \mathcal{I}_{nk}$.

The consistency assumption for $n = 3$, $k = 2$ was considered in [19]. In what follows, we consider only consistent collections of measures. For a consistent collection, the measures μ_β are well-defined for all $\beta \in \mathcal{I}_{nt}$, where $t \le k$. Indeed, denote $\mu_\beta = \mathrm{Pr}_\alpha(\mu_\alpha)$ for arbitrary $\alpha \in \mathcal{I}_{nk}$ containing β. The consistency assumption implies that the result is independent of the choice of α.

Proposition 3.3. *Unlike the multimarginal problem, the consistency assumption is not sufficient for $1 < k < n$.*

Proof. Let $X_i = \{0, 1, \dots, k-1\}$ for all $1 \le i \le n$. For every $\alpha \in \mathcal{I}_{nk}$ let us construct the corresponding measure μ_α on the set X_α. If $\alpha = \{i_1, i_2, \dots, i_k\}$, then every point of X_α is given by coordinates $x = (x_{i_1}, x_{i_2}, \dots, x_{i_k})$, where $x_{i_t} \in \{0, 1, \dots, k-1\}$ for all $1 \le t \le k$. Set $\mu_\alpha(x) = k^{1-k}$, if $\sum_{t=1}^k x_{i_t} \equiv 1 \pmod k$ and $\mu_\alpha(x) = 0$ in the opposite case.

It is easy to check that the consistency assumption of Definition 3.2 holds: the projection of any measure μ_α onto X_β is uniform if $|\beta| < |\alpha|$. Assume that a uniting measure μ exists. Since the projections are non-zero, μ is not zero itself. Take a point $x = (x_1, x_2, \ldots, x_n)$ such that $\mu(x) > 0$. Then for all $\alpha = \{i_1, \ldots, i_k\} \in \mathcal{I}_{nk}$ the relation $\sum_{t=1}^k x_{i_t} \equiv 1 \pmod k$ holds, in the opposite case the μ-mass of the projection of $(x_{i_1}, x_{i_2}, \ldots, x_{i_k})$ onto X_α is zero, hence projection of μ does not coincide with μ_α.

We extract from condition $\sum_{t=1}^k x_{i_t} \equiv 1 \pmod k$, which holds for all $\{i_1, \ldots, i_k\} \in \mathcal{I}_{nk}$, $k < n$ that $x_i \equiv x_j \pmod k$ for all $1 \le i, j \le n$. Then $\sum_{t=1}^k x_{i_t} \equiv k \cdot x_1 \equiv 0 \not\equiv 1 \pmod k$. We obtain a contradiction. □

A different continuous example for $n = 3$, $k = 2$ the reader can find in [19] (Remark 2.3).

3.2. Existence of a signed measure

It follows from the previous proposition that the consistency assumption is not sufficient for existence of a uniting measure. Nevertheless, it is sufficient for existence of a signed measure.

Let $\nu_i \in \mathcal{P}(X_i)$ be an arbitrary family of probability measures.

Definition 3.4. For all $\alpha \in \mathcal{I}_{nt}$, $0 \le t \le k$ let us extend μ_α to X in the following way: $\widetilde{\mu}_\alpha = \mu_\alpha \otimes \prod_{i \notin \alpha} \nu_i$. In addition, set $\widetilde{\mu}_t = \sum_{\alpha \in \mathcal{I}_{nt}} \widetilde{\mu}_\alpha$, where $0 \le t \le k$.

Proposition 3.5. *Let the variables $\{\lambda_t\}_{t=0}^k$ be a solution to the following upper-triangular system of linear equations:*

$$\begin{cases} \displaystyle\sum_{t=i}^{\min(k, n-k+i)} \binom{n-k}{t-i} \lambda_t = 0 \quad \text{for} \quad 0 \le i < k, \\ \displaystyle\sum_{t=k}^{\min(k, n-k+k)} \binom{n-k}{t-k} \lambda_t = 1 \quad \Leftrightarrow \quad \lambda_k = 1. \end{cases} \tag{5}$$

Then for any consistent collection of probability measures $\{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}}$ and for any sequence of probability measures $\{\nu_i\}_{i=1}^n$, the projection of the signed measure $\mu = \sum_{t=0}^k \lambda_t \widetilde{\mu}_t$ on X_α is μ_α for every $\alpha \in \mathcal{I}_{nk}$.

Proof. Introduce the following notation: for $\alpha \in \mathcal{I}_n$, $0 \le t \le k$, and for $\beta \in \mathcal{I}_{nt}$ we define

$$\widetilde{\mu}^\alpha_\beta = \mu_\beta \otimes \prod_{\substack{i \notin \beta \\ i \in \alpha}} \nu_i, \beta \subset \alpha,$$

$$\widetilde{\mu}^\alpha_t = \sum_{\substack{\beta \in \mathcal{I}_{nt} \\ \beta \subset \alpha}} \widetilde{\mu}^\alpha_\beta.$$

Note that this notation generalizes measures $\widetilde{\mu}_\beta$ and $\widetilde{\mu}_t$ introduced in Definition 3.4 in the following sense: if $\alpha = \{1, \ldots, n\}$ then $\widetilde{\mu}^\alpha_\beta = \widetilde{\mu}_\beta$ and $\widetilde{\mu}^\alpha_t = \widetilde{\mu}_t$.

Let us fix $\alpha \in \mathcal{I}_{nk}$. For arbitrary $\beta \in \mathcal{I}_{nt}$, where $t \le k$, the pushforward of $\widetilde{\mu}_\beta$ by the projection of X onto X_α is equal to the product of $\mathrm{Pr}_{\alpha \cap \beta}(\mu_\beta)$ by $\mathrm{Pr}_{\alpha \setminus \beta}\left(\prod_{i \notin \beta} \nu_i\right)$. The first term is $\mu_{\beta \cap \alpha}$, the second term is $\prod_{i \in \alpha \setminus \beta} \nu_i$. It is easy to realize that their product is $\widetilde{\mu}^\alpha_{\beta \cap \alpha}$.

Let us project $\widetilde{\mu}_t$ onto X_α. By the definition of $\widetilde{\mu}_t$, one can get

$$\mathrm{Pr}_\alpha(\widetilde{\mu}_t) = \sum_{\beta \in \mathcal{I}_{nt}} \widetilde{\mu}^\alpha_{\beta \cap \alpha} = \sum_{i=0}^t \sum_{\substack{\gamma \in \mathcal{I}_{ni} \\ \gamma \subset \alpha}} \binom{n-k}{t-i} \widetilde{\mu}^\alpha_\gamma = \sum_{i=0}^t \binom{n-k}{t-i} \widetilde{\mu}^\alpha_i.$$

Then the projection of $\mu = \sum_{t=0}^{k} \lambda_t \widetilde{\mu}_t$ onto the space X_α can be written as a linear combination of $\widetilde{\mu}_t^\alpha$:

$$\mathrm{Pr}_\alpha(\mu) = \sum_{i=0}^{k} c_i \widetilde{\mu}_i^\alpha,$$

where

$$c_i = \sum_{t=i}^{k} \binom{n-k}{t-i} \lambda_t = \sum_{t=i}^{\min(k,n-k+i)} \binom{n-k}{t-i} \lambda_t.$$

Since the coefficients λ_i solve the linear system (5), we conclude that $c_i = 0$ for all $0 \le i < k$ and $c_k = 1$. Thus,

$$\mathrm{Pr}_\alpha(\mu) = \widetilde{\mu}_k^\alpha.$$

By the definition,

$$\widetilde{\mu}_k^\alpha = \sum_{\substack{\beta \in \mathcal{I}_{nk} \\ \beta \subset \alpha}} \widetilde{\mu}_\beta^\alpha,$$

and the only $\beta \in \mathcal{I}_{nk}$ satisfying the property $\beta \subset \alpha$ is $\beta = \alpha$. So,

$$\widetilde{\mu}_k^\alpha = \widetilde{\mu}_\alpha^\alpha = \mu_\alpha \quad \Rightarrow \quad \mathrm{Pr}_\alpha(\mu) = \mu_\alpha$$

for all $\alpha \in \mathcal{I}_{nk}$. □

The system of linear equations (5) introduced in Proposition 3.5 is upper triangular, and the coefficients placed on the main diagonal are equal to 1. Then this system has a unique solution $\{\lambda_t\}_{t=0}^k$. Thus, the following theorem holds.

Theorem 3.6. *There exists a unique family of real coefficients $\{\lambda_t\}_{t=0}^k$ such that in the (n,k)-problem with a consistent family $\{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}}$ of probability measures, a linear combination $\mu = \sum_{t=0}^{k} \lambda_t \widetilde{\mu}_t$ satisfies $\mathrm{Pr}_\alpha(\mu) = \mu_\alpha$ for all $\alpha \in \mathcal{I}_{nk}$.*

Example 3.7. Let us give an example in the $(3,2)$-case. One has

$$\begin{aligned}
\widetilde{\mu}_0 &= \nu_1 \otimes \nu_2 \otimes \nu_3, \\
\widetilde{\mu}_1 &= \mu_1 \otimes \nu_2 \otimes \nu_3 + \nu_1 \otimes \mu_2 \otimes \nu_3 + \nu_1 \otimes \nu_2 \otimes \mu_3, \\
\widetilde{\mu}_2 &= \mu_{12} \otimes \nu_3 + \mu_{13} \otimes \nu_2 + \mu_{23} \otimes \nu_1.
\end{aligned}$$

The projections of these measures onto $X_1 \times X_2$ are given by

$$\begin{aligned}
\mathrm{Pr}_{12}(\widetilde{\mu}_0) &= \nu_1 \otimes \nu_2, \\
\mathrm{Pr}_{12}(\widetilde{\mu}_1) &= \mathrm{Pr}_{12}(\mu_1 \otimes \nu_2 \otimes \nu_3) + \mathrm{Pr}_{12}(\nu_1 \otimes \mu_2 \otimes \nu_3) + \mathrm{Pr}_{12}(\nu_1 \otimes \nu_2 \otimes \mu_3) \\
&= \mu_1 \otimes \nu_2 + \nu_1 \otimes \mu_2 + \nu_1 \otimes \nu_2, \\
\mathrm{Pr}_{12}(\widetilde{\mu}_2) &= \mathrm{Pr}_{12}(\mu_{12} \otimes \nu_3) + \mathrm{Pr}_{12}(\mu_{13} \otimes \nu_2) + \mathrm{Pr}_{12}(\mu_{23} \otimes \nu_1) \\
&= \mu_{12} + \mu_1 \otimes \nu_2 + \nu_1 \otimes \mu_2.
\end{aligned}$$

Thus for arbitrary coefficients λ_0, λ_1, λ_2 one can find projection of $\lambda_0\widetilde{\mu}_0 + \lambda_1\widetilde{\mu}_1 + \lambda_2\widetilde{\mu}_2$ onto $X_1 \times X_2$:

$$\mathrm{Pr}_{12}(\lambda_0\widetilde{\mu}_0 + \lambda_1\widetilde{\mu}_1 + \lambda_2\widetilde{\mu}_2) = (\lambda_0 + \lambda_1)\nu_1 \otimes \nu_2 + (\lambda_1 + \lambda_2)(\mu_1 \otimes \nu_2 + \nu_1 \otimes \mu_2) + \lambda_2\mu_{12}.$$

In order to have equality $\mathrm{Pr}_{12}(\lambda_0\widetilde{\mu}_0+\lambda_1\widetilde{\mu}_1+\lambda_2\widetilde{\mu}_2) = \mu_{12}$ it is sufficient to require $\lambda_0+\lambda_1 = 0$, $\lambda_1+\lambda_2 = 0$, $\lambda_2 = 1$. This system has a unique solution $\lambda_0 = 1$, $\lambda_1 = -1$, $\lambda_2 = 1$. Thus $\mathrm{Pr}_{12}(\widetilde{\mu}_0 - \widetilde{\mu}_1 + \widetilde{\mu}_2) = \mu_{12}$. By the reason of symmetry $\mathrm{Pr}_{13}(\widetilde{\mu}_0 - \widetilde{\mu}_1 + \widetilde{\mu}_2) = \mu_{13}$ and $\mathrm{Pr}_{23}(\widetilde{\mu}_0 - \widetilde{\mu}_1 + \widetilde{\mu}_2) = \mu_{23}$.

3.3. Dual condition for existence of a uniting measure

The following existence criterion for uniting measure is a particular case of a result obtained by Kellerer in [24]. We give an independent proof based on the use of the minimax theorem.

Theorem 3.8. *Let X_1, X_2, $\dots$, X_n be compact metric spaces and let $\mu_\alpha \in \mathcal{P}(X_\alpha)$, $\alpha \in \mathcal{I}_{nk}$ be a fixed family of measures (consistent or not). Then $\Pi(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ is not empty if and only if for every collection of functions $\{f_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$, $f_\alpha \in L^1(X_\alpha, \mu_\alpha)$ satisfying assumption $\sum_{\alpha\in I_{nk}} f_\alpha(x_\alpha) \geq 0$ for all $x \in X$ the following inequality holds:*

$$\sum_{\alpha\in\mathcal{I}_{nk}} \int_{X_\alpha} f_\alpha \, d\mu_\alpha \geq 0.$$

Proof. The existence of a uniting measure trivially implies the inequality. If $\mu \in \Pi(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ and the set of functions f_α satisfies the assumption of the theorem, the function $F(x) = \sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(x_\alpha)$ is integrable with respect to μ and the following inequality holds:

$$\sum_{\alpha\in\mathcal{I}_{nk}} \int_{X_\alpha} f_\alpha \, d\mu_\alpha = \int_X F \, d\mu \geq \int_X 0 \, d\mu = 0.$$

Let us prove the theorem in the other direction. Assume that the collection of measures $\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$ does not satisfy assumptions of Definition 3.2. Then there exists $\alpha, \beta \in \mathcal{I}_{nk}$, such that the measures $\nu_1 = \mathrm{Pr}_{\alpha\cap\beta}\mu_\alpha$ and $\nu_2 = \mathrm{Pr}_{\alpha\cap\beta}\mu_\beta$ are different. Let A be a subset of $X_{\alpha\cap\beta}$ satisfying $\nu_1(A) < \nu_2(A)$. Set: $f_\alpha(x_\alpha) = 1$ if $x_{\alpha\cap\beta} \in A$ and 0 in the opposite case. In addition, set $f_\beta(x_\beta) = -1$ if $x_{\alpha\cap\beta} \in A$, and 0 in the opposite case; $f_\gamma(x_\gamma) = 0$, if $\gamma \notin \{\alpha, \beta\}$. Then $\sum_{\gamma\in\mathcal{I}_{nk}} f_\gamma(x_\gamma) = 0$ for all $x \in X$. On the other hand

$$\sum_{\gamma\in\mathcal{I}_{nk}} \int_{X_\gamma} f_\gamma \, d\mu_\gamma = \int_{X_\alpha} f_\alpha \, d\mu_\alpha + \int_{X_\beta} f_\beta \, d\mu_\beta = \nu_1(A) - \nu_2(A) < 0.$$

Thus, one can assume without loss of generality that the collection of measures $\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$ satisfies Definition 3.2. We apply the following version of the minimax theorem (see [8,36]):

Theorem 3.9 *(Fenchel-Rockafellar Duality). Let E be a normed vector space and E^* be the corresponding dual space. Consider convex functions Φ and Ψ on E, taking values in $\mathbb{R} \cup \{+\infty\}$. Let Φ^* and Ψ^* be the corresponding Legendre transforms. In addition, assume that there exists $z \in E$ satisfying $\Phi(z) < +\infty$, $\Psi(z) < +\infty$. Then*

$$\inf_E [\Psi + \Phi] = \max_{z\in E^*} [-\Phi^*(-z) - \Psi^*(z)].$$

Let E be the space of continuous bounded functions on X equipped with the uniform convergence norm $|| \, ||_\infty$. According to Radon theorem E^* is the space of finite signed measures on X equipped with the full variation norm. Set:

$$\Phi : u \in C_b(X) \to \begin{cases} 0, \text{ if } u \geq 0, \\ +\infty \text{ otherwise.} \end{cases}$$

$$\Psi : u \in C_b(X) \to \begin{cases} \sum_{\alpha \in \mathcal{I}_{nk}} \int_{X_\alpha} u_\alpha \, d\mu_\alpha, \text{ if } u(x) = \sum_{\alpha \in \mathcal{I}_{nk}} u_\alpha(x_\alpha) \text{ for } u_\alpha \in C_b(X_\alpha), \\ +\infty \text{ otherwise.} \end{cases}$$

Function Ψ does not depend on representation of u as sum of u_α. Indeed, if μ is a signed measure satisfying $\mathrm{Pr}_\alpha \mu = \mu_\alpha$ for all $\alpha \in \mathcal{I}_{nk}$, then $\int_X u \, d\mu = \sum_{\alpha \in \mathcal{I}_{nk}} \int_{X_\alpha} u_\alpha \, d\mu_\alpha$. The signed measure μ exists by Theorem 3.6. It is easy to check that functions Ψ and Φ are convex; in addition, function $u \equiv 1$ satisfies $\Phi(u) < +\infty$ and $\Psi(u) < +\infty$, so by the minimax theorem the following equality holds:

$$\inf_E [\Psi + \Phi] = \max_{z \in E^*} [-\Phi^*(-z) - \Psi^*(z)].$$

It is easy to check that

$$\inf_E [\Phi + \Psi] = \inf_{\sum u_\alpha \geq 0} \sum \int_{X_\alpha} u_\alpha \, d\mu_\alpha.$$

Let us find $\Phi^*(-\pi)$.

$$\Phi^*(-\pi) = \sup_{u \geq 0} \left[-\int_X u \, d\pi \right] = -\inf_{u \geq 0} \int_X u \, d\pi$$

If π is nonnegative, then $\int_X u \, d\pi \geq 0$ for all $u \geq 0$. Otherwise $\int_X u \, d\pi$ can take arbitrary small values. Hence

$$\Phi^*(-\pi) = \begin{cases} 0, \text{ if } \pi \geq 0, \\ +\infty, \text{ otherwise.} \end{cases}$$

In the same way we check that

$$\Psi^*(\pi) = \begin{cases} 0, \text{ if } \mathrm{Pr}_\alpha \pi = \mu_\alpha, \\ +\infty, \text{ otherwise.} \end{cases}$$

Thus the maximum $\max_{\pi \in E^*} [-\Phi^*(-\pi) - \Psi^*(\pi)]$ equals 0, if there exists a nonnegative uniting measure, otherwise it equals $-\infty$. In particular, if a uniting measure does not exist, then $\inf_{\sum f_\alpha \geq 0} \sum \int_{X_\alpha} f_\alpha \, d\mu_\alpha = -\infty$. Hence there exist continuous functions f_α satisfying $\sum \int_{X_\alpha} f_\alpha \, d\mu_\alpha < 0$. $\square$

3.4. Sufficient condition for existence of a uniting measure

Let us mention the following trivial sufficient condition for existence of uniting measure.

Proposition 3.10. *Assume that there exists a family of measures $\nu_i \in \mathcal{P}(X_i)$, $1 \leq i \leq n$, such that $\mu_\alpha = \prod_{i \in \alpha} \nu_i$ for all $\alpha \in \mathcal{I}_{nk}$. Then the set $\Pi(\{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}})$ is non-empty and $\prod_{i=1}^n \nu_i$ is a uniting measure.*

We generalize this sufficient condition using Theorem 3.6.

Theorem 3.11 *(Density condition). For given natural numbers $1 \le k < n$ there exists a constant $\lambda_{nk} > 1$ which admits the following property.*

Assume we are given a consistent family of probability measures $\mu_\alpha \in \mathcal{P}(X_\alpha)$, $\alpha \in \mathcal{I}_{nk}$, and another family of probability measures $\nu_i \in \mathcal{P}(X_i)$, $1 \le i \le n$. Assume that every measure μ_α, $\alpha \in \mathcal{I}_{nk}$, is absolutely continuous with respect to $\nu_\alpha = \prod_{i\in\alpha} \nu_i$:

$$\mu_\alpha = \rho_\alpha \cdot \nu_\alpha.$$

Finally, assume that there exist constants $0 < m \le M$ such that every density ρ_α satisfies $m \le \rho_\alpha \le M$ ν_α-almost everywhere for all $\alpha \in \mathcal{I}_{nk}$.

Then $\Pi(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ is not empty provided $\frac{M}{m} \le \lambda_{nk}$.

Proof. The definition of m implies that $\mu_\alpha - m \cdot \nu_\alpha$ is a nonnegative measure for all $\alpha \in \mathcal{I}_{nk}$, hence $m \le 1$, because both μ_α and ν_α are probability measures. In addition, if $m = 1$, the $\mu_\alpha - \nu_\alpha = 0$ for all $\alpha \in \mathcal{I}_{nk}$. In this case the measure $\nu = \prod_{i=1}^n \nu_i$ is uniting.

Consider the case $m < 1$. Note that $\mu'_\alpha = (\mu_\alpha - m \cdot \nu_\alpha)/(1 - m)$ is a probability measure for all $\alpha \in \mathcal{I}_{nk}$, which is absolutely continuous with respect to ν_α and its density is bounded from above by $\frac{m}{1-m}(\lambda_{nk} - 1) > 0$. In addition, the family of measures μ'_α satisfies consistency condition. Theorem 3.6 implies that given measures ν_i and μ'_α one can construct a family of measures $\widetilde{\mu}'_t$ and find numbers λ_t such that the signed measure $\sum_{t=0}^k \lambda_t \widetilde{\mu}'_t$ is uniting. Note that μ'_α is absolutely continuous with respect to ν_α for all $\alpha \in \mathcal{I}_{nt}$, $1 \le t \le k$, moreover, its density is bounded from above by $\frac{m}{1-m}(\lambda_{nk} - 1)$. This means that the same condition holds for $\widetilde{\mu}'_\alpha$, where we consider the corresponding density with respect to $\nu = \prod_{i=1}^n \nu_i$. Hence $\widetilde{\mu}'_t$ is absolutely continuous with respect to ν and its density is bounded almost everywhere by $\binom{n}{t} \cdot \frac{m}{1-m}(\lambda_{nk} - 1)$.

We infer from this that the density of the signed uniting measure $\mu' = \sum_{t=0}^k \lambda_k \widetilde{\mu}'_t$ is bounded from below by $-\sum_{t=0}^k |\lambda_t| \binom{n}{t} \frac{m}{1-m}(\lambda_{nk} - 1) = -C \cdot \frac{m}{1-m}(\lambda_{nk} - 1)$, where C depends on (n, k) only.

Let us prove that the assertion of the theorem holds for $\lambda_{nk} = 1 + \frac{1}{C}$. For the set of measures μ'_α we constructed a uniting signed measure μ' which density with respect to ν is almost everywhere bounded from below by number $-C \cdot \frac{m}{1-m}(\lambda_{nk} - 1) = -\frac{m}{1-m}$. Then $\mu = (1 - m)\mu' + m\nu$ is a uniting measure for the family $\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$, and its density is nonnegative ν-almost everywhere, hence μ is nonnegative. $\square$

Thus we obtained a sufficient condition for existence of uniting measure for a wide class of functions. Moreover, the uniting measure obtained in Theorem 3.11 admits a bounded density. However, it is often helpful to require density to be bounded away from zero.

Definition 3.12. We say that measures μ and ν on the same measurable space $(X, \mathcal{F})$ are *uniformly equivalent*, if there exists a Radon–Nikodym density ρ of μ with respect to ν, which is bounded from above and from below by positive constants: $0 < m \le \rho(x) \le M$ for all $x \in X$.

In particular, uniformly equivalent measures are absolutely continuous with respect to each other.

The existence of uniting measure uniformly equivalent to $\prod_{i=1}^n \nu_i$ is stronger than the existence of any uniting measure (see Example 5.10). We will call such measures **reducible** later (see Definition 5.8) and we will need the existence of a reducible uniting measure to prove the existence of dual solution. Luckily we have a similar theorem.

Theorem 3.13 *(Uniformly equivalent density condition). Under assumption of Theorem 3.11 there exists constant $\widehat{\lambda}_{nk} > 1$ with the following property. If all $\alpha \in \mathcal{I}_{nk}$ satisfy $m \le \rho_\alpha \le M$ ν_α-almost everywhere* *and* $\frac{M}{m} \le \widehat{\lambda}_{nk}$*, then the set* $\Pi(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ *contains at least one measure which is uniformly equivalent to* $\prod_{i=1}^n \nu_i$.

Proof. The proof follows the lines of the proof of Theorem 3.11. There we introduced constant C. One can check that if $\widehat{\lambda}_{nk} < 1 + \frac{1}{C}$ then the density of the constructed measure μ is separated from zero. It is also obvious that this measure is bounded. □

Remark 3.14. We are interested in maximal possible values for λ_{nk} and $\widehat{\lambda}_{nk}$ in Theorem 3.11 and Theorem 3.13. Though their values constructed in the provided proofs are arbitrarily close to each other, that may not be the case for their maximal values.

3.5. *Estimates for* $(3,2)$*-case*

In the $(3,2)$-case one can obtain explicit estimates on the optimal value of λ_{32} from Theorem 3.11.

Proposition 3.15. *For* $\lambda_{32} > 2$ *the conclusion of Theorem 3.11 does not hold.*

Proof. Let $X_1 = X_2 = X_3 = \{0,1\}$ and let every ν_i be the uniform probability measure on X_i. Let us construct measures $\mu_{12}, \mu_{13}, \mu_{23}$ on spaces $X_1 \times X_2$, $X_1 \times X_3$ and $X_2 \times X_3$ respectively. Consider the positive numbers m and M such that $M/m = \lambda_{32}$ and $2(m+M) = 1$. Set $\mu_{ij}(x_i, x_j) = M$, if $x_i + x_j = 1$; and $\mu_{ij}(x_i,x_j) = m$ otherwise. The constructed measures are probability measures: $\mu_{ij}(X_i \times X_j) = 1$, which follows from the equation $2(m+M) = 1$.

Assume that a uniting measure μ exists. Consider the following sums:

$$\begin{aligned}
A = 6m &= \mu_{12}(0,0) + \mu_{12}(1,1) + \mu_{13}(0,0) + \mu_{13}(1,1) + \mu_{23}(0,0) + \mu_{23}(1,1) \\
&= 3\mu(0,0,0) + \mu(1,0,0) + \mu(0,1,0) + \mu(0,0,1) \\
&\quad + \mu(1,1,0) + \mu(1,0,1) + \mu(0,1,1) + 3\mu(1,1,1), \\
B = 6M &= \mu_{12}(0,1) + \mu_{12}(1,0) + \mu_{13}(0,1) + \mu_{13}(1,0) + \mu_{23}(0,1) + \mu_{23}(1,0) \\
&= 2\mu(1,0,0) + 2\mu(0,1,0) + 2\mu(0,0,1) + 2\mu(1,1,0) + 2\mu(1,0,1) + 2\mu(0,1,1).
\end{aligned}$$

On one hand $2A < B$, because $2m < M$. On the other hand, analyzing expressions on the right-hand sides we see that $2A \ge B$. We get a contradiction. □

Proposition 3.16. *The conclusion of Theorem 3.13 holds for* $\widehat{\lambda}_{32} = \frac{3}{2}$*. In particular, there exists a uniting measure* μ*, which is uniformly equivalent to* $\nu = \nu_1 \otimes \nu_2 \otimes \nu_3$.

Proof. Let $0 < m \le M$ be constants from Theorem 3.11: $m \le \rho_{ij} \le M$ for all $1 \le i < j < M$ ν_{ij}-almost everywhere. Clearly, $m \le 1 \le M$. If $m = 1$ or $M = 1$, then $\mu_{ij} = \nu_{ij}$, this means that ν is a uniting measure itself.

For $m < 1 < M$, we claim that the following expression for μ gives us a nonnegative uniting measure.

$$\begin{aligned}
\mu = 4\mu_1 \otimes \mu_2 \otimes \mu_3 &- 2\left(\nu_1 \otimes \mu_2 \otimes \mu_3 + \mu_1 \otimes \nu_2 \otimes \mu_3 + \mu_1 \otimes \mu_2 \otimes \nu_3\right) \\
&+ 2\left(\mu_{12} \otimes \nu_3 + \mu_{13} \otimes \nu_2 + \mu_{23} \otimes \nu_1\right) - \left(\mu_{12} \otimes \mu_3 + \mu_{13} \otimes \mu_2 + \mu_{23} \otimes \mu_1\right).
\end{aligned}$$

Note that contrary to the proof of Theorem 3.11 this measure is not a linear combination of $\widetilde{\mu}_\alpha$ defined in Definition 3.4.

Let us first check that μ is nonnegative. To this end we prove that its density with respect to $\nu = \nu_1 \otimes \nu_2 \otimes \nu_3$ is nonnegative almost everywhere. The density of μ with respect to ν has the form

$$
\begin{aligned}
\frac{d\mu}{d\nu}(x_1,x_2,x_3) = {} & 4\rho_1(x_1)\rho_2(x_2)\rho_3(x_3) - 2\left(\rho_1(x_1)\rho_2(x_2) + \rho_1(x_1)\rho_3(x_3) + \rho_2(x_2)\rho_3(x_3)\right) \\
& + 2\left(\rho_{12}(x_1,x_2) + \rho_{13}(x_1,x_3) + \rho_{23}(x_2,x_3)\right) \\
& - \left(\rho_{12}(x_1,x_2)\rho_3(x_3) + \rho_{13}(x_1,x_3)\rho_2(x_2) + \rho_{23}(x_2,x_3)\rho_1(x_1)\right).
\end{aligned}
$$

Assumption $m \le \rho_{ij}(x_i,x_j) \le M$ implies that, for ν_i-almost all x_i the inequality $m \le \rho_i(x_i) \le M$ holds, where $\rho_i = \frac{d\mu_i}{d\nu_i}$. The assumption of the theorem implies $1 < M \le \widehat{\lambda}_{32} m = \frac{3}{2}m$. Thus, it is sufficient to check inequality

$$
4p_1p_2p_3 - 2(p_1p_2 + p_1p_3 + p_2p_3) + 2(p_{12} + p_{13} + p_{23}) - (p_1p_{23} + p_2p_{13} + p_3p_{12}) \ge 0
$$

for all $m \le p_i, p_{ij} \le \frac{3}{2}m$, $\frac{2}{3} < m < 1$, and for the proof of uniform boundedness it is sufficient to prove that there exists constant $\varepsilon(m) > 0$ such that

$$
d(p_i,p_{ij}) = 4p_1p_2p_3 - 2(p_1p_2 + p_1p_3 + p_2p_3) + 2(p_{12} + p_{13} + p_{23}) - (p_1p_{23} + p_2p_{13} + p_3p_{12}) \ge \varepsilon(m).
$$

For each fixed m, the 6-tuple of variables p_i and p_{ij} lies in a compact K_m given by conditions $m \le p_i, p_{ij} \le \frac{3}{2}m$. Function d is continuous, so it achieves a minimum at some point $p \in K_m$. Note that d is linear in every variable p_i, p_{ij}, thus p can be taken such that every variable equals m or $\frac{3}{2}m$ at p. The coefficient of p_{ij} equals $2 - p_k > 0$ provided $p_k \le \frac{3}{2}m < \frac{3}{2}$, hence this function is strictly increasing in p_{ij}. Then at p one has $p_{ij} = m$ for all $1 \le i,j \le 3$. Finally, we reduce the proof to the following inequality we have to check:

$$
4p_1p_2p_3 - 2(p_1p_2 + p_1p_3 + p_2p_3) - m(p_1 + p_2 + p_3) + 6m \ge \varepsilon(m)
$$

where all $p_i \in \left\{m, \frac{3}{2}m\right\}$, $\frac{2}{3} < m < 1$.

Since the function is symmetric we have to check the following inequalities:

1. $p_1 = p_2 = p_3 = m$: $4m^3 - 9m^2 + 6m > 0$ if $\frac{2}{3} < m < 1$;
2. $p_1 = \frac{3}{2}m$, $p_2 = p_3 = m$: $6m^3 - \frac{23}{2}m^2 + 6m > 0$ if $\frac{2}{3} < m < 1$;
3. $p_1 = p_2 = \frac{3}{2}m$, $p_3 = m$: $9m^3 - \frac{29}{2}m^2 + 6m > 0$ if $\frac{2}{3} < m < 1$;
4. $p_1 = p_2 = p_3 = \frac{3}{2}m$: $\frac{27}{2}m^3 - 18m^2 + 6m > 0$ if $\frac{2}{3} < m < 1$.

Every inequality can be easily checked and we complete the proof of nonnegativity of μ and its uniform equivalence to ν.

It remains to check that μ is uniting for μ_{ij}:

$$
\begin{aligned}
\mathrm{Pr}_{12}(\mu) = {} & 4\mu_1 \otimes \mu_2 - 2\nu_1 \otimes \mu_2 - 2\mu_1 \otimes \nu_2 - 2\mu_1 \otimes \mu_2 \\
& + 2\mu_{12} + 2\mu_1 \otimes \nu_2 + 2\nu_1 \otimes \mu_2 - \mu_{12} - \mu_1 \otimes \mu_2 - \mu_1 \otimes \mu_2 = \mu_{12}.
\end{aligned}
$$

In the same way we check that the desired identities hold for other projections. □

One can prove another estimate for $\lambda_{32} = 2$. Unfortunately, the arguments in our proof can not be used to prove uniform equivalence of μ and ν.

Proposition 3.17. *For the value $\lambda_{32} = 2$ the conclusion of Theorem 3.11 holds. Together with Proposition 3.15, we get that 2 is the greatest possible value for λ_{32}.*

Proof. The proof we have is fairly long. It lasts till page 27 and contains a series of rather technical lemmas.

Let $0 < m \le M$ be constants from Theorem 3.11. Consider the following set:

$$\Delta = \left\{ \xi - \text{nonnegative measure on } X_1 \times X_2 \times X_3 \colon \frac{d(\mu_{ij} - \mathrm{Pr}_{ij}(\xi))}{d\nu_{ij}} \ge_{\text{a.e.}} m \text{ for all } \{i,j\} \in \mathcal{I}_{3,2} \right\}.$$

This set is not empty because it contains the trivial (zero) measure. In addition, Δ is weakly closed. From assumption $m \le \frac{d(\mu_{ij} - \mathrm{Pr}_{ij}(\xi))}{d\nu_{ij}}$ we infer that $\mu_i \ge \mathrm{Pr}_i(\xi)$, hence Δ is uniformly tight and the variations of measures from Δ are uniformly bounded. Then the Prokhorov theorem for non-probability measures (see [5, Volume II, Theorem 8.6.2]) implies that Δ is weakly compact. Hence there exists an extreme measure $\xi_{\max}$, where functional $\xi_{\max}(X)$ attains its maximum.

Remark 3.18. Later $\xi_{\max}$ will be one of the components of our measure, so note that we can say nothing about its Radon–Nikodym density with respect to ν. That means that this argument can not be generalized to prove the existence of uniting measure uniformly equivalent to ν and so says nothing about $\widehat{\lambda}_{32}$. In fact we do not know if $\widehat{\lambda}_{32}$ can be taken arbitrarily close to 2.

For all $\{i,j\} \in \mathcal{I}_{3,2}$ Radon–Nikodym derivative $\frac{d(\mu_{ij} - \mathrm{Pr}_{ij}(\xi_{\max}))}{d\nu_{ij}}$ is defined up to measure 0. So let us fix some realization and use it later.

Lemma 3.19. *For ν-almost all $x \in X$ there exists a couple $\{i,j\} \in \mathcal{I}_{3,2}$ such that*

$$\frac{d(\mu_{ij} - \mathrm{Pr}_{ij}(\xi_{\max}))}{d\nu_{ij}}(x_i, x_j) = m.$$

Proof. Assume the converse. Then for the set

$$E = \left\{ x \in X \colon \frac{d(\mu_{ij} - \mathrm{Pr}_{ij}(\xi_m))}{d\nu_{ij}}(x_i, x_j) > m \quad \text{for all } \{i,j\} \in \mathcal{I}_{3,2} \right\}$$

we have $\nu(E) > 0$. Then there exists $\varepsilon > 0$ such that

$$E_\varepsilon = \left\{ x \in X \colon \frac{d(\mu_{ij} - \mathrm{Pr}_{ij}(\xi_m))}{d\nu_{ij}}(x_i, x_j) \ge m + \varepsilon \quad \text{for all } \{i,j\} \in \mathcal{I}_{3,2} \right\}$$

satisfies $\nu(E_\varepsilon) > 0$. Let ξ_Δ be the measure which density (with respect to ν) equals ε on E_ε and 0 otherwise. It is easy to check that $\xi_{\max} + \xi_\Delta \in \Delta$, $(\xi_{\max} + \xi_\Delta)(X) > \xi_{\max}(X)$ and this contradicts to definition of $\xi_{\max}$. $\square$

Consider the family of probability measures

$$\mu'_{ij} = \frac{\mu_{ij} - \mathrm{Pr}_{ij}(\xi_{\max})}{1 - \xi_{\max}(X)}, 1 \le i, j \le 3.$$

Since $\{\mu_{ij}\}$ is consistent, the family of measures $\{\mu'_{ij}\}$ is consistent too. Since $\xi_{\max} \in \Delta$, we have $m/\alpha \le d\mu'_{ij}/d\nu_{ij} \le M/\alpha$ almost everywhere, where $\alpha = 1 - \xi_{\max}(X)$. Hence, the family $\{\mu'_{ij}\}$ satisfies assumptions of Proposition 3.17. Moreover, if a measure μ' is uniting for μ'_{ij}, then the measure $\mu = \alpha\mu' + \xi_{\max}$ is uniting for μ_{ij}. Thus, it is sufficient to solve the problem only for μ'_{ij}.

Now, we replace μ_{ij} with μ'_{ij}, m and M with m/α and M/α respectively. We may assume that densities $\rho_i = \frac{d\mu_i}{d\nu_i}$, $\rho_{ij} = \frac{d\mu_{ij}}{d\nu_{ij}}$ satisfying the following assumptions:

1. $m \le \rho_{ij}(x_i, x_j) \le M$, $1 \le i, j \le 3$ for all $x \in X$.
2. $\int_{X_j} \rho_{ij}(x_i, x_j)\, \nu_j(dx_j) = \rho_i(x_i)$ for all $x_i \in X_i$.
3. For ν-almost all $x \in X$ at least one of the numbers $\rho_{ij}(x_i, x_j)$, $1 \le i, j \le 3$, equals m.

Assumptions 1 and 2 are always fulfilled after changing ρ_i and ρ_{ij} on a set of zero measure, and the last one follows from Lemma 3.19. Under these assumptions one can prove the following lemma:

Lemma 3.20. *Assume that ρ_i, ρ_{ij} satisfy Assumptions 1-3. Then for ν_{ij}-almost all $(x_i, x_j) \in X_{ij}$ one of the following conditions holds: $\rho_{ij}(x_i, x_j) = m$ or $\rho_i(x_i) + \rho_j(x_j) \le m + M$.*

Proof. Let $k \in \{1, 2, 3\} \setminus \{i, j\}$. Let us denote by X^r_{ij} the set of couples $(x_i, x_j) \in X_{ij}$ such that for ν_k-almost all $x_k \in X_k$ one of the numbers $\rho_{ij}(x_i, x_j)$, $\rho_{ik}(x_i, x_k)$ and $\rho_{jk}(x_j, x_k)$ equals m. Assumption 3 implies that X^r_{ij} has full measure with respect to ν_{ij}.

Let $(x_i, x_j) \in X^r_{ij}$. Assume that $\rho_{ij}(x_i, x_j) > m$. The for ν_k-almost all $x_k \in X_k$ at least one of the numbers $\rho_{ik}(x_i, x_k)$ and $\rho_{jk}(x_j, x_k)$ equals m. In particular, $\rho_{ik}(x_i, x_k) + \rho_{jk}(x_j, x_k) \le m + M$ for ν_k-almost all $x_k \in X_k$. Then we infer from 1, 2

$$\rho_i(x_i) + \rho_j(x_j) = \int\limits_{X_k} \rho_{ik}(x_i, x_k)\, d\nu_k + \int\limits_{X_k} \rho_{jk}(x_j, x_k)\, d\nu_k \le m + M. \quad \square$$

Changing, if necessary, density functions ρ_i, ρ_{ij} on a set of zero measure, we can assume, in addition, that the following holds:

4. For all $(x_i, x_j) \in X_{ij}$ one has $\rho_{ij}(x_i, x_j) = m$ or $\rho_i(x_i) + \rho_j(x_j) \le m + M$, $1 \le i, j \le 3$.

Lemma 3.21. *Let the density functions ρ_i, ρ_{ij} satisfy Assumptions 1-4. Then for all $i \ne j$ and all $x_i \in X_i$ the following inequality holds:*

$$\nu_j\left(x_j \in X_j \colon \rho_j(x_j) \le m + M - \rho_i(x_i)\right) \ge \frac{\rho_i(x_i) - m}{M - m}.$$

Proof. Fix a point $x_i \in X_i$, and denote by A be the set of points $x_j \in X_j$ satisfying $\rho_{ij}(x_i, x_j) = m$. Then $\rho_i(x_i) = \int_{X_j} \rho_{ij}(x_i, x_j)\, dx_j \le m\nu_j(A) + M(1 - \nu_j(A))$, which implies $\nu_j(A) \le \frac{M - \rho_i(x_i)}{M - m}$.

On the other hand Assumption 4 implies that for all $x_j \in X_j \setminus A$ the inequality $\rho_i(x_i) + \rho_j(x_j) \le m + M$ holds. Hence

$$\nu_j\left(x_j \in X_j \colon \rho_j(X_j) \le m + M - \rho_i(x_i)\right) \ge \nu_j(X_j \setminus A) = 1 - \nu_j(A) \ge \frac{\rho_i(x_i) - m}{M - m}. \quad \square$$

Choosing a sequence $x_i^{(n)}$ such that $\rho_i(x_i^{(n)}) \to M_i = \sup_{x_i \in X_i} \rho_i(x_i)$ and passing to the limit one gets the following corollary:

Corollary 3.22. *Let $M_i = \sup_{x_i \in X_i} \rho_i(x_i)$. Then for all $j \ne i$ the following inequality holds:*

$$\nu_j(x_j \in X_j \colon \rho_j(X_j) \le m + M - M_i) \ge \frac{M_i - m}{M - m}.$$

Lemma 3.23. *Let ρ_i, ρ_{ij} satisfy Assumptions 1-4 and $\frac{M}{m} \le 2$. Then inequalities*

$$\frac{2}{3} \le m \le 1, \quad p_i(x_i) \le \frac{m}{2}\left(3 + \sqrt{3 - \frac{2}{m}}\right)$$

hold for all $x_i \in X_i$, $1 \le i \le 3$.

Proof. Let $M_i = \sup_{x_i \in X_i} \rho_i(x_i)$. Assume that $M_1 \ge M_2$ and $M_1 \ge M_3$. It is sufficient to check that $\frac{3}{2}m \ge 1$ and $M_1 \le \frac{m}{2}\left(3 + \sqrt{3 - \frac{2}{m}}\right)$.

Assume that $\frac{3}{2}m \ge M_1$. Then, since $M_1 = \sup_{x_1 \in X_1} \rho_1(x_1)$, one has $M_1 \ge 1$. This implies $\frac{3}{2}m \ge M_1 \ge 1$. Moreover, $M_1 \le \frac{3}{2}m \le \frac{m}{2}\left(3 + \sqrt{3 - \frac{2}{m}}\right)$.

Consider the case $M_1 \ge \frac{3}{2}m$. Set $A = \{x_2 \in X_2 \colon \rho_2(x_2) \le m + M - M_1\}$. Then the following holds:

$$\begin{aligned} 1 = \int\limits_{X_2} \rho_2(X_2)\, d\nu_2 &\le (m + M - M_1)\nu_2(A) + M_2\,(1 - \nu_2(A)) \\ &\le (m + M - M_1)\nu_2(A) + M_1\,(1 - \nu_2(A)) = (m + M - 2M_1)\nu_2(A) + M_1. \end{aligned}$$

Corollary 3.22 implies $\nu_2(A) \ge \frac{M_1 - m}{M - m} \ge \frac{M_1}{m} - 1$ (here we use $M \le 2m$). Applying this inequality and the inequality $M_1 \ge \frac{3}{2}m$ one gets

$$\begin{aligned} 1 &\le (m + M - 2M_1)\nu_2(A) + M_1 \le (3m - 2M_1)\nu_2(A) + M_1 \\ &\le (3m - 2M_1)\left(\frac{M_1}{m} - 1\right) + M_1 = m\left(-2\left(\frac{M_1}{m}\right)^2 + 6\frac{M_1}{m} - 3\right). \end{aligned}$$

The function $-2x^2 + 6x - 3$ is decreasing on $x \ge \frac{3}{2}$, hence

$$1 \le m\left(-2\left(\frac{M_1}{m}\right)^2 + 6\frac{M_1}{m} - 3\right) \le m\left(-2\left(\frac{3}{2}\right)^2 + 6\cdot\frac{3}{2} - 3\right) = \frac{3}{2}m.$$

Moreover, $-2\left(\frac{M_1}{m}\right)^2 + 6\frac{M_1}{m} - 3 \ge \frac{1}{m}$, thus $\frac{M_1}{m} \le \frac{1}{2}\left(3 + \sqrt{3 - \frac{2}{m}}\right)$. $\square$

Let us describe explicit constructions of uniting measures for $m = \frac{2}{3}$ and $\frac{2}{3} < m \le 1$. If $m = \frac{2}{3}$, then $\rho_i(x_i) \le \frac{m}{2}\left(3 + \sqrt{3 - \frac{2}{m}}\right) = 1$ for all $x_i \in X_i$. Measures μ_i and ν_i are probability measures, $\frac{d\mu_i}{d\nu_i} \le 1$. Hence $\mu_i = \nu_i$. The desired measure is given by

$$\mu = \mu_1 \otimes \mu_{23} + \mu_2 \otimes \mu_{13} + \mu_3 \otimes \mu_{12} - 2\mu_1 \otimes \mu_2 \otimes \mu_3.$$

This measure is nonnegative: $\frac{d\mu}{d\nu}(x_1, x_2, x_3) = \rho_{12}(x_1, x_2) + \rho_{13}(x_1, x_3) + \rho_{23}(x_2, x_3) - 2 \ge 0$ since $\rho_{ij}(x_i, x_j) \ge m = \frac{2}{3}$. In addition, it is uniting:

$$\mathrm{Pr}_{12}(\mu) = \mu_1 \otimes \mu_2 + \mu_2 \otimes \mu_1 + \mu_{12} - 2\mu_1 \otimes \mu_2 = \mu_{12},$$

and the same for other projections.

Let us consider the case $\frac{2}{3} < m \le 1$. Set: $u = \sqrt{3 - \frac{2}{m}}$. Then $\frac{1}{m} = \frac{1}{2}(3 - u^2)$; u satisfies $0 < u \le 1$ under assumption $\frac{2}{3} < m \le 1$. The desired measure is given by

$$\begin{aligned} \mu = &-\frac{8}{m^2 u(u+1)^3}\mu_1 \otimes \mu_2 \otimes \mu_3 + 2\frac{5u + 9}{u(u+1)^3}\nu_1 \otimes \nu_2 \otimes \nu_3 \\ &+ 4\frac{u + 3}{mu(u+1)^3}\left(\nu_1 \otimes \mu_2 \otimes \mu_3 + \mu_1 \otimes \nu_2 \otimes \mu_3 + \mu_1 \otimes \mu_2 \otimes \nu_3\right) \end{aligned}$$

$$
\begin{aligned}
&- 2\frac{5u+9}{u(u+1)^3}\left(\mu_1 \otimes \nu_2 \otimes \nu_3 + \nu_1 \otimes \mu_2 \otimes \nu_3 + \nu_1 \otimes \nu_2 \otimes \mu_3\right)\\
&+ 2\frac{u+2}{(u+1)^2}\left(\mu_{23} \otimes \nu_1 + \mu_{13} \otimes \nu_2 + \mu_{12} \otimes \nu_3\right)\\
&- \frac{2}{m(u+1)^2}\left(\mu_{23} \otimes \mu_1 + \mu_{13} \otimes \mu_2 + \mu_{12} \otimes \mu_3\right).
\end{aligned}
$$

This measure is uniting for μ_{ij}:

$$
\begin{aligned}
\mathrm{Pr}_{12}(\mu) = &\left(4\frac{u+3}{mu(u+1)^3} - 2\frac{5u+9}{u(u+1)^3} + 2\frac{u+2}{(u+1)^2}\right)(\nu_1 \otimes \mu_2 + \mu_1 \otimes \nu_2)\\
&+ \left(-\frac{8}{m^2u(u+1)^3} + 4\frac{u+3}{mu(u+1)^3} - \frac{4}{m(u+1)^2}\right)\mu_1 \otimes \mu_2\\
&+ \left(2\frac{5u+9}{u(u+1)^3} - 2\frac{5u+9}{u(u+1)^3}\right)\nu_1 \otimes \nu_2 + \left(2\frac{u+2}{(u+1)^2} - \frac{2}{m(u+1)^2}\right)\mu_{12}\\
= &\ \mu_{12}.
\end{aligned}
$$

To prove the desired equality we substitute $\frac{1}{m} = \frac{1}{2}(3 - u^2)$ and check that all the terms are zero except the last one. In addition, the coefficient of μ_{12} equals 1. We do the same for the other projections.

To check nonnegativity of μ it is sufficient to check that the following expression is nonnegative:

$$
\begin{aligned}
&- 8p_1p_2p_3 + 4m(u+3)(p_1p_2 + p_1p_3 + p_2p_3) - 2m^2(5u+9)(p_1 + p_2 + p_3)\\
&+ 2m^2u(u+1)(u+2)(p_{12} + p_{13} + p_{23}) - 2mu(u+1)(p_1p_{23} + p_2p_{13} + p_3p_{12})\\
&+ 2m^2(5u+9),
\end{aligned}
$$

where $p_i = \rho_i(x_i)$, $p_{ij} = \rho_{ij}(x_i, x_j)$. One has $m \le p_{ij} \le 2m$ by our assumption, $m \le p_i \le \frac{m}{2}\left(3 + \sqrt{3 - \frac{2}{m}}\right) = \frac{m}{2}(u+3)$ and $\frac{2}{3} < m \le 1$ by Lemma 3.23.

This function is linear in p_{ij} with the coefficient

$$
2m^2u(u+1)(u+2) - 2mu(u+1)p_k \ge 2m^2u(u+1)(u+2) - m^2u(u+1)(u+3) \ge 0
$$

(here we use that $u \le 1$), hence one can set $p_{ij} = m$ for all $1 \le i, j \le 3$. In this case the expression is equal to

$$
\begin{aligned}
&- 8p_1p_2p_3 + 4m(u+3)(p_1p_2 + p_1p_3 + p_2p_3) - 2m^2(5u+9)(p_1 + p_2 + p_3)\\
&+ 6m^3u(u+1)(u+2) - 2m^2u(u+1)(p_1 + p_2 + p_3) + 2m^2(5u+9)\\
= &- 8p_1p_2p_3 + 4m(u+3)(p_1p_2 + p_1p_3 + p_2p_3) - 2m^2(u+3)^2(p_1 + p_2 + p_3)\\
&+ 6m^3u(u+1)(u+2) - m^3(u^2-3)(5u+9)\\
= &\ (m(u+3) - 2p_1)(m(u+3) - 2p_2)(m(u+3) - 2p_3) \ge 0.
\end{aligned}
$$

This completes the proof of the well-posedness.

So, indeed, for the value $\lambda_{32} = 2$ the conclusion of Theorem 3.11 holds. This completes the proof of Proposition 3.17. □

One can prove many other sufficient conditions of existence of uniting measures. One of the examples is given in the next theorem.

Theorem 3.24. *Assume that a consistent family of measures $\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$ satisfies $\mu_{ij} \geq \frac{2}{3}\mu_i \otimes \mu_j$, $1 \leq i,j \leq 3$. Then there exists a uniting measure.*

Proof. The desired measure is given by

$$\mu = \left(\mu_{12} - \frac{2}{3}\mu_1 \otimes \mu_2\right) \otimes \mu_3 + \left(\mu_{13} - \frac{2}{3}\mu_1 \otimes \mu_3\right) \otimes \mu_2 + \left(\mu_{23} - \frac{2}{3}\mu_2 \otimes \mu_3\right) \otimes \mu_1.$$

Indeed, one has

$$\mathrm{Pr}_{12}(\mu) = \mu_{12} - \frac{2}{3}\mu_1 \otimes \mu_2 + \mu_1 \otimes \mu_2 - \frac{2}{3}\mu_1 \otimes \mu_2 + \mu_2 \otimes \mu_1 - \frac{2}{3}\mu_1 \otimes \mu_2 = \mu_{12},$$

analogously for other projections. Thus μ is uniting. $\square$

4. Connection to the Monge–Kantorovich problem with linear constraints

4.1. Monge–Kantorovich problem with linear constraints: definitions and basic facts

D. Zaev considered in [38] the multimarginal transportation problem with additional linear constraints. In this subsection we formulate basic definitions and theorems of his paper.

Let $X_1, X_2, \dots, X_n$ be Polish spaces equipped with Borel σ-algebras, $X := X_1 \times \dots \times X_n$, $\mu_1, \dots, \mu_n$ are probability measures on $X_1, \dots, X_n$ respectively.

Let W be an arbitrary linear subspace in $C_L(X, \{\mu_k\}_{k=1}^n)$. Let us consider the following subspace in the set of measures:

$$\Pi_W(\{\mu_k\}_{k=1}^n) = \left\{\pi \in \Pi(\{\mu_k\}_{k=1}^n)\colon \int \omega \, d\pi = 0 \text{ for all } \omega \in W\right\}.$$

Finally, we are ready to formulate our constrained problem:

Problem 4.1 *(Monge–Kantorovich problem with linear constraints). Given Polish spaces $X = X_1 \times \dots X_n$, Borel probability measures $\mu_k \in \mathcal{P}(X_k)$, a cost function $c \in C_L(X, \{\mu_k\}_{k=1}^n)$, and a linear subspace $W \subset C_L(X, \{\mu_k\}_{k=1}^n)$ find*

$$\inf_{\pi\in\Pi_W(\{\mu_k\}_{k=1}^n)} \left\{\int_X c(x)\, d\pi\right\}.$$

The following theorems are the main results of [38]:

Theorem 4.2. *Problem with additional linear constraints has a solution if the set $\Pi_W(\{\mu_k\}_{k=1}^n)$ is not empty.*

Theorem 4.3 *(Kantorovich duality with additional linear constraints). Let $X_1, \dots, X_n$, and $X = X_1 \times \dots \times X_n$ be Polish spaces, let $\mu_k \in \mathcal{P}(X_k)$, $k = 1, \dots, n$, and let W be a linear subspace of $C_L(X, \{\mu_k\}_{k=1}^n)$, $c \in C_L(X, \{\mu_k\}_{k=1}^n)$. Then*

$$\inf_{\pi\in\Pi_W(\{\mu_k\}_{k=1}^n)} \int_X c\, d\pi = \sup_{f+\omega\leq c} \sum_{k=1}^n \int_{X_k} f_k(x_k)\, d\mu_k,$$

where $f(x_1, \dots, x_n) = \sum_{k=1}^n f_k(x_k)$ and $f_k \in C_L(X_k, \mu_k)$, $\omega \in W$. Moreover, if $c \in C_b(X)$ and $W \subset C_b(X)$, then the supremum can be taken on the set of bounded continuous functions $f_k \in C_b(X_k)$.

4.2. A problem with linear constraints which is equivalent to the multistochastic problem

Let us consider again the multistochastic Monge–Kantorovich problem on Polish spaces $X_1, \ldots, X_n$. We are given $\binom{n}{k}$ probability measures μ_α on X_α, where $\alpha \in \mathcal{I}_{nk}$, and a cost function $c : X \to \mathbb{R}$, $X = X_1 \times \cdots \times X_n$. Our aim is to construct an equivalent Monge–Kantorovich problem with linear constraints. Then we can apply Kantorovich duality from Theorem 4.3 for the $\binom{n}{k}$ spaces X_α with marginals μ_α (note that these spaces are themselves "composite") and restrictions correspond to the dependencies of $\{X_\alpha\}$ in X.

In what follows we denote

$$\widetilde{X} = \prod_{\alpha \in \mathcal{I}_{nk}} X_\alpha.$$

For every $\alpha \in \mathcal{I}_{nk}$ we define the corresponding natural projection $\mathrm{Pr}_\alpha \colon \widetilde{X} \to X_\alpha$.

Definition 4.4. For all $\alpha \in \mathcal{I}_{nk}$ and $i \in \alpha$ let us consider projection $\widetilde{x}^i_\alpha := \mathrm{Pr}_{X_i} \circ \mathrm{Pr}_{X_\alpha}$. In what follows $\widetilde{x}^i_\alpha$ denotes the projection operator and, at the same time, the image of $\widetilde{x} \in \widetilde{X}$ under action of this operator. The set $\{\widetilde{x}\}^i_\alpha$ can be viewed as a set of coordinates of $\widetilde{x}$ in $\widetilde{X}$.

Definition 4.5. The subspace $P \subset \widetilde{X}$ will be defined as follows:

$$P = \left\{ \widetilde{x} \in \widetilde{X} \colon \widetilde{x}^i_\alpha = \widetilde{x}^i_\beta \text{ for all } \alpha, \beta \in \mathcal{I}_{nk}, i \in \alpha \cap \beta \right\}.$$

The subspace P can be characterized in terms of a diagonal operator. The space $\widetilde{X}$ is isomorphic to $(X_1 \times \cdots \times X_n)^{\binom{n-1}{k-1}} = X^{\binom{n-1}{k-1}}$: to verify this it is sufficient to interchange factors in the product of spaces $X_\alpha = \prod_{i \in \alpha} X_i$. Let Δ be the diagonal mapping from X onto $\widetilde{X} = X^{\binom{n-1}{k-1}}$. It is easy to see that this mapping is well–defined, because it does not depend on permutation of spaces in the isomorphism $\widetilde{X} \cong (X_1 \times \cdots \times X_n)^{\binom{n-1}{k-1}}$. Hence P is the image of X under action Δ and restriction of Δ on P acts bijectively.

The following properties of Δ are direct consequences of its definition:

Proposition 4.6. *Operator Δ generates an operator $\Delta_* : \mathcal{P}(X) \to \mathcal{P}(\widetilde{X})$ acting on measures, which has the following properties:*

1. *For every measure $\mu \in \mathcal{P}(X)$ the support of $\Delta_*(\mu)$ is a subset of P.*
2. *Operator Δ_* is a bijection between $\mathcal{P}(X)$ and the set of measures $\mu \in \mathcal{P}(\widetilde{X})$ with the property $\mathrm{supp}(\mu) \subset P$.*
3. *Every $\mu \in \mathcal{P}(X)$ and every $\alpha \in \mathcal{I}_{nk}$ satisfy $\mathrm{Pr}_\alpha(\mu) = \mathrm{Pr}_{X_\alpha}(\Delta_*(\mu))$.*
4. *Let μ be an arbitrary probability measure on X and let $c \in L^1(X, \mu)$. Let $\widetilde{c}$ be a measurable function on $\widetilde{X}$ such that $\widetilde{c}(\widetilde{x}) = c(\Delta^{-1}(\widetilde{x}))$ for all $\widetilde{x} \in P$. Then $\widetilde{c} \in L^1(\widetilde{X}, \Delta_*(\mu))$ and $\int_X c \, d\mu = \int_{\widetilde{X}} \widetilde{c} \, d\Delta_*(\mu) = \int_P \widetilde{c} \, d\Delta_*(\mu)$.*

The following theorem is an immediate corollary of these properties

Theorem 4.7. *Let $c \in C_L(X, \{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}})$ be a function on X and $\widetilde{c} \in C_L(\widetilde{X}, \{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}})$ be any extension of $c \circ \Delta^{-1} : P \subset \widetilde{X} \to \mathbb{R}$ to the whole space $\widetilde{X}$. Then*

$$\inf_{\pi \in \Pi(X, \{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}})} \int_X c \, d\pi = \inf_{\substack{\xi \in \Pi(\widetilde{X}, \{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}}) \\ \mathrm{supp}(\xi) \subset P}} \int \widetilde{c} \, d\xi.$$

The minimum on the left-hand side is attained if and only if the minimum on the right-hand side is attained.

Remark 4.8. It is worth to remind the reader that $\widetilde{c} \in C_L(\widetilde{X}, \{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ means that $\widetilde{c}$ is a continuous function on $\widetilde{X}$ and there exists such a collection of functions f_α such that $|\widetilde{c}(\widetilde{x})| \leq \sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(\widetilde{x}_\alpha)$. The existence of such $\widetilde{c}$ extending $c \circ \Delta^{-1}$ is not obvious. We prove it in Lemma 4.10.

Consider the distance function d_i on X_i and the family of functions $\omega^i_{\alpha\beta} : \widetilde{X} \to \mathbb{R}$,

$$\omega^i_{\alpha\beta}(\widetilde{x}) := \min(d_i(\widetilde{x}^i_\alpha, \widetilde{x}^i_\beta), 1)$$

for all $\alpha, \beta \in \mathcal{I}_{nk}$, $i \in \alpha\cap\beta$. Note that every $\omega^i_{\alpha\beta}$ is a nonnegative, continuous, bounded from above function, hence $\omega^i_{\alpha\beta} \in C_b(\widetilde{X}) \subset C_L(\widetilde{X}, \{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$. In addition, if some measure $\mu \in \mathcal{P}(\widetilde{X})$ satisfies $\int \omega^i_{\alpha\beta}\, d\mu = 0$, then $\operatorname{supp}(\mu) \subset (\omega^i_{\alpha\beta})^{-1}(0) = \{\widetilde{x} \in \widetilde{X} : \widetilde{x}^i_\alpha = \widetilde{x}^i_\beta\}$.

Let us define the space of linear restrictions:

$$W := \operatorname{span}\{\omega^i_{\alpha\beta}\} \subset C_b(\widetilde{X}) \subset C_L(\widetilde{X}, \{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}}).$$

It follows from the observations collected above that for every $\pi \in \mathcal{P}(\widetilde{X})$ the equality $\int \omega\, d\pi = 0$ holds for all $\omega \in W$ if and only if $\operatorname{supp}(\pi) \subset P$. Hence

$$\Pi_W(\widetilde{X}, \{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}}) = \{\pi \in \Pi(\widetilde{X}, \{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}}) : \operatorname{supp}(\pi) \subset P\}.$$

Having this in mind, we can give another formulation of Theorem 4.7:

Theorem 4.9. *Let $c \in C_L(X, \{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ be a function on X and $\widetilde{c} \in C_L(\widetilde{X}, \{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ be any extension of $c \circ \Delta^{-1} : P \subset \widetilde{X} \to \mathbb{R}$ to the entire space $\widetilde{X}$. Then*

$$\inf_{\pi\in\Pi(X,\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})} \int_X c\, d\pi = \inf_{\xi\in\Pi_W(\widetilde{X},\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})} \int \widetilde{c}\, d\xi,$$

and the minimum on the left-hand side is attained if and only if it is attained on the right-hand side.

This theorem gives another formulation of the transportation problem with linear constraints which is equivalent to our multistochastic problem. It remains to prove that there exists a function $\widetilde{c}$ which satisfies our requirement.

Lemma 4.10.

a) *Let $c \in C_b(X)$. There exists a function $\widetilde{c} \in C_b(\widetilde{X})$ which is an extension of $c \circ \Delta^{-1}$ onto $\widetilde{X}$.*
b) *Let $c \in C_L(X, \{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$. There exists a function $\widetilde{c} \in C_L(\widetilde{X}, \{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ (note that c and $\widetilde{c}$ belong to different spaces) which is an extension of $c \circ \Delta^{-1}$ onto $\widetilde{X}$.*

Proof. Let pr be the projection of $\widetilde{X} \cong X^{\binom{n-1}{k-1}}$ onto a fixed factor. It is easy to see that pr is continuous and $\operatorname{pr} \circ \Delta = \operatorname{id}$ on X.

a) Assume that $c \in C_b(X)$ and $|c| \leq M$ for some number M. Set $\widetilde{c}(\widetilde{x}) := c(\operatorname{pr}(\widetilde{x}))$. Function $\widetilde{c}$ is continuous, $|\widetilde{c}| \leq M$ and $\widetilde{c}(\widetilde{x}) = c(\Delta^{-1}(\widetilde{x}))$ for all $\widetilde{x} \in P$. Thus, $\widetilde{c}$ is an extension of $c \circ \Delta^{-1}$ onto $\widetilde{X}$ and $\widetilde{c} \in C_b(\widetilde{X})$.

b) Assume that $c \in C_L(X, \{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$. Then $|c(x)| \le \sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(x_\alpha)$. Set

$$\widetilde{c}(\widetilde{x}) := \begin{cases} -\sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(\widetilde{x}_\alpha), & \text{if } c(\mathrm{pr}(\widetilde{x})) < -\sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(\widetilde{x}_\alpha), \\ \sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(\widetilde{x}_\alpha), & \text{if } c(\mathrm{pr}(\widetilde{x})) > \sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(\widetilde{x}_\alpha), \\ c(\mathrm{pr}(\widetilde{x})), & \text{otherwise.} \end{cases}$$

The function $\widetilde{c}$ constructed in this way is continuous, $|\widetilde{c}(\widetilde{x})| \le \sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(\widetilde{x}_\alpha)$ and $\widetilde{c}(\widetilde{x}) = c(\Delta^{-1}(\widetilde{x}))$ for all $\widetilde{x} \in P$. Thus, $\widetilde{c}$ is an extension of $c \circ \Delta^{-1}$ onto $\widetilde{X}$ and $\widetilde{c} \in C_L(\widetilde{X}, \{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$. $\square$

Theorem 4.3 implies the following duality relation:

Proposition 4.11. *Under assumptions of the previous theorem*

$$\inf_{\pi\in\Pi(X,\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})} \int_X c\, d\pi = \sup_{f+\omega\le\widetilde{c}} \sum_{\alpha\in\mathcal{I}_{nk}} \int_{X_\alpha} f_\alpha(x_\alpha)\, d\mu_\alpha,$$

where $f(\widetilde{x}) = \sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(\widetilde{x}_\alpha)$, $f_\alpha \in C_L(X_\alpha, \mu_\alpha)$ *(or* $C_b(X_\alpha)$, *if* $c \in C_b(X)$*),* $\omega \in W$.

Assume that for the family of functions f_α there exists $\omega \in W$ such that $\sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(\widetilde{x}_\alpha) + \omega(\widetilde{x}) \le \widetilde{c}(\widetilde{x})$ for all $\widetilde{x} \in \widetilde{X}$. In particular, this equality holds for all $\widetilde{x} \in P$. Then for all $x \in X$

$$\sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(\Delta(x)_\alpha) + \omega(\Delta(x)) \le \widetilde{c}(\Delta(x)).$$

Moreover, $\widetilde{c}(\Delta(x)) = c(x)$, $\omega(\Delta(x)) = 0$, $\Delta(x)_\alpha = x_\alpha$, hence $\sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(x_\alpha) \le c(x)$ for all $x \in X$. One gets

$$\sup_{f+\omega\le\widetilde{c}} \sum_{\alpha\in\mathcal{I}_{nk}} \int_{X_\alpha} f_\alpha(x_\alpha)\, d\mu_\alpha \le \sup_{f\le c} \sum_{\alpha\in\mathcal{I}_{nk}} \int_{X_\alpha} f_\alpha(x_\alpha)\, d\mu_\alpha.$$

In addition, the following inequality holds:

$$\inf_{\pi\in\Pi(X,\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})} \int_X c\, d\pi \ge \sup_{f\le c} \sum_{\alpha\in\mathcal{I}_{nk}} \int_{X_\alpha} f_\alpha(x_\alpha)\, d\mu_\alpha.$$

Summarizing these results we get the following final version of our duality theorem that generalizes the duality theorem for compact spaces proven in the paper [19, Theorem 3.2]:

Theorem 4.12 *(Kantorovich duality). Assume we are given Polish spaces* $X_1, \dots, X_n$ *and a family of measures* $\mu_\alpha \in \mathcal{P}(X_\alpha)$, *where* $\alpha \in \mathcal{I}_{nk}$. *Let* $c \in C_L(X, \{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ *be a cost function on* X. *Then*

$$\inf_{\pi\in\Pi(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})} \int_X c\, d\pi = \sup_{f\le c} \sum_{\alpha\in\mathcal{I}_{nk}} \int_{X_\alpha} f_\alpha\, d\mu_\alpha,$$

where $f(x) = \sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(x_\alpha)$ *and* $f_\alpha \in C_L(X_\alpha, \mu_\alpha)$ *for all* $\alpha \in \mathcal{I}_{nk}$. *Moreover, if* $c \in C_b(X)$, *then the supremum can be taken on the set of bounded continuous functions* $f_\alpha \in C_b(X_\alpha)$. *If the set* $\Pi(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ *is non-empty, the infimum on the left-hand side is attained.*

Remark 4.13. Note that the statement of this theorem holds even when the set $\Pi(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ is empty. In this case, we define the infimum in the left-hand side to be $+\infty$. The supremum in the right-hand side is also equal to $+\infty$. For the case of compact spaces, it follows from Theorem 3.8, and for the case of non-compact X_i, this is a part of the Kantorovich duality Theorem 4.3 proven by Zaev.

5. Sufficient conditions for existence of a dual solution

5.1. Definition and properties of (n, k)-functions

Definition 5.1. Assume we are given Polish spaces $X_1, \dots, X_n$ and a positive integer $1 \le k < n$. A function $F\colon X \to [-\infty, +\infty)$ is called an (n, k)-function if there exists a collection of functions $\{f_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$, $f_\alpha\colon X_\alpha \to [-\infty, +\infty)$ satisfying

$$F(x) = \sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(x_\alpha)$$

for all $x \in X$. If $F(x) > -\infty$ for each x (and therefore $f_\alpha(x_\alpha) > -\infty$ for all $x_\alpha \in X_\alpha$), F is called a finite (n, k)-function.

This definition is given without any additional assumptions on the functions f_α and the function F. We prove that for every (n, k)-function F there exists a "regular" collection of functions $\{f_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$ such that $F(x) = \sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(x_\alpha)$ for all $x \in X$.

Let us introduce more notations. For $x_\alpha \in X_\alpha$, $x_\beta \in X_\beta$, such that $\alpha \cap \beta = \varnothing$, we denote by $x_\alpha x_\beta$ a point from the space $X_{\alpha\sqcup\beta}$, whose coordinates will be the union of the coordinates x_α and x_β. In addition, we write $\mathbf{n} = \{1, 2, \dots, n\}$.

Proposition 5.2. *Let F be a finite (n, k)-function defined on the space X. Fix $y \in X$. For each $\alpha \in \mathcal{I}_n$ we define a function $F_\alpha\colon x_\alpha \mapsto F(x_\alpha y_{\boldsymbol{n}\setminus\alpha})$ on the space X_α.*

Then there exists a sequence of real numbers $\{\lambda_i\}_{i=0}^k$ depending only on n and k such that $F(x) = \sum_{\alpha\in\mathcal{I}_{nk}} \widehat{f}_\alpha(x_\alpha)$ for each $x \in X$, where

$$\widehat{f}_\alpha(x_\alpha) = \sum_{\beta\subseteq\alpha} \lambda_{|\beta|} F_\beta(x_\beta),\ \ \alpha \in \mathcal{I}_{nk}.$$

This representation of F is regular in the following sense: if F is a measurable / continuous / bounded function, then for all $\alpha \in \mathcal{I}_{nk}$ the function $\widehat{f}_\alpha$ is measurable / continuous / bounded too.

Example 5.3. Let F be a finite $(n, 1)$-function. Fix $y = (y_1, y_2, \dots, y_n) \in X$. Let $\lambda_0 = \frac{1}{n} - 1$ and $\lambda_1 = 1$. Then

$$\widehat{f}_i(x_i) = F_i(x_i) - \frac{n-1}{n} F_\varnothing = F_i(y_1, \dots, y_{i-1}, x_i, y_{i+1}, \dots, y_n) - \frac{n-1}{n} F(y_1, \dots, y_n).$$

Since F is a finite $(n, 1)$-function, there exists a sequence of functions $f_i\colon X_i \to \mathbb{R}$ such that $F(x_1, \dots, x_n) = f_1(x_1) + \dots + f_n(x_n)$ for all $x \in X$. One can easily verify that

$$\widehat{f}_i(x_i) = f_i(x_i) - f_i(y_i) + \frac{1}{n}(f_1(y_1) + \dots + f_n(y_n)),$$

and therefore $F(x) = \sum_{i=1}^n \widehat{f}_i(x_i)$ for all $x \in X$.

Example 5.4. Let F be a finite $(3,2)$-function. Fix $(y_1,y_2,y_3) \in X$. Let $\lambda_0 = 1/3$, $\lambda_1 = -1/2$ and $\lambda_2 = 1$. Then by construction

$$\begin{aligned}
\widehat{f}_{12}(x_1,x_2) &= F(x_1,x_2,y_3) - \frac{1}{2}F(x_1,y_2,y_3) - \frac{1}{2}F(y_1,x_2,y_3) + \frac{1}{3}F(y_1,y_2,y_3),\\
\widehat{f}_{13}(x_1,x_3) &= F(x_1,y_2,x_3) - \frac{1}{2}F(x_1,y_2,y_3) - \frac{1}{2}F(y_1,y_2,x_3) + \frac{1}{3}F(y_1,y_2,y_3),\\
\widehat{f}_{23}(x_2,x_3) &= F(y_1,x_2,x_3) - \frac{1}{2}F(y_1,x_2,y_3) - \frac{1}{2}F(y_1,y_2,x_3) + \frac{1}{3}F(y_1,y_2,y_3).
\end{aligned}$$

Similarly to Example 5.3 we can verify that

$$F(x_1,x_2,x_3) = \widehat{f}_{12}(x_1,x_2) + \widehat{f}_{13}(x_1,x_3) + \widehat{f}_{23}(x_2,x_3)$$

for all $x \in X$.

Proof of Proposition 5.2. Consider a function $\widehat{F}\colon X \to \mathbb{R}$ defined as follows:

$$\widehat{F}(x) = \sum_{\alpha\in\mathcal{I}_{nk}} \widehat{f}_\alpha(x_\alpha).$$

Since by construction $\widehat{f}(x_\alpha) = \sum_{\beta\subseteq\alpha} \lambda_{|\beta|} F_\beta(x_\beta)$, one has

$$\widehat{F}(x) = \sum_{\beta\in\mathcal{I}_n} \sum_{\alpha\in\mathcal{I}_{nk}\colon\, \beta\subseteq\alpha} \lambda_{|\beta|} F_\beta(x_\beta).$$

For every $\beta \in \mathcal{I}_n$, let us find the amount A_β of numbers $\alpha \in \mathcal{I}_{nk}$ satisfying $\beta \subseteq \alpha$. If $|\beta| > k$, then there is no such α. Otherwise, it can be easily verified that $A_\beta = \binom{n-|\beta|}{k-|\beta|}$. Hence,

$$\widehat{F}(x) = \sum_{\beta\in\mathcal{I}_n\colon\, |\beta|\le k} \binom{n-|\beta|}{k-|\beta|} \lambda_{|\beta|} F_\beta(x_\beta) = \sum_{t=0}^{k} \lambda_t \binom{n-t}{k-t} \sum_{\beta\in\mathcal{I}_{nt}} F_\beta(x_\beta).$$

Since F is a finite (n,k)-function, there exists a collection of functions $\{f_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$, $f_\alpha\colon X_\alpha \to \mathbb{R}$, such that for all $x \in X$ we have

$$\sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(x_\alpha) = F(x).$$

For each $\beta \in \mathcal{I}_n$ the function $F_\beta(x_\beta)$ can be represented as follows:

$$F_\beta(x_\beta) = \sum_{\gamma,\delta\in\mathcal{I}_n} f_{\gamma\sqcup\delta}(x_\gamma y_\delta),$$

where the sum is taken for all pairs of disjoint sets of indices $\gamma, \delta \in \mathcal{I}_n$ satisfying $\gamma \subseteq \beta$, $\delta \subseteq \boldsymbol{n}\backslash\beta$ and $|\gamma| + |\delta| = k$. Hence, the function $\widehat{F}(x)$ can be represented as follows:

$$\widehat{F}(x) = \sum_{t=0}^{k} \lambda_t \binom{n-t}{k-t} \sum_{\beta\in\mathcal{I}_{nt}} F_\beta(x_\beta) = \sum_{\gamma,\delta\in\mathcal{I}_n} c_{\gamma,\delta} f_{\gamma\sqcup\delta}(x_\gamma y_\delta), \tag{6}$$

where the last sum is taken for all pairs of disjoint sets of indices γ and δ such that $|\gamma| + |\delta| = k$, and $c_{\gamma,\delta}$ is a linear combination of $\{\lambda_i\}_{i=0}^{k}$ with constant coefficients.

Let us find the coefficient $c_{\gamma,\delta}$. To this end, let us find for each $0 \le t \le k$ the amount of indices $\beta \in \mathcal{I}_{nt}$ satisfying $\gamma \subseteq \beta$ and $\delta \subseteq \boldsymbol{n}\backslash\beta$. If $t < |\gamma|$, then this quantity is trivially zero. Similarly, it is zero if $t > n - |\delta| = n - k + |\gamma|$. Otherwise, exactly $|\gamma|$ indices of β are fixed, and we need to choose $t - |\gamma|$ indices from $n - |\gamma| - |\delta| = n - k$ available items. Hence, the amount of such β is $\binom{n-k}{t-|\gamma|}$. Substituting this into equation (6) we get

$$c_{\gamma,\delta} = \sum_{t=|\gamma|}^{\min(k,n-k+|\gamma|)} \lambda_t \binom{n-t}{k-t}\binom{n-k}{t-|\gamma|}.$$

In particular, the coefficient $c_{\gamma,\delta}$ depends only on $|\gamma|$.

In order for the equality $F(x) = \widehat{F}(x)$ to hold, it is sufficient to require that the coefficients $c_{\gamma,\delta}$ satisfy the following equalities:

$$c_{\gamma,\delta} = \begin{cases} 1, & \text{if } |\gamma| = k, \\ 0, & \text{otherwise.} \end{cases}$$

We obtain the system of linear equations on λ

$$\begin{cases} \displaystyle\sum_{t=a}^{\min(k,n-k+a)} \lambda_t \binom{n-t}{k-t}\binom{n-k}{t-a} = 0 \quad \text{for } 0 \le a < k, \\ \displaystyle\sum_{t=k}^{\min(k,n-k+k)} \lambda_t \binom{n-t}{k-t}\binom{n-k}{t-k} = \lambda_k = 1. \end{cases}$$

The matrix of this linear system is upper-triangular and all diagonal elements are not equal to 0. Hence, this system admits a unique solution $\{\widehat{\lambda}_i\}_{i=0}^k$. Thus, if $\widehat{f}_\alpha(x_\alpha) = \sum_{\beta\subseteq\alpha} \widehat{\lambda}_{|\beta|} F_\beta(x_\beta)$, then $F(x) = \sum_{\alpha\in\mathcal{I}_{nk}} \widehat{f}_\alpha(x_\alpha)$ for all $x \in X$. $\square$

For $1 \le i \le n$, we fix a probability measure μ_i on the space X_i. For each $\alpha \in \mathcal{I}_n$ we denote by μ_α the probability measure $\prod_{i\in\alpha} \mu_i$ on the space X_α, and we denote by μ the probability measure $\prod_{1\le i\le n} \mu_i$ on the space X. If a finite (n, k)-function F is integrable (with respect to μ), one might expect that there exists a collection of integrable functions $\{f_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$ (with respect to μ_α) such that $F(x) = \sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(x_\alpha)$. We will prove it in Theorem 5.6 using a collection of integrable functions $\{\widehat{f}_\alpha\}$ constructed in Proposition 5.2.

To achieve this let us first verify the following lemma:

Lemma 5.5. *Let X_i, $1 \le i \le n$, be Polish spaces equipped with the Borel σ-algebras, and for every i let μ_i be a probability measure one X_i. Let $c\colon X \to \mathbb{R}$ be an integrable function on X. Fix a point $y \in X$, and for each $\alpha \in \mathcal{I}_n$ let us denote by c_α the function $x_\alpha \mapsto c(x_\alpha y_{\boldsymbol{n}\backslash\alpha})$ defined on X_α. For $\alpha = \varnothing$ the function $c_\varnothing$ is a constant function on the one-point space $X_\varnothing$ which is equal to $c(y)$, and $\|c_\varnothing\|_1$ is just the absolute value of $c(y)$.*

Then there exists a point $y \in X$ such that $\|c_\alpha\|_1 \le 2^{n+1} \|c\|_1$ for all $\alpha \in \mathcal{I}_n$.

Proof. For each $\alpha \in \mathcal{I}_n$ the spaces $X_{\boldsymbol{n}\backslash\alpha} \times X_\alpha$ and X are canonically isomorphic, and therefore the function c can be viewed as a function of two arguments $c(x_{\boldsymbol{n}\backslash\alpha}, y_\alpha)$, where $x_{\boldsymbol{n}\backslash\alpha} \in X_{\boldsymbol{n}\backslash\alpha}$ and $y_\alpha \in X_\alpha$.

By the Fubini-Tonelli theorem, the function $|c(\cdot, y_\alpha)|$ is integrable for μ_α-almost all y_α and

$$\|c\|_1 = \int\limits_{X_\alpha} \left(\int\limits_{X_{\boldsymbol{n}\backslash\alpha}} |c(x_{\boldsymbol{n}\backslash\alpha}, y_\alpha)|\, \mu_{\boldsymbol{n}\backslash\alpha}(dx_{\boldsymbol{n}\backslash\alpha}) \right) \mu_\alpha(dy_\alpha). \tag{7}$$

Consider the internal function from this expression:

$$C_\alpha(y_\alpha) = \int\limits_{X_{\boldsymbol{n}\setminus\alpha}} |c(x_{\boldsymbol{n}\setminus\alpha}, y_\alpha)|\, \mu_{\boldsymbol{n}\setminus\alpha}(dx_{\boldsymbol{n}\setminus\alpha}).$$

This function is non-negative. In addition, it follows from (7), $C_\alpha \in L^1(X_\alpha, \mu_\alpha)$ and $\|C_\alpha\|_1 = \|c\|_1$. Let

$$A_\alpha = \left\{ y_\alpha \in X_\alpha \colon C_\alpha(y_\alpha) > 2^{n+1} \|c\|_1 \right\}.$$

If $\|c\|_1 = 0$, then $C_\alpha(y_\alpha)$ is equal to 0 for μ_α-almost all points y_α, and therefore $\mu_\alpha(A_\alpha) = 0$. Otherwise, it follows from Markov's inequality that

$$\mu_\alpha(A_\alpha) \le \frac{1}{2^{n+1}\|c\|_1} \int\limits_{X_\alpha} C_\alpha(y_\alpha)\, \mu_\alpha(d_\alpha) = \frac{\|C_\alpha\|_1}{2^{n+1}\|c\|_1} = \frac{1}{2^{n+1}}.$$

In both cases we conclude that $\mu_\alpha(A_\alpha) \le 2^{-n-1}$.

If $y \in \mathrm{Pr}_\alpha^{-1}(X_\alpha \backslash A_\alpha)$, then

$$C_\alpha(y_\alpha) \le 2^{n+1} \|c\|,$$

and therefore the function $c_{\boldsymbol{n}\setminus\alpha} \colon x_{\boldsymbol{n}\setminus\alpha} \mapsto c(x_{\boldsymbol{n}\setminus\alpha} y_\alpha)$ is integrable with respect to $\mu_{\boldsymbol{n}\setminus\alpha}$ and

$$\left\| c_{\boldsymbol{n}\setminus\alpha} \right\|_1 = C_\alpha(y_\alpha) \le 2^{n+1} \|c\|_1.$$

Let us define

$$A = \bigcap_{\alpha \in \mathcal{I}_n} \mathrm{Pr}_\alpha^{-1}(X_\alpha \backslash A_\alpha).$$

Then if $y \in A$, for all $\alpha \in \mathcal{I}_n$ the function $c_{\boldsymbol{n}\setminus\alpha} \colon x_{\boldsymbol{n}\setminus\alpha} \mapsto c(x_{\boldsymbol{n}\setminus\alpha} y_\alpha)$ is integrable and $\left\| c_{\boldsymbol{n}\setminus\alpha} \right\|_1 \le 2^{n+1} \|c\|_1$.

We only need to verify that A is non-empty. We have

$$\mu\left(\mathrm{Pr}_\alpha^{-1}(X_\alpha \backslash A_\alpha)\right) = \mu_\alpha\left(X_\alpha \backslash A_\alpha\right) = 1 - \mu_\alpha(A_\alpha) \ge 1 - \frac{1}{2^{n+1}},$$

and therefore

$$\mu(A) \ge 1 - \frac{|\mathcal{I}_n|}{2^{n+1}} \ge 1 - \frac{2^n}{2^{n+1}} = \frac{1}{2}.$$

Thus, A is a set of positive measure, and therefore $A \ne \varnothing$. $\square$

Theorem 5.6. *For every $1 \le i \le n$, let X_i be a Polish space equipped with the Borel σ-algebra, and let μ_i be a probability measure on X_i. There exists a constant C depending only on n and k such that for any finite (n,k)-function $F \in L^1(X, \mu)$ there exists a collection of integrable functions $\{\widehat{f}_\alpha\}_{\alpha \in \mathcal{I}_{nk}}$, $\widehat{f}_\alpha \in L^1(X_\alpha, \mu_\alpha)$, such that*

$$F(x) = \sum_{\alpha \in \mathcal{I}_{nk}} \widehat{f}_\alpha(x_\alpha)$$

for all $x \in X$ and $\|\widehat{f}_\alpha\|_1 \le C \cdot \|F\|_1$ for all $\alpha \in \mathcal{I}_{nk}$.

Proof. Consider a finite (n,k)-function F defined on the space X. By Lemma 5.5 there exists a point $y \in X$ such that the function $F_\alpha \colon x_\alpha \mapsto F(x_\alpha y_{\boldsymbol{n}\setminus\alpha})$ is integrable and $\|F_\alpha\|_1 \le 2^{n+1}\|F\|_1$ for all $\alpha \in \mathcal{I}_n$.

By Proposition 5.2 there exists a sequence of real numbers $\{\lambda_i\}_{i=0}^k$ such that $F(x) = \sum_{\alpha\in\mathcal{I}_{nk}} \widehat{f}_\alpha(x_\alpha)$ for all $x \in X$, where

$$\widehat{f}_\alpha(x_\alpha) = \sum_{\beta\subseteq\alpha} \lambda_{|\beta|} F_\beta(x_\beta),\ \ \alpha \in \mathcal{I}_{nk}.$$

Since $F_\beta \in L^1(X_\beta,\mu_\beta)$ for all $\beta \in \mathcal{I}_n$, we conclude that $\widehat{f}_\alpha \in L^1(X_\alpha,\mu_\alpha)$. In addition,

$$\left\|\widehat{f}_\alpha\right\|_1 \le \sum_{\beta\subseteq\alpha} \left|\lambda_{|\beta|}\right| \cdot \|F_\beta\|_1 \le 2^{n+1}\|F\|_1 \sum_{\beta\subseteq\alpha} \left|\lambda_{|\beta|}\right| = 2^{n+1}\|F\|_1 \sum_{t=0}^{k} \binom{k}{t} |\lambda_t|.$$

Thus, we conclude that $\left\|\widehat{f}_\alpha\right\|_1 \le C\cdot\|F\|_1$, where

$$C = 2^{n+1}\sum_{t=0}^{k}\binom{k}{t}|\lambda_t|,$$

and this constant depends only on n and k. $\square$

Example 5.7. Let us find a constant C explicitly for the case of the $(3,2)$-problem. Consider a finite integrable $(3,2)$-function F. There exists a point $y \in X = X_1\times X_2\times X_3$ such that $\|F_\alpha\|_1 \le 16\|F\|_1$ for all $\alpha\in\mathcal{I}_3$. By Example 5.4 the functions

$$\begin{aligned}
\widehat{f}_{12}(x_1,x_2) &= F(x_1,x_2,y_3) - \frac{1}{2}F(x_1,y_2,y_3) - \frac{1}{2}F(y_1,x_2,y_3) + \frac{1}{3}F(y_1,y_2,y_3),\\
\widehat{f}_{13}(x_1,x_3) &= F(x_1,y_2,x_3) - \frac{1}{2}F(x_1,y_2,y_3) - \frac{1}{2}F(y_1,y_2,x_3) + \frac{1}{3}F(y_1,y_2,y_3),\\
\widehat{f}_{23}(x_2,x_3) &= F(y_1,x_2,x_3) - \frac{1}{2}F(y_1,x_2,y_3) - \frac{1}{2}F(y_1,y_2,x_3) + \frac{1}{3}F(y_1,y_2,y_3),
\end{aligned}$$

satisfy the equation $F(x_1,x_2,x_3) = \widehat{f}_{12}(x_1,x_2) + \widehat{f}_{13}(x_1,x_3) + \widehat{f}_{23}(x_2,x_3)$ for all $(x_1,x_2,x_3) \in X$. All functions $\{\widehat{f}_{ij}\}$ are integrable with respect to $\mu_i\otimes\mu_j$. In addition,

$$\begin{aligned}
\left\|\widehat{f}_{12}\right\|_1 &\le \|F(\cdot,\cdot,y_3)\|_1 + \frac{1}{2}\|F(\cdot,y_2,y_3)\|_1 + \frac{1}{2}\|F(y_1,\cdot,y_3)\|_1 + \frac{1}{3}|F(y_1,y_2,y_3)|\\
&\le 16\left(1+\frac{1}{2}+\frac{1}{2}+\frac{1}{3}\right)\|F\|_1 < 38\|F\|_1.
\end{aligned}$$

Similarly, $\left\|\widehat{f}_{13}\right\|_1 < 38\|F\|_1$ and $\left\|\widehat{f}_{23}\right\|_1 < 38\|F\|_1$, and therefore we can put $C = 38$. This constant estimate is crude, but we do not need to know the optimal value.

We want to generalize this property to a wider class of measures that are uniformly equivalent to the product of their projections to one-dimensional spaces.

Definition 5.8. We call the probability measure μ on the space X *reducible* if for $1\le i\le n$ there exists a probability measure ν_i on spaces X_i such that μ is uniformly equivalent to $\prod_{1\le i\le n}\nu_i$.

We call the consistent set of probability measures $\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$ *reducible* if there exists a uniting reducible measure $\mu\in\Pi(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$.

If the probability measures μ and ν on the space X are uniformly equivalent, then their projections are also uniformly equivalent: $\mathrm{Pr}_\alpha \mu$ is uniformly equivalent to $\mathrm{Pr}_\alpha \nu$ for all $\alpha \in \mathcal{I}_n$. In particular, if the set of measures $\{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}}$ is reducible, then $\mu_i = \mathrm{Pr}_i(\mu)$ is uniformly equivalent to ν_i. Then the measure $\prod_{1 \le i \le n} \mu_i$ is uniformly equivalent to the measure $\prod_{1 \le i \le n} \nu_i$. Hence, the following is true:

Proposition 5.9. *A collection of probability measures $\{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}}$ is reducible if and only if there exists a uniting measure $\mu \in \Pi(\{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}})$, which is uniformly equivalent to $\prod_{1 \le i \le n} \mu_i$.*

If the set of measures μ_α is reducible, then for all $\beta \in \mathcal{I}_{nt}$, $t \le k$, the measure μ_β is uniformly equivalent to $\prod_{i \in \beta} \mu_i$. It is easy to see that this condition is not sufficient.

Example 5.10. Let X_1, X_2 and X_3 be discrete spaces, each consisting of two elements $\{0, 1\}$. Define a probability measure μ_{ij} on the space $X_i \times X_j$ as follows:

$$\mu_{ij}(x_i, x_j) = \begin{cases} \frac{1}{3}, & \text{if } x_i \ne x_j, \\ \frac{1}{6}, & \text{otherwise.} \end{cases}$$

The triple of measures $\{\mu_{ij}\}_{\{i,j\} \in \mathcal{I}_{3,2}}$ is consistent. In addition, every measure μ_{ij}, $\{i, j\} \in \mathcal{I}_{3,2}$, is uniformly equivalent to $\mu_i \otimes \mu_j$. The set $\Pi(\{\mu_{ij}\}_{\{i,j\} \in \mathcal{I}_{3,2}})$ is non-empty: consider the following measure μ on the space $X_1 \times X_2 \times X_3$: $\mu(x_1, x_2, x_3) = 0$ if $x_1 = x_2 = x_3$, otherwise $\mu(x_1, x_2, x_3) = 1/6$. It is easy to check that $\mu \in \Pi(\{\mu_{ij}\}_{\{i,j\} \in \mathcal{I}_{3,2}})$.

Let $\nu \in \Pi(\{\mu_{ij}\}_{\{i,j\} \in \mathcal{I}_{3,2}})$. Then the following equations hold:

$$\nu(0,0,0) + \nu(0,0,1) = \mu_{12}(0,0) = \frac{1}{6},$$
$$\nu(0,0,1) + \nu(0,1,1) = \mu_{13}(0,1) = \frac{1}{3},$$
$$\nu(0,1,1) + \nu(1,1,1) = \mu_{23}(1,1) = \frac{1}{6}.$$

From these equations we get $\nu(0,0,0) + \nu(1,1,1) = 0$. From the non-negativity of the measure we get $\nu(0,0,0) = \nu(1,1,1) = 0$, and then we easily verify that $\nu(x_1, x_2, x_3) = 1/6$ for the remaining points. Thus $\Pi(\{\mu_{ij}\}_{\{i,j\} \in \mathcal{I}_{3,2}})$ consists of a single measure that is not uniformly equivalent to $\mu_1 \otimes \mu_2 \otimes \mu_3$.

The following theorem generalizes Theorem 5.6 to reducible collections of measures.

Theorem 5.11. *For $1 \le i \le n$, let X_i be a Polish space equipped with the Borel σ-algebra, and let μ be a reducible probability measure on X. Denote $\mu_\alpha = \mathrm{Pr}_\alpha(\mu)$. Then there exists a constant C_μ depending on the measure μ and parameters (n, k) such that for any finite (n, k)-function $F \in L^1(X, \mu)$ there exists a collection of integrable functions $\{\widehat{f}_\alpha\}_{\alpha \in \mathcal{I}_{nk}}$, $\widehat{f}_\alpha \in L^1(X_\alpha, \mu_\alpha)$, such that*

$$F(x) = \sum_{\alpha \in \mathcal{I}_{nk}} \widehat{f}_\alpha(x_\alpha)$$

for all $x \in X$ and

$$\left\| \widehat{f}_\alpha \right\|_{L^1(\mu_\alpha)} \le C_\mu \cdot \|F\|_{L^1(\mu)}$$

for all $\alpha \in \mathcal{I}_{nk}$.

Proof. Since μ is reducible, there exist probability measures $\nu_i \in \mathcal{P}(X_i)$ and positive reals m and M such that $m \cdot \nu \leq \mu \leq M \cdot \nu$, where $\nu = \prod_{1 \leq i \leq n} \nu_i$.

Consider a finite (n,k)-function $F \in L^1(X,\mu)$. Since $\mu \geq m \cdot \nu$, the function F is integrable with respect to ν and

$$\|F\|_{L^1(\nu)} \leq \frac{1}{m} \|F\|_{L^1(\mu)} .$$

Denote $\nu_\alpha = \prod_{i \in \alpha} \nu_i$. It follows from Theorem 5.6 that there exists a collection of integrable functions $\{\widehat{f}_\alpha\}_{\alpha \in \mathcal{I}_{nk}}$, $\widehat{f}_\alpha \in L^1(X_\alpha, \nu_\alpha)$ such that

$$F(x) = \sum_{\alpha \in \mathcal{I}_{nk}} \widehat{f}_\alpha(x_\alpha)$$

for all $x \in X$ and

$$\left\|\widehat{f}_\alpha\right\|_{L^1(\nu_\alpha)} \leq C \cdot \|F\|_{L^1(\nu)} \leq \frac{C}{m} \|F\|_{L^1(\mu)}$$

for all $\alpha \in \mathcal{I}_{nk}$, where C is a constant depending only on n and k

Since $M \cdot \nu \geq \mu$, we have $M \cdot \nu_\alpha \geq \mu_\alpha$ for all $\alpha \in \mathcal{I}_{nk}$. Hence, the function $\widehat{f}_\alpha$ is integrable with respect to μ_α and

$$\left\|\widehat{f}_\alpha\right\|_{L^1(\mu_\alpha)} \leq M \left\|\widehat{f}_\alpha\right\|_{L^1(\nu_\alpha)} \leq \frac{M}{m} C \|F\|_{L^1(\mu)}$$

for all $\alpha \in \mathcal{I}_{nk}$. Thus, we can put $C_\mu = \frac{M}{m} C$. $\square$

5.2. Existence of a dual solution for reducible collections of measures

First, we generalize the notion of the proper thickness of the set introduced in [35].

Definition 5.12. Let $X_1, \dots, X_n$ be Polish spaces, and for each $\alpha \in \mathcal{I}_{nk}$ let μ_α be a probability measure on the space X_α. For a measurable set $A \subset X$ define its proper (n,k)-thickness as

$$\operatorname{sth}(A) = \inf \left\{ \sum_{\alpha \in \mathcal{I}_{nk}} \mu_\alpha(Y_\alpha) \colon Y_\alpha \subseteq X_\alpha, A \subseteq \bigcup_{\alpha \in \mathcal{I}_{nk}} \operatorname{Pr}_\alpha^{-1}(Y_\alpha) \right\}. \tag{8}$$

We are going to use this notion in the particular case of sets with zero proper thickness.

Proposition 5.13. *If* $\operatorname{sth}(A) = 0$*, then the infimum in* (8) *is attained: there exist measurable subsets* $Y_\alpha \subseteq X_\alpha$*,* $\alpha \in \mathcal{I}_{nk}$*, such that* $\mu_\alpha(Y_\alpha) = 0$ *and* $A \subseteq \bigcup_{\alpha \in \mathcal{I}_{nk}} \operatorname{Pr}_\alpha^{-1}(Y_\alpha)$.

Proof. The proof follows the proof of [35, Lemma 2.5.4]. If for a collection of measurable subsets $\{Y_\alpha\}_{\alpha \in \mathcal{I}_{nk}}$ we have $A \subseteq \bigcup_{\alpha \in \mathcal{I}_{nk}} \operatorname{Pr}_\alpha^{-1}(Y_\alpha)$, then $f_\alpha = \mathbb{1}[Y_\alpha]$ satisfy the inequality

$$\sum_{\alpha \in \mathcal{I}_{nk}} f_\alpha(x_\alpha) \geq \mathbb{1}[A](x)$$

for all $x \in X$, where $\mathbb{1}[A]$ is the characteristic function of the set A. Moreover, it is clear that

$$\sum_{\alpha\in\mathcal{I}_{nk}} \int\limits_{X_\alpha} f_\alpha(x_\alpha)\,\mu_\alpha(dx_\alpha) = \sum_{\alpha\in\mathcal{I}_{nk}} \mu_\alpha(Y_\alpha).$$

Since $\mathrm{sth}(A) = 0$, we can consider a minimizing sequence of collections of functions $\{f_\alpha^{(t)}\}_\alpha$, $f_\alpha^{(t)} : X_\alpha \to [0,1]$, such that

$$\sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha^{(t)}(x_\alpha) \geq \mathbb{1}[A](x)$$

for all $x \in X$ and

$$\sum_{\alpha\in\mathcal{I}_{nk}} \int\limits_{X_\alpha} f_\alpha^{(t)}(x_\alpha)\,\mu_\alpha(dx_\alpha) \xrightarrow[t\to\infty]{} 0.$$

Since $f_\alpha^{(t)}$ is non-negative for all $\alpha \in \mathcal{I}_{nk}$ and for all t, we conclude that

$$\int\limits_{X_\alpha} f_\alpha^{(t)}(x_\alpha)\,\mu_\alpha(dx_\alpha) \xrightarrow[t\to\infty]{} 0 \text{ for all } \alpha \in \mathcal{I}_{nk}.$$

Let us recall the formulation of the Komlós theorem.

Theorem 5.14 *([5, Theorem 4.7.24]). Let μ be a finite nonnegative measure on a space X, let $\{f_n\} \subset L^1(\mu)$, and let*

$$\sup_n \|f_n\|_{L^1(\mu)} < \infty.$$

Then, one can find a subsequence $\{g_n\} \subseteq \{f_n\}$ and a function $g \in L^1(\mu)$ such that, for every sequence $\{h_n\} \subseteq \{g_n\}$, the arithmetic means $(h_1 + \cdots + h_n)/n$ converge almost everywhere to g.

Using this theorem and passing, if necessary, to subsequences, we may assume that the sequence

$$g_\alpha^{(t)} = \frac{1}{t}\left(f_\alpha^{(1)} + \cdots + f_\alpha^{(t)}\right)$$

converges to some integrable function g_α μ_α-almost everywhere in X_α for all $\alpha \in \mathcal{I}_{nk}$. Thus, we can suppose that

$$g_\alpha(x_\alpha) = \limsup_{t\to\infty} g_\alpha^{(t)}(x_\alpha) \text{ for all } x_\alpha \in X_\alpha.$$

By construction we obtain $0 \leq g_\alpha(x_\alpha) \leq 1$ for all $x_\alpha \in X_\alpha$. Also, since $\sum_{\alpha\in\mathcal{I}_{nk}} g_\alpha^{(t)}(x_\alpha) \geq \mathbb{1}[A](x)$ for all $x \in X$ and for all t, we conclude that

$$\sum_{\alpha\in\mathcal{I}_{nk}} g_\alpha(x_\alpha) \geq \mathbb{1}[A](x) \text{ for all } x \in X. \tag{9}$$

In addition, since $|g_\alpha^{(t)}(x_\alpha)| \leq 1$ it follows from the Lebesgue's dominated convergence theorem that

$$\int\limits_{X_\alpha} g_\alpha(x_\alpha)\,\mu_\alpha(dx_\alpha) = \lim_{t\to\infty} \int\limits_{X_\alpha} g_\alpha^{(t)}(x_\alpha)\,\mu_\alpha(dx_\alpha) = \lim_{t\to\infty} \int\limits_{X_\alpha} f_\alpha^{(t)}(x_\alpha)\,\mu_\alpha(dx_\alpha) = 0.$$

Thus, since the function g_α is non-negative, we conclude that $g_\alpha(x_\alpha) = 0$ for μ_α-almost all $x_\alpha \in X_\alpha$.

Consider the collection of sets $\{Y_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$:

$$Y_\alpha = \{x_\alpha \in X_\alpha \colon g_\alpha(x_\alpha) > 0\}.$$

Since g_α is equal to 0 almost everywhere on X_α, we have $\mu_\alpha(Y_\alpha) = 0$. In addition, if $x \in A$, then it follows from inequality (9) that $\sum_{\alpha\in\mathcal{I}_{nk}} g_\alpha(x_\alpha) \geq 1$, and therefore there exists at least one $\alpha \in \mathcal{I}_{nk}$ such that $g_\alpha(x_\alpha) > 0$ or equivalently $x_\alpha \in Y_\alpha$. Thus, $A \subseteq \bigcup_{\alpha\in\mathcal{I}_{nk}} \mathrm{Pr}_\alpha^{-1}(Y_\alpha)$. $\square$

Definition 5.15. We say that a measurable set $A \subset X$ is a zero (n,k)-thickness set if $\mathrm{sth}(A) = 0$, or equivalently if there exist a collection of measurable subsets $Y_\alpha \subset X_\alpha$, $\alpha \in \mathcal{I}_{nk}$ such that $\mu_\alpha(Y_\alpha) = 0$ for all α and $A \subseteq \cup_{\alpha\in\mathcal{I}_{nk}} \mathrm{Pr}_\alpha^{-1}(Y_\alpha)$.

In addition to the standard dual multistochastic problem, we consider a more convenient relaxed dual problem.

Definition 5.16 *(Relaxed dual problem).* Let c be a measurable cost function on the space X. Denote by

$$\Psi_c(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$$

the set of collections of integrable functions $\{f_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$, $f_\alpha \colon X_\alpha \to \mathbb{R}$ such that inequality

$$\sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(x_\alpha) \leq c(x)$$

holds at all points $x \in X$ except a zero (n,k)-thickness set. Then, in the relaxed dual problem we are looking for

$$J = \sup\left\{ \sum_{\alpha\in\mathcal{I}_{nk}} \int_{X_\alpha} f_\alpha(x_\alpha)\,\mu_\alpha(dx_\alpha) \colon \{f_\alpha\} \in \Psi_c(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}}) \right\}$$

First, let us verify that Kantorovich duality also holds for the relaxed dual problem.

Theorem 5.17 *(Kantorovich duality for the relaxed dual problem). Assume we are given Polish spaces $X_1, \ldots, X_n$ and a family of measures $\mu_\alpha \in \mathcal{P}(X_\alpha)$, where $\alpha \in \mathcal{I}_{nk}$. Let $c \in C_L(X, \{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ be a cost function on X. Then*

$$\inf_{\pi\in\Pi(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})} \int_X c\,d\pi = \sup_{\{f_\alpha\}\in\Psi_c(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})} \sum_{\alpha\in\mathcal{I}_{nk}} \int_{X_\alpha} f_\alpha\,d\mu_\alpha.$$

If the set $\Pi(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ is non-empty, the infimum on the left-hand side is attained.

Proof. If $\{f_\alpha\} \in \Psi_c(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$, then there exists a collection of measurable subsets $Y_\alpha \subset X_\alpha$ such that $\mu_\alpha(Y_\alpha) = 0$ and

$$\sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(x_\alpha) \leq c(x) \text{ for all } x \notin \bigcup_{\alpha\in\mathcal{I}_{nk}} \mathrm{Pr}_\alpha^{-1}(Y_\alpha).$$

Consider the collection of functions $\{\widehat{f_\alpha}\}$ defined as follows: $\widehat{f_\alpha}(x_\alpha) = f_\alpha(x_\alpha)$ if $x_\alpha \notin Y_\alpha$ and $\widehat{f_\alpha}(x_\alpha) = -\infty$ otherwise. For all $\alpha \in \mathcal{I}_{nk}$ the function $\widehat{f_\alpha}$ coincides with f_α almost everywhere with respect to μ_α, and therefore

$$\sum_{\alpha\in\mathcal{I}_{nk}}\int \widehat{f}_\alpha(x_\alpha)\,\mu_\alpha(dx_\alpha)=\sum_{\alpha\in\mathcal{I}_{nk}}\int f_\alpha(x_\alpha)\,\mu_\alpha(dx_\alpha)$$

In addition, the inequality $\sum_{\alpha\in\mathcal{I}_{nk}}\widehat{f}_\alpha(x_\alpha)\le c(x)$ holds for all $x\in X$. Thus, for any uniting measure $\pi\in\Pi(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ we have

$$\sum_{\alpha\in\mathcal{I}_{nk}}\int f_\alpha(x_\alpha)\,\mu_\alpha(dx_\alpha)=\sum_{\alpha\in\mathcal{I}_{nk}}\int \widehat{f}_\alpha(x_\alpha)\,\mu_\alpha(dx_\alpha)=\int\limits_X\left(\sum_{\alpha\in\mathcal{I}_{nk}}\widehat{f}_\alpha(x_\alpha)\right)d\pi\le\int\limits_X c(x)\,d\pi.$$

In particular, we conclude that

$$\inf_{\pi\in\Pi(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})}\int\limits_X c\,d\pi\ge\sup_{\{f_\alpha\}\in\Psi_c(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})}\sum_{\alpha\in\mathcal{I}_{nk}}\int\limits_{X_\alpha}f_\alpha\,d\mu_\alpha, \tag{10}$$

which is usually called “the weak duality”. Since every collection of functions $\{f_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$ satisfying the conditions $f_\alpha\in C_L(X_\alpha,\mu_\alpha)$ and $\sum_{\alpha\in\mathcal{I}_{nk}}f_\alpha(x_\alpha)\le c(x)$ for all $x\in X$ belongs to the space $\Psi_c(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$, we conclude that

$$\sup_{\{f_\alpha\}\in\Psi_c(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})}\sum_{\alpha\in\mathcal{I}_{nk}}\int\limits_{X_\alpha}f_\alpha\,d\mu_\alpha\ge\sup_{\substack{f_\alpha\in C_L(X_\alpha,\mu_\alpha),\\ \sum f_\alpha\le c}}\sum_{\alpha\in\mathcal{I}_{nk}}\int\limits_{X_\alpha}f_\alpha\,d\mu_\alpha.$$

By Theorem 4.12,

$$\inf_{\pi\in\Pi(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})}\int\limits_X c\,d\pi=\sup_{\substack{f_\alpha\in C_L(X_\alpha,\mu_\alpha),\\ \sum f_\alpha\le c}}\sum_{\alpha\in\mathcal{I}_{nk}}\int\limits_{X_\alpha}f_\alpha\,d\mu_\alpha;$$

thus, the equality is achieved in expression (10). □

In [25] the following theorem was proved, establishing the existence of a dual solution in the multi-marginal case.

Theorem 5.18 *(Kellerer). For every $1\le i\le n$, let X_i be a Polish space equipped with a Borel probability measure μ_i. Let $c\colon X_1\times\cdots\times X_n\to[-\infty,+\infty]$ be a measurable cost function on the space $X_1\times\cdots\times X_n$. Suppose that there exists a sequence of integrable functions $\{c_i\}_{i=1}^n$, $c_i: X_i\to(-\infty,+\infty]$ such that inequality*

$$|c(x_1,\dots,x_n)|\le\sum_{i=1}^n c_i(x_i)$$

holds for all $(x_1,\dots,x_n)\in X$.

Then the supremum in the relaxed dual Monge-Kantorovich problem

$$\sup\left\{\sum_{i=1}^n\int\limits_{X_i}\varphi_i(x_i)\,\mu_i(dx_i)\colon\{\varphi_i\}_{i=1}^n\in\Psi_c(\{\mu_i\}_{i=1}^n)\right\}$$

is finite and attained.

We prove the multistochastic generalization of this theorem for the case of reducible collection of projections.

Theorem 5.19 *(Supremum reachability in relaxed dual problem). For every $1 \le i \le n$, let X_i be a Polish space, let $\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$, $\mu_\alpha \in \mathcal{P}(X_\alpha)$ be a reducible collection of probability measures, and let $c\colon X \to [-\infty, +\infty]$ be a measurable cost function on the space X. Suppose that there exists a collection of integrable functions $\{c_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$, $c_\alpha\colon X_\alpha \to (-\infty, +\infty]$ such that the inequality*

$$|c(x)| \le \sum_{\alpha\in\mathcal{I}_{nk}} c_\alpha(x_\alpha)$$

holds for all $x \in X$.

Then the supremum in the relaxed dual multistochastic Monge-Kantorovich problem

$$J = \sup\left\{ \sum_{\alpha\in\mathcal{I}_{nk}} \int_{X_\alpha} f_\alpha(x_\alpha)\,\mu_\alpha(dx_\alpha)\colon \{f_\alpha\}_{\alpha\in\mathcal{I}_{nk}} \in \Psi_c(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}}) \right\} \tag{11}$$

is finite and attained.

Proof. Replacing c_α with $|c_\alpha|$ we may assume that the function c_α is non-negative for all $\alpha \in \mathcal{I}_{nk}$. Let $c^*_\alpha\colon X_\alpha \to [0, +\infty)$ be an arbitrary finite integrable function such that $c^*_\alpha(x_\alpha) = c_\alpha(x_\alpha)$ for μ_α-almost all $x_\alpha \in X_\alpha$. Consider a function c^* on the space X:

$$c^*(x) = \begin{cases} c(x), & \text{if } c^*_\alpha(x_\alpha) = c_\alpha(x_\alpha) \text{ for all } x_\alpha \in X_\alpha, \\ 0, & \text{otherwise.} \end{cases}$$

It trivially follows from the construction that $c^*(x) = c(x)$ for all $x \in X$ except a zero (n, k)-thickness set. Hence,

$$\Psi_c(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}}) = \Psi_{c^*}(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}}).$$

In addition, $|c^*(x)| \le \sum_{\alpha\in\mathcal{I}_{nk}} c^*_\alpha(x_\alpha)$ for all $x \in X$. In particular, since $c^*_\alpha(x_\alpha) < +\infty$ for all $x_\alpha \in X_\alpha$ and for all $\alpha \in \mathcal{I}_{nk}$, we conclude that $|c^*(x)| < +\infty$ for all $x \in X$. Thus, replacing c with c^* and replacing c_α with c^*_α for all $\alpha \in \mathcal{I}_{nk}$, we may assume that $|c(x)| < +\infty$ for all $x \in X$ and $0 \le c_\alpha(x_\alpha) < +\infty$ for all $x_\alpha \in X_\alpha$ and for all $\alpha \in \mathcal{I}_{nk}$.

Denote

$$\widehat{J} = \sum_{\alpha\in\mathcal{I}_{nk}} \int_{X_\alpha} c_\alpha(x_\alpha)\,\mu_\alpha(dx_\alpha).$$

The function $c_\alpha\colon X_\alpha \to [0, +\infty)$ is finite and integrable with respect to μ_α for all $\alpha \in \mathcal{I}_{nk}$; in addition,

$$\sum_{\alpha\in\mathcal{I}_{nk}} (-c_\alpha(x_\alpha)) \le c(x) \text{ for all } x \in X.$$

Thus, $\{-c_\alpha\}_{\alpha\in\mathcal{I}_{nk}} \in \Psi_c(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$, and therefore the set $\Psi_c(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ is non-empty and

$$J \ge \sum_{\alpha\in\mathcal{I}_{nk}} \int_{X_\alpha} (-c_\alpha(x_\alpha))\,\mu_\alpha(dx_\alpha) = -\widehat{J}.$$

Since the collection of measures $\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$ is reducible, there exists a reducible measure $\mu \in \Pi(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$. Since $c_\alpha \in L^1(X_\alpha, \mu_\alpha)$, the extension of c_α to the space X is integrable with respect to μ. Thus, since $|c(x)| \le \sum_{\alpha\in\mathcal{I}_{nk}} c_\alpha(x_\alpha) \in L^1(X, \mu)$, we conclude that $c \in L^1(X, \mu)$.

Let $\{f_\alpha\}_{\alpha\in\mathcal{I}_{nk}} \in \Psi_c(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$. Since $f_\alpha \in L^1(X_\alpha, \mu_\alpha)$, the extension of f_α to the space X is integrable with respect to μ. Hence,

$$\sum_{\alpha\in\mathcal{I}_{nk}} \int_{X_\alpha} f_\alpha(x_\alpha)\,\mu_\alpha(dx_\alpha) = \int_X \sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(x_\alpha)\,\mu(dx)$$

We have $\sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(x_\alpha) \le c(x)$ at all points except a zero (n,k)-thickness set. Since μ is a uniting measure, every set of zero (n,k)-thickness has zero measure with respect to μ. Hence, $\sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(x_\alpha) \le c(x)$ for μ-almost all $x \in X$, and therefore

$$\int_X \sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(x_\alpha)\,\mu(dx) \le \int_X c(x)\,\mu(dx) \le \sum_{\alpha\in\mathcal{I}_{nk}} \int_{X_\alpha} c_\alpha(x_\alpha)\,\mu_\alpha(dx_\alpha) = \widehat{J}.$$

Thus, we conclude that

$$\sum_{\alpha\in\mathcal{I}_{nk}} \int_{X_\alpha} f_\alpha(x_\alpha)\,\mu_\alpha(dx_\alpha) \le \widehat{J} \text{ for all } \{f_\alpha\} \in \Psi_c(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}}),$$

and therefore $J \le \widehat{J}$. In particular, the supremum in (11) is finite.

Consider the maximizing sequence of collections of functions $\{f_\alpha^{(t)}\}_{\alpha\in\mathcal{I}_{nk}} \in \Psi_c(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ such that

$$\sum_{\alpha\in\mathcal{I}_{nk}} \int_{X_\alpha} f_\alpha^{(t)}(x_\alpha)\,\mu_\alpha(dx_\alpha) \xrightarrow[n\to\infty]{} J.$$

We may assume that

$$\sum_{\alpha\in\mathcal{I}_{nk}} \int_{X_\alpha} f_\alpha^{(t)}(x_\alpha)\,\mu_\alpha(dx_\alpha) \ge -\widehat{J} \text{ for all } t. \tag{12}$$

For each t consider the finite (n,k)-function $F^{(t)}(x) = \sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha^{(t)}(x_\alpha)$. Let us bound the norm of the function $F^{(t)}$ from above. Since $F^{(t)}(x) \le c(x)$ for all points except a zero (n,k)-thickness set, and $c(x) \le \sum_{\alpha\in\mathcal{I}_{nk}} c_\alpha(x_\alpha)$ for all $x \in X$, we conclude that $F^{(t)}(x) \le \sum_{\alpha\in\mathcal{I}_{nk}} c_\alpha(x_\alpha)$ for μ-almost all $x \in X$. Finally, since $\sum_{\alpha\in\mathcal{I}_{nk}} c_\alpha(x_\alpha) \ge 0$, we have

$$F^{(t)}(x) + |F^{(t)}(x)| = \max(0, 2F^{(t)}(x)) \le 2\sum_{\alpha\in\mathcal{I}_{nk}} c_\alpha(x_\alpha)$$

for μ-almost all $x \in X$. Combining this with inequality (12) we get

$$\left\|F^{(t)}\right\|_{L^1(\mu)} = \int_X |F^{(t)}(x)|\,\mu(dx) \le 2\sum_{\alpha\in\mathcal{I}_{nk}} \int_{X_\alpha} c_\alpha(x_\alpha)\mu_\alpha(dx_\alpha) - \int_X F^{(t)}(x)\,\mu(dx) \le 3\widehat{J}.$$

Since μ is reducible, for each t by Theorem 5.11 there exists a collection of finite integrable functions $\{\widehat{f}_\alpha^{(t)}\}_{\alpha\in\mathcal{I}_{nk}}$ such that the equation

$$F^{(t)}(x) = \sum_{\alpha\in\mathcal{I}_{nk}} \widehat{f}_\alpha^{(t)}(x_\alpha)$$

holds for all $x \in X$ and

$$\left\|\widehat{f}_\alpha^{(t)}\right\|_{L^1(\mu_\alpha)} \leq C_\mu \left\|F^{(t)}\right\|_{L^1(\mu)} \leq 3C_\mu \widehat{J} = C$$

for all $\alpha \in \mathcal{I}_{nk}$. In particular, $\{\widehat{f}_\alpha^{(t)}\} \in \Psi_c(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ for all t, and this sequence of collections is also maximizing. Thus, replacing $\{f_\alpha^{(t)}\}$ with $\{\widehat{f}_\alpha^{(t)}\}$, we may assume that the inequality

$$\left\|f_\alpha^{(t)}\right\|_{L^1(\mu_\alpha)} \leq C$$

holds for all $\alpha \in \mathcal{I}_{nk}$ and for all t.

In particular,

$$\sup_t \left\|f_\alpha^{(t)}\right\|_{L^1(\mu_\alpha)} < +\infty$$

for all $\alpha \in \mathcal{I}_{nk}$. Hence, using the Komloś theorem and passing, if necessary, to subsequences, we may assume that the sequence of functions

$$g_\alpha^{(t)}(x_\alpha) = \frac{1}{t}\left(f_\alpha^{(1)} + \cdots + f_\alpha^{(t)}\right), \quad t \in \mathbb{N},$$

converges to some function $g_\alpha \in L^1(X_\alpha, \mu_\alpha)$ μ_α-almost everywhere in X_α for all $\alpha \in \mathcal{I}_{nk}$.

For each t consider the finite (n,k)-function

$$G^{(t)}(x) = \sum_{\alpha\in\mathcal{I}_{nk}} g_\alpha^{(t)}(x_\alpha) = \frac{1}{t}\left(F^{(1)}(x) + \cdots + F^{(t)}(x)\right).$$

We have $G^{(t)}(x) \leq c(x)$ for all $x \in X$ except a zero (n,k)-thickness set, and therefore $\{g_\alpha^{(t)}\}_{\alpha\in\mathcal{I}_{nk}} \in \Psi_c(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ for all t. In addition, it follows from the properties of the Ceśaro mean that the sequence of collections $\{g_\alpha^{(t)}\}_{\alpha\in\mathcal{I}_{nk}}$ is maximizing as well as $\{f_\alpha^{(t)}\}_{\alpha\in\mathcal{I}_{nk}}$.

Let us verify that $\{g_\alpha\}_{\alpha\in\mathcal{I}_{nk}} \in \Psi_c(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$. For every t there exists a collection of measurable subsets $\{A_\alpha^{(t)}\}_{\alpha\in\mathcal{I}_{nk}}$, $A_\alpha^{(t)} \subseteq X_\alpha$ such that $\mu_\alpha(A_\alpha^{(t)}) = 1$ and $G^{(t)}(x) \leq c(x)$ for all $x \in X$ such that $x_\alpha \in A_\alpha^{(t)}$ for all $\alpha \in \mathcal{I}_{nk}$. In addition, for each $\alpha \in \mathcal{I}_{nk}$ there exists a measurable subset $A'_\alpha \subseteq X_\alpha$ such that $\mu_\alpha(A'_\alpha) = 1$ and if $x_\alpha \in A'_\alpha$, then $g^{(t)}(x_\alpha) \to g_\alpha(x_\alpha)$ as $t \to \infty$.

For $\alpha \in \mathcal{I}_{nk}$, let

$$A_\alpha = A'_\alpha \cap \left(\bigcap_{t=1}^{\infty} A_\alpha^{(t)}\right).$$

For any $x \in \bigcap_{\alpha\in\mathcal{I}_{nk}} \mathrm{Pr}_\alpha^{-1}(A_\alpha)$ we have $\sum_{\alpha\in\mathcal{I}_{nk}} g_\alpha^{(t)}(x_\alpha) \leq c(x)$ for all t and

$$\sum_{\alpha\in\mathcal{I}_{nk}} g_\alpha^{(t)}(x_\alpha) \xrightarrow[t\to\infty]{} \sum_{\alpha\in\mathcal{I}_{nk}} g_\alpha(x_\alpha).$$

Thus, if $x \in \bigcap_{\alpha\in\mathcal{I}_{nk}} \mathrm{Pr}_\alpha^{-1}(A_\alpha)$, then $\sum_{\alpha\in\mathcal{I}_{nk}} g_\alpha(x_\alpha) \leq c(x)$, and therefore, since $\mu_\alpha(A_\alpha) = 1$, we conclude that $\{g_\alpha\}_{\alpha\in\mathcal{I}_{nk}} \in \Psi_c(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$.

Consider the finite (n,k)-function $G(x) = \sum_{\alpha\in\mathcal{I}_{nk}} g_\alpha(x_\alpha)$. We have

$$G^{(t)}(x) \xrightarrow[t\to\infty]{} G(x)$$

for all $x \in X$ except a zero (n,k)-thickness set, and therefore the sequence of functions $\{G^{(t)}\}$ converges pointwise to G μ-almost everywhere. In addition,

$$G^{(t)}(x) \leq \sum_{\alpha \in \mathcal{I}_{nk}} c_\alpha(x_\alpha) \in L^1(X, \mu)$$

for μ-almost all x, and therefore it follows from the reverse Fatou lemma that

$$J = \lim_{t \to \infty} \int\limits_X G^{(t)}(x)\, \mu(dx) \leq \int\limits_X G(x)\, \mu(dx) = \sum_{\alpha \in \mathcal{I}_{nk}} \int\limits_{X_\alpha} g_\alpha(x_\alpha)\, \mu_\alpha(x_\alpha).$$

Thus, the supremum in (11) is attained on the collection of functions $\{g_\alpha\}_{\alpha \in \mathcal{I}_{nk}}$. $\square$

Combining this result with Theorem 4.12, we get the following general duality theorem for the case of reducible projections.

Theorem 5.20 *(General duality theorem). For every $1 \leq i \leq n$, let X_i be a Polish space, let $\{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}}$, $\mu_\alpha \in \mathcal{P}(X_\alpha)$ be a reducible collection of probability measures, and let $c \in C_L(X, \{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}})$ be a continuous cost function on the space X. Then there exists a uniting measure $\pi \in \Pi(\{\mu_\alpha\}_{\alpha \in \mathcal{I}_{nk}})$ and a collection of integrable functions $\{f_\alpha\}_{\alpha \in \mathcal{I}_{nk}}$, $f_\alpha \colon X_\alpha \to [-\infty, +\infty)$, such that*

$$\sum_{\alpha \in \mathcal{I}_{nk}} f_\alpha(x_\alpha) \leq c(x) \text{ for all } x \in X$$

and

$$\int\limits_X c(x)\, \pi(dx) = \sum_{\alpha \in \mathcal{I}_{nk}} \int\limits_{X_\alpha} f_\alpha(x_\alpha)\, \mu_\alpha(dx_\alpha).$$

In particular, π is a solution to the related primal (n, k)-problem, and $\{f_\alpha\}_{\alpha \in \mathcal{I}_{nk}}$ is a solution to the related dual (n, k)-problem.

5.3. Unreachability of the supremum in the dual problem in the irreducible case

In contrast to the multi-marginal case, in the theorem proved above, the essential requirement is the irreducibility of the set of measures μ_α. In the following paragraph we construct a multistochastic $(3, 2)$-problem with a bounded continuous cost function such that the supremum in the corresponding dual problem can not be attained.

Let $X_1 = X_2 = X_3 = \mathbb{N}$. For $1 \leq i \leq 3$, the space X_i is a Polish space equipped with the discrete topology. For each n denote

$$A_n = \{(n+1, n, n), (n, n+1, n), (n, n, n+1)\}.$$

One can easily verify that these sets are pairwise disjoint.

Consider the measure μ_p on the space $X = X_1 \times X_2 \times X_3$ defined as follows:

$$\mu_p(n_1, n_2, n_3) = \begin{cases} \dfrac{2}{(\pi n)^2}, & \text{if } (n_1, n_2, n_3) \in A_n \text{ for some } n, \\ 0, & \text{otherwise.} \end{cases}$$

We have

$$\mu_p(X) = \sum_{n=1}^{\infty} |A_n| \cdot \frac{2}{(\pi n)^2} = \frac{6}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} = 1,$$

and therefore the measure μ_p is a probability measure.

Consider another measure μ_ε on the space X: let $\mu_\varepsilon(n_1, n_2, n_3) = 2^{-n_1-n_2-n_3}$ for all $(n_1, n_2, n_3) \in X$. We have

$$\mu_\varepsilon(X) = \sum_{(n_1,n_2,n_3)\in X} \frac{1}{2^{n_1+n_2+n_3}} = \left(\sum_{n_1=1}^{\infty} \frac{1}{2^{n_1}}\right) \cdot \left(\sum_{n_2=1}^{\infty} \frac{1}{2^{n_2}}\right) \cdot \left(\sum_{n_3=1}^{\infty} \frac{1}{2^{n_3}}\right) = 1,$$

and therefore μ_ε is a probability measure too.

Lemma 5.21. *Consider the probability measure $\mu = (1-\alpha)\mu_p + \alpha\mu_\varepsilon$, where $0 \le \alpha \le 1$. For $\{i,j\} \in \mathcal{I}_{3,2}$, denote $\mu_{ij} = \mathrm{Pr}_{ij}(\mu)$. If $\gamma \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$ is a uniting measure for the triple of projections $\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$, which means it has the same projections as μ, then*

$$\gamma(x) \ge \frac{2(1-\alpha)}{\pi^2}\left(\frac{1}{n^2} - \frac{1}{(n+1)^2}\right) - \frac{\alpha}{2^n}$$

for all $x \in A_n$ for all n.

Proof. First, let us find μ_{ij} explicitly. We have

$$\mathrm{Pr}_{ij}(\mu_\varepsilon)(n_i, n_j) = \sum_{n=1}^{\infty} \frac{1}{2^{n_i+n_j+n}} = \frac{1}{2^{n_i+n_j}} \text{ for all } (n_i, n_j) \in \mathbb{N}^2.$$

In addition, one can easily verify that

$$\mathrm{Pr}_{ij}(\mu_p)(n_i, n_j) = \begin{cases} 0, & \text{if } |n_i - n_j| \ge 2, \\ \frac{2}{(\pi n)^2}, & \text{if } |n_i - n_j| \le 1 \text{ and } \min(n_i, n_j) = n. \end{cases}$$

In particular, since $\mu_{ij} = (1-\alpha)\mathrm{Pr}_{ij}(\mu_p) + \alpha\mathrm{Pr}_{ij}(\mu_\varepsilon)$, we obtain the following equations:

$$\mu_{ij}(n_i, n_j) = \begin{cases} \dfrac{\alpha}{2^{n_i+n_j}} & \text{if } |n_i - n_j| \ge 2, \\ \frac{2(1-\alpha)}{(\pi n)^2} + \frac{\alpha}{2^{n_i+n_j}} & \text{if } |n_i - n_j| \le 1 \text{ and } \min(n_i, n_j) = n. \end{cases} \tag{13}$$

Fix a positive integer m. Consider the following functions $f_{ij} \colon \mathbb{N}^2 \to \mathbb{R}$:

$$f_{12}(n_1, n_2) = \begin{cases} 1, & \text{if } (n_1, n_2) = (m+1, m), \\ 0, & \text{otherwise;} \end{cases}$$

$$f_{13}(n_1, n_3) = \begin{cases} -1, & \text{if } n_1 = m+1 \text{ and } n_3 \in \{m-1, m+1\}, \\ 0, & \text{otherwise;} \end{cases}$$

$$f_{23}(n_2, n_3) = \begin{cases} -1, & \text{if } n_2 = m \text{ and } n_3 \notin \{m-1, m, m+1\}, \\ 0, & \text{otherwise.} \end{cases}$$

The function f_{ij} is bounded, and therefore is integrable with respect to μ_{ij}. Using equations (13) we get

$$\int f_{12}\, d\mu_{12} = \mu_{12}(m+1, m) = \frac{2(1-\alpha)}{\pi^2 m^2} + \frac{\alpha}{2^{2m+1}},$$
$$\int f_{13}\, d\mu_{13} = -\mu_{13}(m+1, m-1) - \mu_{13}(m+1, m+1) = -\frac{\alpha}{2^{2m}} - \frac{2(1-\alpha)}{\pi^2(m+1)^2} - \frac{\alpha}{2^{2m+2}},$$
$$\int f_{23}\, d\mu_{23} = -\sum_{n \notin \{m-1, m, m+1\}} \mu_{23}(m, n) = -\sum_{n=1}^{\infty} \frac{\alpha}{2^{m+n}} + \frac{\alpha}{2^{2m-1}} + \frac{\alpha}{2^{2m}} + \frac{\alpha}{2^{2m+1}} > -\frac{\alpha}{2^m} + \frac{\alpha}{2^{2m}}.$$

Summarizing this, we obtain

$$\begin{aligned}
&\int f_{12}\, d\mu_{12} + \int f_{13}\, d\mu_{13} + \int f_{23}\, d\mu_{23} \\
&\qquad > \frac{2(1-\alpha)}{\pi^2}\left(\frac{1}{m^2} - \frac{1}{(m+1)^2}\right) + \left(\frac{\alpha}{2^{2m+1}} - \frac{\alpha}{2^{2m+2}}\right) - \frac{\alpha}{2^m} \\
&\qquad > \frac{2(1-\alpha)}{\pi^2}\left(\frac{1}{m^2} - \frac{1}{(m+1)^2}\right) - \frac{\alpha}{2^m}.
\end{aligned} \tag{14}$$

Consider the $(3, 2)$-function

$$F(n_1, n_2, n_3) = f_{12}(n_1, n_2) + f_{13}(n_1, n_3) + f_{23}(n_2, n_3).$$

Let us verify that $F(n_1, n_2, n_3) \le 0$ if $(n_1, n_2, n_3) \ne (m+1, m, m)$. Indeed, since $f_{13} \le 0$ and $f_{23} \le 0$, we conclude that if $F(n_1, n_2, n_3) > 0$, then $f_{12}(n_1, n_2) > 0$, and therefore $(n_1, n_2) = (m+1, m)$. If $n_3 \notin \{m-1, m, m+1\}$, then by construction $f_{23}(m, n_3) = -1$, and $f_{13}(m+1, n_3) = 0$, and therefore $F(m+1, m, n_3) = 0$. Otherwise, if $n_3 \in \{m-1, m+1\}$, then $f_{13}(m+1, n_3) = -1$ and $f_{23}(m, n_3) = 0$, and therefore $F(m+1, m, n_3) = 0$ too.

In addition, $F(m+1, m, m) = 1$, and therefore if γ is a probability measure on the space X, then

$$\int_X F(n_1, n_2, n_3)\, \gamma(dn_1, dn_2, dn_3) \le \gamma(m+1, m, m).$$

Combining this with inequality (14), we conclude that if $\gamma \in \Pi(\{\mu_{ij}\}_{\{i,j\} \in \mathcal{I}_{3,2}})$, then

$$\begin{aligned}
\gamma(m+1, m, m) &\ge \int_X F(n_1, n_2, n_3)\, \gamma(dn_1, dn_2, dn_3) \\
&= \int f_{12}\, d\mu_{12} + \int f_{13}\, d\mu_{13} + \int f_{23}\, d\mu_{23} \\
&> \frac{2(1-\alpha)}{\pi^2}\left(\frac{1}{m^2} - \frac{1}{(m+1)^2}\right) - \frac{\alpha}{2^m}.
\end{aligned}$$

For the remaining points of A_m the inequality is proved in the same manner. □

Corollary 5.22. *There exists a real $\alpha_0 \in (0, 1)$ such that if $\gamma \in \Pi(\{\mu_{ij}\}_{\{i,j\} \in \mathcal{I}_{3,2}})$, then $\gamma(x) > 0$ for all $x \in A_n$ for all n, where $\mu_{ij} = \mathrm{Pr}_{ij}((1-\alpha_0)\mu_p + \alpha_0 \mu_\varepsilon)$.*

Proof. By Lemma 5.21 we only need to prove that there exists $\alpha_0 \in (0, 1)$ such that the inequality

$$\frac{2(1-\alpha_0)}{\pi^2}\left(\frac{1}{n^2} - \frac{1}{(n+1)^2}\right) - \frac{\alpha_0}{2^n} > 0$$

holds for all $n \in \mathbb{N}$, or equivalently

$$\frac{2(1-\alpha_0)}{\pi^2\alpha_0} > \frac{2^{-n}}{\frac{1}{n^2} - \frac{1}{(n+1)^2}}. \tag{15}$$

One can easily verify that the function in the right hand-side of the inequality converges to 0, and therefore there exists a constant M such that the inequality

$$M \geq \frac{2^{-n}}{\frac{1}{n^2} - \frac{1}{(n+1)^2}}$$

holds for all positive integer n. Thus, the inequality (15) follows from

$$\frac{2(1-\alpha_0)}{\pi^2\alpha_0} > M$$

and therefore every α_0 such that $0 < \alpha_0 < 2/(M\pi^2 + 2)$ is suitable. □

Theorem 5.23. *Let α_0 be the constant constructed in Corollary 5.22. Let $\mu = (1-\alpha_0)\mu_p + \alpha_0\mu_\varepsilon$, and for $\{i,j\} \in \mathcal{I}_{3,2}$ let $\mu_{ij} = \mathrm{Pr}_{ij}(\mu)$. Consider the cost function $c\colon X \to \{0,1\}$: $c(x) = 1$ if $x \in A_n$ for some n, and $c(x) = 0$ otherwise. Then the supremum in the corresponding relaxed dual $(3,2)$-problem (see Definition 5.16) can not be attained.*

Proof. The cost function c is a bounded continuous function on the space X equipped with the discrete topology. Since $\mu_{ij}(n_i, n_j) > 0$ for all $(n_i, n_j) \in X_{ij}$, the subset A of X is a zero $(3,2)$-thickness set if and only if $A = \varnothing$. Thus, the triple of functions $\{f_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$ belongs to $\Psi_c(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$ if and only if the inequality

$$f_{12}(n_1,n_2) + f_{13}(n_1,n_3) + f_{23}(n_2,n_3) \leq c(n_1,n_2,n_3)$$

holds for all $(n_1,n_2,n_3) \in X$.

The set $\Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$ is non-empty, and therefore it follows from Theorem 4.12 that

$$\min_{\gamma\in\Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})} \int_X c\, d\gamma = \sup\left\{ \sum \int_{X_{ij}} f_{ij}\, d\mu_{ij} \colon \sum f_{ij}(x_i,x_j) \leq c(x_1,x_2,x_3) \right\}.$$

Assume that the supremum in the dual problem is attained. Then there exists a uniting measure $\gamma \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$ and a triple of integrable functions $\{f_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$, $f_{ij}\colon X_{ij} \to [-\infty,+\infty)$ such that

$$f_{12}(n_1,n_2) + f_{13}(n_1,n_3) + f_{23}(n_2,n_3) \leq c(n_1,n_2,n_3)$$

for all $(n_1,n_2,n_3) \in X$ and

$$\int_{X_{12}} f_{12}\, d\mu_{12} + \int_{X_{13}} f_{13}\, d\mu_{13} + \int_{X_{23}} f_{23}\, d\mu_{23} = \int_X c\, d\gamma.$$

It follows from equation (13) that $\mu_{ij}(n_i,n_j) > 0$ for all pairs of positive integers (n_i,n_j). Hence, since f_{ij} is integrable with respect to μ_{ij}, we conclude that f_{ij} can not take value $-\infty$.

Consider the finite $(3,2)$-function

$$F(n_1,n_2,n_3) = f_{12}(n_1,n_2) + f_{13}(n_1,n_3) + f_{23}(n_2,n_3). \tag{16}$$

Since f_{ij} is integrable with respect to μ_{ij} and the measure γ is uniting, the function F is integrable with respect to γ and

$$\int_X F\,d\gamma = \int_{X_{12}} f_{12}\,d\mu_{12} + \int_{X_{13}} f_{13}\,d\mu_{13} + \int_{X_{23}} f_{23}\,d\mu_{23} = \int_X c\,d\gamma.$$

Since in addition $F(n_1,n_2,n_3) \le c(n_1,n_2,n_3)$ for all $(n_1,n_2,n_3) \in X$, we conclude that $F(n_1,n_2,n_3) = c(n_1,n_2,n_3)$ γ-almost everywhere. It follows from Corollary 5.22 that $\gamma(x) > 0$ if $x \in A_n$ for some n, and therefore

$$F(n+1,n,n) = F(n,n+1,n) = F(n,n,n+1) = 1 \tag{17}$$

for all $n \in \mathbb{N}$.

One can easily verify using equation (16) that for all $n \in \mathbb{N}$ we have

$$\begin{aligned} F(n+1,&n+1,n+1) - F(n,n,n) \\ &= \quad F(n,n+1,n+1) + F(n+1,n,n+1) + F(n+1,n+1,n) \\ &\quad - F(n+1,n,n) - F(n,n+1,n) - F(n,n,n+1). \end{aligned}$$

Since, $F(n_1,n_2,n_3) \le c(n_1,n_2,n_3)$ and $c(n_1,n_2,n_3) = 0$ if the point (n_1,n_2,n_3) is not contained in the set $\sqcup_{n=1}^{\infty} A_n$, the inequality

$$F(n,n+1,n+1) + F(n+1,n,n+1) + F(n+1,n+1,n) \le 0$$

holds for every positive integer n. In addition, it follows from equation (17) that

$$F(n+1,n,n) + F(n,n+1,n) + F(n,n,n+1) = 3.$$

Summarizing this, we conclude that $F(n+1,n+1,n+1) \le F(n,n,n) - 3$, and therefore

$$F(n,n,n) \le F(1,1,1) - 3(n-1) \le c(1,1,1) - 3(n-1) = -3(n-1),$$

for all $n \in \mathbb{N}$.

In particular, we conclude that for all $n \in \mathbb{N}$ the following inequality holds:

$$|f_{12}(n,n)| + |f_{13}(n,n)| + |f_{23}(n,n)| \ge 3(n-1).$$

Using this inequality and equation (13), we can bound from below the $\sum \|f_{ij}\|_{L^1(\mu_{ij})}$:

$$\begin{aligned} &\|f_{12}\|_{L^1(\mu_{12})} + \|f_{13}\|_{L^1(\mu_{13})} + \|f_{23}\|_{L^1(\mu_{23})} \\ &\qquad \ge \sum_{n=1}^{\infty} \left(|f_{12}(n,n)| \cdot \mu_{12}(n,n) + |f_{13}(n,n)| \cdot \mu_{13}(n,n) + |f_{23}(n,n)| \cdot \mu_{23}(n,n)\right) \\ &\qquad > \frac{2(1-\alpha_0)}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} \left(|f_{12}(n,n)| + |f_{13}(n,n)| + |f_{23}(n,n)|\right) \end{aligned}$$

$$\geq \frac{2(1-\alpha_0)}{\pi^2} \sum_{n=1}^{\infty} \frac{3(n-1)}{n^2} = +\infty.$$

Thus, at least one the functions f_{ij} is not integrable, and this contradiction proves Theorem 5.23. $\square$

The measure μ constructed in Theorem 5.23 is strictly positive at every point of the space X. In particular, this means that μ is equivalent to $\mathrm{Pr}_1(\mu)\otimes\mathrm{Pr}_2(\mu)\otimes\mathrm{Pr}_3(\mu)$. Thus, we obtain the following proposition, which demonstrates that in Theorem 5.19, we cannot replace "uniform equivalence" with simple equivalence.

Proposition 5.24. *Let $X_1 = X_2 = X_3 = \mathbb{N}$. There exists a probability measure μ on the space $X = X_1 \times X_2 \times X_3$ and a cost function $c : X \to \{0, 1\}$ such that the following conditions hold:*

(i) *measure μ is equivalent (but not uniformly equivalent) to $\mu_1 \otimes \mu_2 \otimes \mu_3$, where $\mu_i = \mathrm{Pr}_i\mu$;*
(ii) *there is no optimal solution to the relaxed dual problem for the cost function c and projections μ_{ij}, where $\mu_i = \mathrm{Pr}_i\mu$.*

In the classical Monge-Kantorovich problem the dual solution may not exist provided c is unbounded. In [34,3] authors introduce the concept of strong c-monotonicity, which generalizes the c-monotonicity and allows us to find a generalized dual solution.

Definition 5.25. A Borel set $\Gamma \subseteq X \times Y$ is *strongly c-monotone* if there exist Borel measurable functions $\varphi\colon X \to [-\infty, +\infty)$, $\psi\colon Y \to [-\infty, +\infty)$ such that $\varphi(x) + \psi(y) \leq c(x, y)$ for all $(x, y) \in X \times Y$ and $\varphi(x) + \psi(y) = c(x, y)$ holds if $(x, y) \in \Gamma$. A transport plan $\pi \in \Pi(\mu, \nu)$ is *strongly c-monotone* if π is concentrated on a strongly c-monotone Borel set.

One can easily verify that strong c-monotonicity implies c-monotonicity, and if there exists a solution to the dual problem, then every optimal transport plan is strongly c-monotone. In [3] authors prove that under general assumptions on the cost function the transport plan π is optimal if and only if π is strongly c-monotone.

Theorem 5.26 *([3, Theorem 3]). Let X, Y be Polish spaces equipped with Borel probability measures μ, ν, and let $c\colon X \times Y \to [0, \infty]$ be Borel measurable and $\mu\otimes\nu$-a.e. finite. Then a finite transport plan $\pi \in \Pi(\mu, \nu)$ is optimal if and only if it is strongly c-monotone.*

In particular, for every finite optimal transport plan π there exist (not necessary integrable) functions φ, ψ such that $\varphi(x) + \psi(x) \leq c(x, y)$ and the equality holds π-a.e. We can naturally generalize the concept of strong c-monotonicity to the multistochastic Monge-Kantorovich problem as follows.

Definition 5.27. A Borel set $\Gamma \subset X$ is *strongly c-monotone* if there exist Borel measurable functions $\{f_\alpha\}_{\alpha\in\mathcal{I}_{nk}}$, $f_\alpha\colon X_\alpha \to [-\infty, +\infty)$ such that the inequality

$$\sum_{\alpha\in\mathcal{I}_{nk}} f_\alpha(x_\alpha) \leq c(x)$$

holds for all $x \in X$ and the equality is achieved if $x \in \Gamma$. A transport plan $\pi \in \Pi(\{\mu_\alpha\}_{\alpha\in\mathcal{I}_{nk}})$ is *strongly c-monotone* if π is concentrated on a strongly c-monotone Borel set Γ.

We do not know whether there exists a strongly c-monotone transport plan in the problem considered in Theorem 5.23. In what follows, we construct another example of the $(3, 2)$-problem and prove that in this example there is no strongly c-monotone optimal transport plan.

As in the previous example, let $X_1 = X_2 = X_3 = \mathbb{N}$. For each n denote

$$B_n = \{(n, n+1, n+1), (n+1, n, n+1), (n+1, n+1, n)\}.$$

Consider the following measure μ defined on the space $X_1 \times X_2 \times X_3$ as follows:

$$\mu(n_1, n_2, n_3) = \begin{cases} \dfrac{1}{(\pi n)^2} & \text{if } (n_1, n_2, n_3) \in A_n \sqcup B_n \text{ for some } n, \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

One can check that μ is a probability measure. Finally, for $\{i, j\} \in \mathcal{I}_{3,2}$ denote $\mu_{ij} = \mathrm{Pr}_{ij}(\mu)$.

Lemma 5.28. *The measure μ is the only uniting measure for the triple of projections $\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$.*

Proof. Let $\gamma \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$. For $\{i, j\} \in \mathcal{I}_{3,2}$, the projection μ_{ij} is concentrated on the set $\{(n_i, n_j) \in \mathbb{N}^2 \colon |n_i - n_j| \le 1\}$, and therefore the transport plan γ is concentrated on the set

$$S = \{(n_1, n_2, n_3) \in \mathbb{N}^3 \colon \max\{n_1, n_2, n_3\} - \min\{n_1, n_2, n_3\} \le 1\} = \bigsqcup_{k=1}^{\infty} \left(\{(k, k, k)\} \sqcup A_k \sqcup B_k\right).$$

Let us prove that γ is uniquely defined by its values on the diagonal, and if we denote $a_k = \gamma(k, k, k)$, then we have

$$\gamma(n_1, n_2, n_3) = \begin{cases} \mu(n_1, n_2, n_3) - (a_1 + \dots + a_n) & \text{if } (n_1, n_2, n_3) \in A_n \text{ for some } n, \\ \mu(n_1, n_2, n_3) + (a_1 + \dots + a_n) & \text{if } (n_1, n_2, n_3) \in B_n \text{ for some } n. \end{cases} \tag{19}$$

Since $\gamma \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$, for all $n \in \mathbb{N}$ we have

$$\mathrm{Pr}_{12}(\gamma)(n, n) = \mu_{12}(n, n).$$

The left-hand side of the equation is equal to $\sum_{k\ge1} \gamma(n, n, k) = a_n + \gamma(n, n, n-1) + \gamma(n, n, n+1)$, and the right-hand side is equal to $\sum_{k\ge1} \mu(n, n, k) = \mu(n, n, n-1) + \mu(n, n, n+1)$. Here, we assume that $\mu(n_1, n_2, n_3) = \gamma(n_1, n_2, n_3) = 0$ if at least one of the variables (n_1, n_2, n_3) is less than 1. Thus,

$$\begin{aligned} &\gamma(n, n, n+1) + \gamma(n, n, n-1) + a_n = \mu(n, n, n+1) + \mu(n, n, n-1) \\ &\qquad \Leftrightarrow \gamma(n, n, n+1) - \mu(n, n, n+1) = \mu(n, n, n-1) - \gamma(n, n, n-1) - a_n. \end{aligned} \tag{20}$$

In addition,

$$\begin{aligned} \mathrm{Pr}_{13}(\gamma)(n, n-1) &= \mu_{13}(n, n-1) \\ &\Leftrightarrow \gamma(n, n-1, n-1) + \gamma(n, n, n-1) = \mu(n, n-1, n-1) + \mu(n, n, n-1) \\ &\Leftrightarrow \mu(n, n, n-1) - \gamma(n, n, n-1) = \gamma(n, n-1, n-1) - \mu(n, n-1, n-1). \end{aligned} \tag{21}$$

Thus, it follows from equations (20) and (21) that

$$\gamma(n, n, n+1) - \mu(n, n, n+1) = \gamma(n, n-1, n-1) - \mu(n, n-1, n-1) - a_n. \tag{22}$$

Due to the symmetry, one can interchange the second and the third variable:

$$\gamma(n, n+1, n) - \mu(n, n+1, n) = \gamma(n, n-1, n-1) - \mu(n, n-1, n-1) - a_n. \tag{23}$$

After that one can interchange the first and the third variable:

$$\gamma(n,n+1,n)-\mu(n,n+1,n)=\gamma(n-1,n-1,n)-\mu(n-1,n-1,n)-a_n. \tag{24}$$

Equations (22) and (23) have identical right-hand sides, and equations (23) and (24) have identical left-hand sides. Thus, the left-hand side of (22) is equal to the right-hand side of (24), i.e. the following equation holds for all $n\in\mathbb{N}$:

$$\gamma(n,n,n+1)-\mu(n,n,n+1)=\gamma(n-1,n-1,n)-\mu(n-1,n-1,n)-a_n.$$

Applying this formula n times, we conclude that

$$\gamma(n,n,n+1)-\mu(n,n,n+1)=\gamma(0,0,1)-\mu(0,0,1)-(a_1+\cdots+a_n)=-(a_1+\cdots+a_n).$$

So, due to the symmetry,

$$\gamma(n_1,n_2,n_3)=\mu(n_1,n_2,n_3)-(a_1+\cdots+a_n)\text{ for all }(n_1,n_2,n_3)\in A_n.$$

Substituting the last equation into equation (20), we conclude that

$$\mu(n,n,n-1)-\gamma(n,n,n-1)=-(a_1+\cdots+a_{n-1});$$

therefore, due to the symmetry,

$$\gamma(n_1,n_2,n_3)=\mu(n_1,n_2,n_3)+(a_1+\cdots+a_n)\text{ for all }(n_1,n_2,n_3)\in B_n,$$

and this completes the proof of formula (19).

We have $\mu(n_1,n_2,n_3)=(\pi n)^{-2}$ for all $(n_1,n_2,n_3)\in A_n$, and therefore $a_1+\cdots+a_n\le(\pi n)^{-2}$ for all n. Thus, since all a_n are nonnegative, we conclude that $\gamma(k,k,k)=a_k=0$ for all k, and therefore $\gamma=\mu$ by equation (19). $\square$

It follows from the previous lemma that μ is the unique optimal solution to the multistochastic problem with arbitrary bounded cost function. Next, we construct the cost function c such that μ is not strongly c-monotone. The existence of this example demonstrates that we can not generalize the equivalence of optimality and strongly c-monotonicity to the multistochastic case.

Theorem 5.29. *Let μ be the measure on $\mathbb{N}^3$ defined in equation (18), and let $\mu_{ij}=\mathrm{Pr}_{ij}(\mu)$. Consider the cost function $c\colon\mathbb{N}^3\to\{0,1\}$: $c(x)=1$ if $x\in B_n$ for some n, and $c(x)=0$ otherwise. Then there are no functions $\{f_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$, $f_{ij}\colon\mathbb{N}^2\to[-\infty,+\infty)$ such that*

$$f_{12}(n_1,n_2)+f_{13}(n_1,n_3)+f_{23}(n_2,n_3)\le c(n_1,n_2,n_3)$$

for all $(n_1,n_2,n_3)\in\mathbb{N}^3$ and the equality holds μ-a.e.

Proof. Assuming the opposite, consider the following $(3,2)$-function:

$$F(n_1,n_2,n_3)=f_{12}(n_1,n_2)+f_{13}(n_1,n_3)+f_{23}(n_2,n_3). \tag{25}$$

Since $\{f_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$ satisfy the assumptions of the theorem, we have $F(n_1,n_2,n_3)=c(n_1,n_2,n_3)$ μ-a.e. Hence, since $\mu(n_1,n_2,n_3)>0$ for all $(n_1,n_2,n_3)\in A_n\sqcup B_n$, we get

$$\begin{aligned} &F(n+1,n,n) = F(n,n+1,n) = F(n,n,n+1) = 0, \\ &F(n,n+1,n+1) = F(n+1,n,n+1) = F(n+1,n+1,n) = 1 \end{aligned} \tag{26}$$

for all $n \in \mathbb{N}$.

Applying (25) one can easily verify the following equation:

$$\begin{aligned} F(n,n,n) &+ F(n,n+1,n+1) + F(n+1,n,n+1) + F(n+1,n+1,n) \\ &= F(n+1,n+1,n+1) + F(n+1,n,n) + F(n,n+1,n) + F(n,n,n+1). \end{aligned}$$

Combining this with equation (26), we get $F(n,n,n) + 3 = F(n+1,n+1,n+1)$ for all n, and therefore the inequality

$$F(n,n,n) = F(n+k,n+k,n+k) - 3k \leq c(n+k,n+k,n+k) - 3k \leq 1 - 3k$$

holds for all $n,k \in \mathbb{N}$. Thus, $F(n,n,n) = -\infty$ for all n. In particular, $F(1,1,1) = -\infty$, and therefore $f_{ij}(1,1) = -\infty$ for some $\{i,j\} \in \mathcal{I}_{3,2}$. Without loss of generality we may assume that $f_{12}(1,1) = -\infty$. Then $F(1,1,2)$ is also equal to $-\infty$, and this contradicts equation (26). $\square$

6. Properties of the dual solution in $(3,2)$-problem

6.1. Boundedness of the dual solution

In the classical Monge-Kantorovich problem for the bounded cost function $c(x,y)$ we can transform every solution to the dual problem to the bounded one, using Legendre transformation.

Proposition 6.1. *Let X and Y be Polish spaces, let $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$, and let $c\colon X \times Y \to \mathbb{R}_+$ be a cost function. If c is a bounded continuous cost function, then there exists a solution (φ, ψ) to the related dual problem such that both $\varphi(x), \psi(y)$ lie between $-\|c\|_\infty$ and $\|c\|_\infty$ for all $x \in X$ and $y \in Y$.*

Proof. The proof is an adaptation of the argument from the proof of [36, Theorem 1.3]. Let (φ, ψ) be a solution to the dual problem provided by [32, Theorem 2.4.3]. If π is a solution to the related primal problem, then $\varphi(x) + \psi(y) = c(x,y)$ π-a.e. In particular, there exists a point $(x_0, y_0) \in X \times Y$ such that $\varphi(x_0) + \psi(y_0) = c(x_0,y_0) \geq 0$. For any real number s the pair of functions $(\varphi - s, \psi + s)$ is also a solution to the dual problem. By a proper choice of s, we can ensure

$$\varphi(x_0) \geq 0, \;\; \psi(y_0) \geq 0.$$

Since $\varphi(x) + \psi(y) \leq c(x,y)$, we have $\varphi(x) \leq c(x,y_0) - \varphi(y_0) \leq c(x,y_0)$ for all x, and $\psi(y) \leq c(x_0,y) - \varphi(x_0) \leq c(x_0,y)$ for all y. Consider the Legendre transformation of the function φ:

$$\overline{\varphi}(x) = \inf_{y \in Y} (c(x,y) - \psi(y)).$$

By construction, $\overline{\varphi}(x) + \psi(y) \leq c(x,y)$ for all $x \in X$ and for all $y \in Y$. From the inequality $\varphi(x) \leq c(x,y) - \psi(y)$ we see that $\overline{\varphi}(x) \geq \varphi(x)$ for all x. Since $\overline{\varphi}(x) \leq c(x,y) - \psi(y)$ for all y, we have

$$\overline{\varphi}(x) \leq c(x,y_0) - \psi(y_0) \leq \|c\|_\infty,$$

and it follows from the inequality $\psi(y) \leq c(x_0,y)$ that

$$\overline{\varphi}(x) \geq \inf_{y \in Y} (c(x,y) - c(x_0,y)) \geq -\|c\|_\infty .$$

Hence, $\overline{\varphi}$ is an integrable function; since $\overline{\varphi}(x) \geq \varphi(x)$ for all x, we have

$$\int_X \overline{\varphi}(x)\,\mu(dx) + \int_Y \psi(y)\,\nu(dy) \geq \int_X \varphi(x)\,\mu(dx) + \int_Y \psi(y)\,\nu(dy),$$

and therefore $(\overline{\varphi}, \psi)$ is a solution to the dual problem.

Finally, define

$$\overline{\psi}(y) = \inf_{x \in X} (c(x,y) - \overline{\varphi}(x)).$$

By the same arguments we conclude that $(\overline{\varphi}, \overline{\psi})$ is a solution to the dual problem and $-\|c\|_\infty \leq \overline{\psi}(y) \leq \|c\|_\infty$ for all $y \in Y$. $\square$

We want to generalize this observation to the multistochastic case.

Definition 6.2. Given finite measures μ and ν on the space X, we say that $\mu \ll_b \nu$ (where “b” means “bounded”) if there exists a positive real M such that $\mu \leq M \cdot \nu$.

The following properties trivially follow from the definition.

Proposition 6.3. *Let μ and ν be finite measures on the space X. Suppose that $\mu \ll_b \nu$. Then*

(a) *μ is absolutely continuous with respect to ν;*
(b) *$L^1(X,\mu) \supseteq L^1(X,\nu)$;*
(c) *if $X = X_1 \times \cdots \times X_n$, then $\mathrm{Pr}_\alpha \mu \ll_b \mathrm{Pr}_\alpha \nu$ for all $\alpha \in \mathcal{I}_n$.*

Definition 6.4. Let $X_1, \ldots, X_n$ be Polish spaces, let $\pi \in \mathcal{P}(X)$, and let ν_α be a probability measure on X_α for some $\alpha \in \mathcal{I}_n$ such that $\nu_\alpha \ll_b \pi_\alpha$. Let ρ be a density function of ν_α with respect to π_α. Then denote by $\mathrm{Up}_\alpha(\nu_\alpha, \pi)$ the measure $\rho^*(x) \cdot \pi$, where $\rho^*(x) = \rho(x_\alpha)$ for all $x \in X$.

Proposition 6.5. *Let $X_1, \ldots, X_n$ be Polish spaces, let $\pi \in \mathcal{P}(X)$, and let ν_α be a probability measure on X_α for some $\alpha \in \mathcal{I}_n$ such that $\nu_\alpha \ll_b \pi_\alpha$. Then*

(a) *the measure $\mathrm{Up}_\alpha(\nu_\alpha, \pi)$ is well-defined;*
(b) *$\mathrm{Up}_\alpha(\nu_\alpha, \pi) \ll_b \pi$;*
(c) *if $\beta \supseteq \alpha$, then $\mathrm{Pr}_\beta(\mathrm{Up}_\alpha(\nu_\alpha, \pi)) = \mathrm{Up}_\alpha(\nu_\alpha, \pi_\beta)$;*
(d) *if $\beta \subseteq \alpha$, then $\mathrm{Pr}_\beta(\mathrm{Up}_\alpha(\nu_\alpha, \pi)) = \mathrm{Pr}_\beta(\nu_\alpha)$;*
(e) *if $\pi = \mu_\alpha \otimes \mu_\beta$ for $\mu_\alpha \in \mathcal{P}(X_\alpha)$ and $\mu_\beta \in \mathcal{P}(X_\beta)$, then $\mathrm{Up}_\alpha(\nu_\alpha, \pi) = \nu_\alpha \otimes \mu_\beta$.*

Proof. Assertion 6.5(a) is trivial: if $\nu_\alpha = \rho_1 \cdot \pi_\alpha = \rho_2 \cdot \pi_\alpha$, then $\rho_1(x_\alpha) = \rho_2(x_\alpha)$ for π_α-a.e. $x_\alpha \in X_\alpha$, and therefore $\rho_1^*(x) = \rho_2^*(x)$ for π-a.e. $x \in X$. In addition, since $\nu_\alpha \ll_b \pi_\alpha$, there exists a positive real M such that $\rho(x_\alpha) \leq M$ for π_α-a.e. $x_\alpha \in X_\alpha$, and therefore $\rho^*(x) \leq M$ for π-a.e. $x \in X$. Hence, $\rho^* \in L^1(X, \pi)$ and the measure $\rho^* \cdot \pi$ is well-defined. Furthermore, since $\rho^* \leq M$ π-a.e, we have $\mathrm{Up}(\nu_\alpha, \pi) \leq M \cdot \pi$; thus, $\mathrm{Up}(\nu_\alpha, \pi) \ll_b \pi$ and assertion 6.5(b) holds.

We have $\mathrm{Up}(\nu_\alpha, \pi) = \rho(x_\alpha) \cdot \pi$. The function ρ does not depend on coordinates x_i for all $i \notin \alpha$. Hence, if $\beta \supseteq \alpha$ and $\beta \in \mathcal{I}_n$, then $\mathrm{Pr}_\beta(\rho(x_\alpha) \cdot \pi) = \rho(x_\alpha) \cdot \pi_\beta$. Since $\mathrm{Pr}_\alpha(\pi_\beta) = \pi_\alpha$, we conclude that $\mathrm{Up}_\alpha(\nu_\alpha, \pi_\beta) =$

$\rho(x_\alpha) \cdot \pi_\beta$. Thus, if $\beta \supseteq \alpha$, then $\mathrm{Pr}_\beta(\mathrm{Up}_\alpha(\nu_\alpha, \pi)) = \mathrm{Up}_\alpha(\nu_\alpha, \mathrm{Pr})$, and this implies assertion 6.5(c). In addition, we have $\mathrm{Pr}_\alpha(\rho(x_\alpha) \cdot \pi) = \rho(x_\alpha) \cdot \pi_\alpha = \nu_\alpha$, and therefore $\mathrm{Pr}_\alpha(\mathrm{Up}_\alpha(\nu_\alpha, \pi)) = \nu_\alpha$. Hence, if $\beta \subseteq \alpha$, then $\mathrm{Pr}_\beta(\mathrm{Up}_\alpha(\nu_\alpha, \pi)) = \mathrm{Pr}_\beta \circ \mathrm{Pr}_\alpha(\mathrm{Up}_\alpha(\nu_\alpha, \pi)) = \mathrm{Pr}_\beta(\nu_\alpha)$, and this implies assertion 6.5(d).

Finally, suppose that $\pi = \mu_\alpha \otimes \mu_\beta$. Then $\pi_\alpha = \mu_\alpha$, and therefore $\nu_\alpha = \rho \cdot \pi_\alpha = \rho \cdot \mu_\alpha$. Thus, $\nu_\alpha \otimes \mu_\beta = (\rho(x_\alpha) \cdot \mu_\alpha) \otimes \mu_\beta = \rho(x_\alpha) \cdot \pi = \mathrm{Up}_\alpha(\nu_\alpha, \pi)$, and this implies assertion 6.5(e). $\square$

Let X_1, X_2, X_3 be Polish spaces, let $\mu_i \in \mathcal{P}(X_i)$ for $1 \le i \le i$, and let $\mu_{ij} = \mu_i \otimes \mu_j$ for all $\{i, j\} \in \mathcal{I}_{3,2}$. Let $c\colon X \to \mathbb{R}_+$ be a nonnegative bounded continuous cost function. The space $\Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$ is non-empty, since $\mu_1 \otimes \mu_2 \otimes \mu_3 \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$, and therefore by Theorem 4.12 there is no duality gap. In addition, since the family of measures $\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$ is reducible, by Theorem 5.19 there exists a solution to the related dual problem. Thus, there exists a solution $\pi \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$ to the primal problem and a solution $\{f_{ij}\}$, $f_{ij} \in L^1(X_{ij}, \mu_{ij})$ to the dual problem, and

$$\int_X c\, d\pi = \int_{X_{12}} f_{12}\, d\mu_{12} + \int_{X_{13}} f_{13}\, d\mu_{13} + \int_{X_{23}} f_{23}\, d\mu_{23}.$$

Lemma 6.6. *Let $\widetilde{\pi}$ be a probability measure on X. Suppose that there exists $\gamma \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$ such that $\widetilde{\pi} \ll_b \gamma$. Then extensions of all f_{12}, f_{13} and f_{23} to the space X are integrable with respect to the measure $\widetilde{\pi}$.*

Proof. The extension of f_{ij} is integrable with respect to $\widetilde{\pi}$ if and only if $f_{ij} \in L^1(X_{ij}, \mathrm{Pr}_{ij}(\widetilde{\pi}))$. Since $\widetilde{\pi} \ll_b \gamma$, by assertion 6.3(c) we have $\mathrm{Pr}_{ij}(\widetilde{\pi}) \ll_b \mathrm{Pr}_{ij}(\gamma) = \mu_{ij}$, and therefore by assertion 6.3(b) we conclude that $L^1(X_{ij}, \mathrm{Pr}_{ij}(\widetilde{\pi})) \supseteq L^1(X_{ij}, \mu_{ij}) \ni f_{ij}$. $\square$

Denote $F(x_1, x_2, x_3) = f_{12}(x_1, x_2) + f_{13}(x_1, x_3) + f_{23}(x_2, x_3)$.

Lemma 6.7. *Let $\widetilde{\pi}$ be a probability measure on X. Suppose that there exists $\gamma \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$ such that $\widetilde{\pi} \ll_b \gamma$. Then*

(a) *the function F and the extensions of all f_{12}, f_{13} and f_{23} to the space X are integrable with respect to the measure $\widetilde{\pi}$;*
(b) $\displaystyle\int_X F\, d\widetilde{\pi} \le \|c\|_\infty$;
(c) *if $\widetilde{\pi} \ll_b \pi$, then* $\displaystyle\int_X F\, d\widetilde{\pi} \ge 0$.

Proof. Assertion 6.7(a) trivially follows from Lemma 6.6. Since $\{f_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$ is a solution to the dual problem, we have $F(x_1, x_2, x_3) \le c(x_1, x_2, x_3)$ for all $x \in X$. In particular,

$$\int_X F\, d\widetilde{\pi} \le \int_X c\, d\widetilde{\pi} \le \|c\|_\infty\,,$$

and this implies assertion 6.7(b).

Since $\mathrm{Pr}_{ij}(\pi) = \mu_{ij}$, by assertion 6.7(a) the function $F \in L^1(X, \pi)$ and

$$\begin{aligned}\int_X F\, d\pi &= \int_{X_{12}} f_{12}\, d\pi + \int_{X_{13}} f_{13}\, d\pi + \int_{X_{23}} f_{23}\, d\pi \\ &= \int_{X_{12}} f_{12}\, d\mu_{12} + \int_{X_{13}} f_{13}\, d\mu_{13} + \int_{X_{23}} f_{23}\, d\mu_{23} = \int_X c\, d\pi.\end{aligned}$$

Since in addition $F(x_1,x_2,x_3) \le c(x_1,x_2,x_3)$ for all $x \in X$, we conclude that $F(x_1,x_2,x_3) = c(x_1,x_2,x_3)$ for π-a.e. $x \in X$. Thus, if $\widetilde{\pi} \ll_b \pi$, then $F(x_1,x_2,x_3) = c(x_1,x_2,x_3)$ $\widetilde{\pi}$-a.e., and therefore

$$\int\limits_X F\,d\widetilde{\pi} = \int\limits_X c\,d\widetilde{\pi} \ge 0$$

since $c \ge 0$. This implies assertion 6.7(c). □

Lemma 6.8. *Let (i,j,k) be a permutation of indices $(1,2,3)$. Let ν_i be a probability measure on X_i such that $\nu_i \ll_b \mu_i$. Then $F \in L^1(X, \nu_i \otimes \mu_j \otimes \mu_k)$ and*

$$\int\limits_X F\,d(\nu_i \otimes \mu_j \otimes \mu_k) \ge \int\limits_X F\,d\pi - \|c\|_\infty .$$

Proof. Since $\nu_i \ll_b \mu_i$, we have $\nu_i \ll_b \mathrm{Pr}_i(\pi)$, and therefore the measure $\mathrm{Up}_i(\nu_i,\pi)$ is well-defined. Consider the following measure:

$$\gamma = \nu_i \otimes \mu_j \otimes \mu_k - \mathrm{Up}_i(\nu_i,\pi) + \mu_i \otimes \mathrm{Pr}_{jk}(\mathrm{Up}_i(\nu_i,\pi)) - \pi. \tag{27}$$

We claim that all the projections of γ to the spaces X_{ij}, X_{ik} and X_{jk} are zero measures. First, by assertion 6.5(c) and assertion 6.5(e) we have

$$\begin{aligned}
\mathrm{Pr}_{ij}(\mathrm{Up}_i(\nu_i,\pi)) &= \mathrm{Up}_i(\nu_i, \mathrm{Pr}_{ij}(\pi)) = \mathrm{Up}_i(\nu_i, \mu_i \otimes \mu_j) = \nu_i \otimes \mu_j,\\
\mathrm{Pr}_{ik}(\mathrm{Up}_i(\nu_i,\pi)) &= \mathrm{Up}_i(\nu_i, \mathrm{Pr}_{ik}(\pi)) = \mathrm{Up}_i(\nu_i, \mu_i \otimes \mu_k) = \nu_i \otimes \mu_k.
\end{aligned}$$

Next, we find the projections of $\mathrm{Up}_i(\nu_i,\pi)$ to the spaces X_j and X_k:

$$\begin{aligned}
\mathrm{Pr}_j(\mathrm{Up}_i(\nu_i,\pi)) &= \mathrm{Pr}_j \circ \mathrm{Pr}_{ij}(\mathrm{Up}_i(\nu_i,\pi)) = \mathrm{Pr}_j(\nu_i \otimes \mu_j) = \mu_j,\\
\mathrm{Pr}_k(\mathrm{Up}_i(\nu_i,\pi)) &= \mathrm{Pr}_k \circ \mathrm{Pr}_{ik}(\mathrm{Up}_i(\nu_i,\pi)) = \mathrm{Pr}_k(\nu_i \otimes \mu_k) = \mu_k.
\end{aligned}$$

Finally, we find the projections of γ to the spaces X_{ij}, X_{ik} and X_{jk}:

$$\begin{aligned}
\mathrm{Pr}_{ij}(\gamma) &= \mathrm{Pr}_{ij}(\nu_i \otimes \mu_j \otimes \mu_k) - \mathrm{Pr}_{ij}(\mathrm{Up}_i(\nu_i,\pi)) + \mathrm{Pr}_{ij}(\mu_i \otimes \mathrm{Pr}_{jk}(\mathrm{Up}_i(\nu_i,\pi))) - \mathrm{Pr}_{ij}(\pi)\\
&= \nu_i \otimes \mu_j - \nu_i \otimes \mu_j + \mu_i \otimes \mathrm{Pr}_j(\mathrm{Up}_i(\nu_i,\pi)) - \mu_i \otimes \mu_j\\
&= \nu_i \otimes \mu_j - \nu_i \otimes \mu_j + \mu_i \otimes \mu_j - \mu_i \otimes \mu_j = 0;\\
\mathrm{Pr}_{ik}(\gamma) &= \mathrm{Pr}_{ik}(\nu_i \otimes \mu_j \otimes \mu_k) - \mathrm{Pr}_{ik}(\mathrm{Up}_i(\nu_i,\pi)) + \mathrm{Pr}_{ik}(\mu_i \otimes \mathrm{Pr}_{jk}(\mathrm{Up}_i(\nu_i,\pi))) - \mathrm{Pr}_{ik}(\pi)\\
&= \nu_i \otimes \mu_k - \nu_i \otimes \mu_k + \mu_i \otimes \mathrm{Pr}_k(\mathrm{Up}_i(\nu_i,\pi)) - \mu_i \otimes \mu_k\\
&= \nu_i \otimes \mu_k - \nu_i \otimes \mu_k + \mu_i \otimes \mu_k - \mu_i \otimes \mu_k = 0;\\
\mathrm{Pr}_{jk}(\gamma) &= \mathrm{Pr}_{jk}(\nu_i \otimes \mu_j \otimes \mu_k) - \mathrm{Pr}_{jk}(\mathrm{Up}_i(\nu_i,\pi)) + \mathrm{Pr}_{jk}(\mu_i \otimes \mathrm{Pr}_{jk}(\mathrm{Up}_i(\nu_i,\pi))) - \mathrm{Pr}_{jk}(\pi)\\
&= \mu_j \otimes \mu_k - \mathrm{Pr}_{jk}(\mathrm{Up}_i(\nu_i,\pi)) + \mathrm{Pr}_{jk}(\mathrm{Up}_i(\nu_i,\pi)) - \mu_j \otimes \mu_k = 0.
\end{aligned}$$

Since $\nu_i \ll_b \mu_i$, we have

$$\nu_i \otimes \mu_j \otimes \mu_k \ll_b \mu_i \otimes \mu_j \otimes \mu_k \in \Pi(\{\mu_{ij}\}_{\{i,j\} \in \mathcal{I}_{3,2}}).$$

Besides, it follows from assertion 6.5(b) that

$$\mathrm{Up}_i(\nu_i,\pi) \ll_b \pi \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}).$$

Hence, by assertion 6.3(c) we have $\mathrm{Pr}_{jk}(\mathrm{Up}_i(\nu_i,\pi)) \ll_b \mathrm{Pr}_{jk}(\pi) = \mu_{jk} = \mu_j \otimes \mu_k$, and therefore

$$\mu_i \otimes \mathrm{Pr}_{jk}(\mathrm{Up}_i(\nu_i,\pi)) \ll_b \mu_i \otimes \mu_j \otimes \mu_k \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}).$$

Thus, it follows from assertion 6.7(a) that the function F and the extension of all f_{12}, f_{13}, and f_{23} to the space X are integrable with respect to all of the summands of equation (27), and therefore that functions are integrable with respect to γ. In particular,

$$\int\limits_X F\,d\gamma = \int\limits_{X_{ij}} f_{ij}\,d\mathrm{Pr}_{ij}(\gamma) + \int\limits_{X_{ik}} f_{ik}\,d\mathrm{Pr}_{ik}(\gamma) + \int\limits_{X_{jk}} f_{jk}\,d\mathrm{Pr}_{jk}(\gamma) = 0.$$

On the other hand, we have

$$\int\limits_X F\,d\gamma = \int\limits_X F\,d(\nu_i\otimes\mu_j\otimes\mu_k) - \int\limits_X F\,d\mathrm{Up}_i(\nu_i,\pi) + \int\limits_X F\,d(\mu_i\otimes\mathrm{Pr}_{jk}(\mathrm{Up}_i(\nu_i,\pi))) - \int\limits_X F\,d\pi.$$

Since $\mathrm{Up}_i(\nu_i,\pi) \ll_b \pi$, by assertion 6.7(c) we have

$$\int\limits_X F\,d\mathrm{Up}_i(\nu_i,\pi) \geq 0.$$

By assertion 6.7(b) we have

$$\int\limits_X F\,d(\mu_i\otimes\mathrm{Pr}_{jk}(\mathrm{Up}_i(\nu_i,\pi))) \leq \|c\|_\infty.$$

Thus, we get

$$0 = \int\limits_X F\,d\gamma \leq \int\limits_X F\,d(\nu_i\otimes\mu_j\otimes\mu_k) - \int\limits_X F\,d\pi + \|c\|_\infty. \quad \square$$

Lemma 6.9. *For $1 \leq i \leq 3$, let ν_i be a probability measure on X_i such that $\nu_i \ll_b \mu_i$. Then $F \in L^1(X, \nu_1 \otimes \nu_2 \otimes \nu_3)$ and*

$$\int\limits_X F\,d(\nu_1\otimes\nu_2\otimes\nu_3) \geq -12\,\|c\|_\infty.$$

Proof. The proof is similar to the proof of Lemma 6.8. We have $\nu_i \otimes \nu_j \ll_b \mu_i \otimes \mu_j = \mathrm{Pr}_{ij}(\pi)$, and therefore the measure $\mathrm{Up}_{ij}(\nu_i\otimes\nu_j,\pi)$ is well-defined for all $\{i,j\}\in\mathcal{I}_{3,2}$. Consider the following measures:

$$\begin{aligned}
\gamma^{(0)} &= \sum_{\{i,j\}\in\mathcal{I}_{3,2}} \mathrm{Up}_{ij}(\nu_i\otimes\nu_j,\pi);\\
\gamma^{(1)} &= \sum_{(i,j,k)\in S_3} \mu_i\otimes\mathrm{Pr}_{jk}(\mathrm{Up}_{ij}(\nu_i\otimes\nu_j,\pi));\\
\gamma^{(2)} &= \sum_{\substack{\{i,j\}\in\mathcal{I}_{3,2}\\ \{i,j,k\}=\{1,2,3\}}} \mu_i\otimes\mu_j\otimes\mathrm{Pr}_k(\mathrm{Up}_{ij}(\nu_i\otimes\nu_j,\pi));
\end{aligned}$$

$$\gamma^{(3)} = \sum_{\substack{\{i,j\}\in\mathcal{I}_{3,2}\\ \{i,j,k\}=\{1,2,3\}}} \mu_i \otimes \mu_j \otimes \nu_k;$$

$$\gamma = \nu_1 \otimes \nu_2 \otimes \nu_3 - \gamma^{(0)} + \gamma^{(1)} - \gamma^{(2)} - \gamma^{(3)} + 2\pi.$$

We claim that $\mathrm{Pr}_{ij}(\gamma) = 0$ for all $\{i,j\} \in \mathcal{I}_{3,2}$. Let (i,j,k) be a permutation of indices $(1,2,3)$. By construction,

$$\gamma^{(0)} = \mathrm{Up}_{ij}(\nu_i \otimes \nu_j, \pi) + \mathrm{Up}_{ik}(\nu_i \otimes \nu_k, \pi) + \mathrm{Up}_{jk}(\nu_j \otimes \nu_k, \pi).$$

It follows from assertion 6.5(c) that

$$\mathrm{Pr}_{ij}(\mathrm{Up}_{ij}(\nu_i \otimes \nu_j, \pi)) = \nu_i \otimes \nu_j,$$

and therefore

$$\mathrm{Pr}_{ij}(\gamma^{(0)}) = \nu_i \otimes \nu_j + \mathrm{Pr}_{ij}(\mathrm{Up}_{ik}(\nu_i \otimes \nu_k, \pi)) + \mathrm{Pr}_{ij}(\mathrm{Up}_{jk}(\nu_j \otimes \nu_k, \pi)). \tag{28}$$

Next, let us find the projection of $\gamma^{(1)}$ onto the space X_{ij}. The measure $\gamma^{(1)}$ can be written as follows:

$$\begin{aligned}\gamma^{(1)} &= \mu_i \otimes \mathrm{Pr}_{jk}(\mathrm{Up}_{ij}(\nu_i \otimes \nu_j, \pi)) + \mu_j \otimes \mathrm{Pr}_{ik}(\mathrm{Up}_{ij}(\nu_i \otimes \nu_j, \pi))\\ &+ \mu_i \otimes \mathrm{Pr}_{jk}(\mathrm{Up}_{ik}(\nu_i \otimes \nu_k, \pi)) + \mu_k \otimes \mathrm{Pr}_{ij}(\mathrm{Up}_{ik}(\nu_i \otimes \nu_k, \pi))\\ &+ \mu_j \otimes \mathrm{Pr}_{ik}(\mathrm{Up}_{jk}(\nu_j \otimes \nu_k, \pi)) + \mu_k \otimes \mathrm{Pr}_{ij}(\mathrm{Up}_{jk}(\nu_j \otimes \nu_k, \pi)).\end{aligned}$$

It follows from assertion 6.5(d) that

$$\begin{aligned}\mathrm{Pr}_{ij}(\mu_i \otimes \mathrm{Pr}_{jk}(\mathrm{Up}_{ij}(\nu_i \otimes \nu_j, \pi))) &= \mu_i \otimes \mathrm{Pr}_j(\mathrm{Up}_{ij}(\nu_i \otimes \nu_j, \pi)) = \mu_i \otimes \nu_j,\\ \mathrm{Pr}_{ij}(\mu_j \otimes \mathrm{Pr}_{ik}(\mathrm{Up}_{ij}(\nu_i \otimes \nu_j, \pi))) &= \mathrm{Pr}_i(\mathrm{Up}_{ij}(\nu_i \otimes \nu_j, \pi)) \otimes \mu_j = \nu_i \otimes \mu_j,\end{aligned}$$

and we trivially have

$$\begin{aligned}\mathrm{Pr}_{ij}(\mu_i \otimes \mathrm{Pr}_{jk}(\mathrm{Up}_{ik}(\nu_i \otimes \nu_k, \pi))) &= \mu_i \otimes \mathrm{Pr}_j(\mathrm{Up}_{ik}(\nu_i \otimes \nu_k, \pi)),\\ \mathrm{Pr}_{ij}(\mu_k \otimes \mathrm{Pr}_{ij}(\mathrm{Up}_{ik}(\nu_i \otimes \nu_k, \pi))) &= \mathrm{Pr}_{ij}(\mathrm{Up}_{ik}(\nu_i \otimes \nu_k, \pi)),\\ \mathrm{Pr}_{ij}(\mu_j \otimes \mathrm{Pr}_{ik}(\mathrm{Up}_{jk}(\nu_j \otimes \nu_k, \pi))) &= \mathrm{Pr}_i(\mathrm{Up}_{jk}(\nu_j \otimes \nu_k, \pi)) \otimes \mu_j,\\ \mathrm{Pr}_{ij}(\mu_k \otimes \mathrm{Pr}_{ij}(\mathrm{Up}_{jk}(\nu_j \otimes \nu_k, \pi))) &= \mathrm{Pr}_{ij}(\mathrm{Up}_{jk}(\nu_j \otimes \nu_k, \pi)).\end{aligned}$$

Thus, we get

$$\begin{aligned}\mathrm{Pr}_{ij}(\gamma^{(1)}) &= \mu_i \otimes \nu_j + \nu_i \otimes \mu_j + \mathrm{Pr}_{ij}(\mathrm{Up}_{ik}(\nu_i \otimes \nu_k, \pi)) + \mathrm{Pr}_{ij}(\mathrm{Up}_{jk}(\nu_j \otimes \nu_k, \pi))\\ &+ \mathrm{Pr}_i(\mathrm{Up}_{jk}(\nu_j \otimes \nu_k, \pi)) \otimes \mu_j + \mu_i \otimes \mathrm{Pr}_j(\mathrm{Up}_{ik}(\nu_i \otimes \nu_k, \pi))\end{aligned} \tag{29}$$

Finally, by construction

$$\gamma^{(3)} = \nu_i \otimes \mu_j \otimes \mu_k + \mu_i \otimes \nu_j \otimes \mu_k + \mu_i \otimes \mu_j \otimes \nu_k,$$

so we get

$$\mathrm{Pr}_{ij}(\gamma^{(3)}) = \nu_i \otimes \mu_j + \mu_i \otimes \nu_j + \mu_i \otimes \mu_j. \tag{30}$$

Similarly, we conclude that

$$\mathrm{Pr}_{ij}(\gamma^{(2)}) = \mathrm{Pr}_i(\mathrm{Up}_{jk}(\nu_j \otimes \nu_k, \pi)) \otimes \mu_j + \mu_i \otimes \mathrm{Pr}_j(\mathrm{Up}_{ik}(\nu_i \otimes \nu_k, \pi)) + \mu_i \otimes \mu_j. \tag{31}$$

Thus, from equations (28)-(31) we get

$$\begin{aligned}
\mathrm{Pr}_{ij}(\gamma) &= \mathrm{Pr}_{ij}(\nu_i \otimes \nu_j \otimes \nu_k) - \mathrm{Pr}_{ij}(\gamma^{(0)}) + \mathrm{Pr}_{ij}(\gamma^{(1)}) - \mathrm{Pr}_{ij}(\gamma^{(2)}) - \mathrm{Pr}_{ij}(\gamma^{(3)}) + 2\mathrm{Pr}_{ij}(\pi) \\
&= \nu_i \otimes \nu_j - \nu_i \otimes \nu_j - \mathrm{Pr}_{ij}(\mathrm{Up}_{ik}(\nu_i \otimes \nu_k, \pi)) - \mathrm{Pr}_{ij}(\mathrm{Up}_{jk}(\nu_j \otimes \nu_k, \pi)) \\
&+ \mu_i \otimes \nu_j + \nu_i \otimes \mu_j + \mathrm{Pr}_{ij}(\mathrm{Up}_{ik}(\nu_i \otimes \nu_k, \pi)) + \mathrm{Pr}_{ij}(\mathrm{Up}_{jk}(\nu_j \otimes \nu_k, \pi)) \\
&+ \mathrm{Pr}_i(\mathrm{Up}_{jk}(\nu_j \otimes \nu_k, \pi)) \otimes \mu_j + \mu_i \otimes \mathrm{Pr}_j(\mathrm{Up}_{ik}(\nu_i \otimes \nu_k, \pi)) \\
&- \mathrm{Pr}_i(\mathrm{Up}_{jk}(\nu_j \otimes \nu_k, \pi)) \otimes \mu_j - \mu_i \otimes \mathrm{Pr}_j(\mathrm{Up}_{ik}(\nu_i \otimes \nu_k, \pi)) - \mu_i \otimes \mu_j \\
&- \nu_i \otimes \mu_j - \mu_i \otimes \nu_j - \mu_i \otimes \mu_j + 2\mu_i \otimes \mu_j \\
&= 0.
\end{aligned}$$

Let us verify that the functions F and the extensions of f_{ij} to the space X for all $\{i,j\} \in \mathcal{I}_{3,2}$ are integrable with respect to γ. First, since $\nu_t \ll_b \mu_t$ for $1 \le t \le 3$, we have

$$\nu_1 \otimes \nu_2 \otimes \nu_3 \ll_b \mu_1 \otimes \mu_2 \otimes \mu_3 \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}).$$

Let (i,j,k) be a permutation of indices $(1,2,3)$. It follows from assertion 6.5(b) that $\mathrm{Up}_{ij}(\nu_i \otimes \nu_j, \pi) \ll_b \pi$, and therefore $\gamma^{(0)} \ll_b \pi$. Next, since $\mathrm{Up}_{ij}(\nu_i \otimes \nu_j, \pi) \ll_b \pi$, it follows from assertion 6.3(c) that $\mathrm{Pr}_{jk}(\mathrm{Up}_{ij}(\nu_i \otimes \nu_j, \pi)) \ll_b \mathrm{Pr}_{jk}(\pi) = \mu_j \otimes \mu_k$ and $\mathrm{Pr}_k(\mathrm{Up}_{ij}(\nu_i \otimes \nu_j, \pi)) \ll_b \mathrm{Pr}_k(\pi) = \mu_k$. Hence, $\mu_i \otimes \mathrm{Pr}_{jk}(\mathrm{Up}_{ij}(\nu_i \otimes \nu_j, \pi)) \ll_b \mu_1 \otimes \mu_2 \otimes \mu_3$ and $\mu_i \otimes \mu_j \otimes \mathrm{Pr}_k(\mathrm{Up}_{ij}(\nu_i \otimes \nu_j, \pi)) \ll_b \mu_1 \otimes \mu_2 \otimes \mu_3$, and therefore

$$\gamma^{(1)} \ll_b \mu_1 \otimes \mu_2 \otimes \mu_3 \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}) \ \text{ and } \ \gamma^{(2)} \ll_b \mu_1 \otimes \mu_2 \otimes \mu_3 \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}).$$

Finally, since $\nu_k \ll_b \mu_k$, we have $\mu_i \otimes \mu_j \otimes \nu_k \ll_b \mu_1 \otimes \mu_2 \otimes \mu_3$, and therefore

$$\gamma^{(3)} \ll_b \mu_1 \otimes \mu_2 \otimes \mu_3 \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}).$$

Thus, by assertion 6.7(a) the function F and the extension of f_{ij} to the space X for all $\{i,j\} \in \mathcal{I}_{3,2}$ are integrable with respect to all summands from the definition of γ, and therefore that functions are integrable with respect to γ. In particular,

$$\int\limits_X F\,d\gamma = \int\limits_{X_{12}} f_{12}\,d\mathrm{Pr}_{12}(\gamma) + \int\limits_{X_{13}} f_{13}\,d\mathrm{Pr}_{13}(\gamma) + \int\limits_{X_{23}} f_{23}\,d\mathrm{Pr}_{23}(\gamma) = 0.$$

Since $\mathrm{Up}_{ij}(\nu_i \otimes \nu_j, \pi) \ll_b \pi$ for all $\{i,j\} \in \mathcal{I}_{3,2}$, it follows from assertion 6.7(c) that

$$\int\limits_X F\,d\gamma^{(0)} \ge 0.$$

Applying assertion 6.7(b) to all terms of the definition of $\gamma^{(1)}$, we conclude that

$$\int\limits_X F\,d\gamma^{(1)} \le 6\,\|c\|_\infty\,.$$

Finally, applying Lemma 6.8 to all terms of $\gamma^{(2)}$ and $\gamma^{(3)}$, we get

$$\int_X F\,d\gamma^{(2)} \geq 3\int_X F\,d\pi - 3\,\|c\|_\infty \quad\text{and}\quad \int_X F\,d\gamma^{(3)} \geq 3\int_X F\,d\pi - 3\,\|c\|_\infty\,.$$

Thus, we get the following inequality:

$$\begin{aligned}\int_X F\,d\gamma &\leq \int_X F\,d(\nu_1\otimes\nu_2\otimes\nu_3) + 6\,\|c\|_\infty + 2\left(3\,\|c\|_\infty - 3\int_X F\,d\pi\right) + 2\int_X F\,d\pi \\ &= 12\,\|c\|_\infty - 4\int_X F\,d\pi + \int_X F\,d(\nu_1\otimes\nu_2\otimes\nu_3),\end{aligned}$$

and therefore

$$\int_X F\,d(\nu_1\otimes\nu_2\otimes\nu_3) \geq 4\int_X F\,d\pi - 12\,\|c\|_\infty\,.$$

It follows from assertion 6.7(c) that $\int_X F\,d\pi \geq 0$; hence,

$$\int_X F\,d(\nu_1\otimes\nu_2\otimes\nu_3) \geq 4\int_X F\,d\pi - 12\,\|c\|_\infty \geq -12\,\|c\|_\infty\,. \quad \square$$

Theorem 6.10. *Let X_1, X_2, X_3 be Polish spaces, let $\mu_i \in \mathcal{P}(X_i)$ for $1 \leq i \leq 3$, and let $\mu_{ij} = \mu_i \otimes \mu_j$ for all $\{i,j\} \in \mathcal{I}_{3,2}$. Let $c\colon X \to \mathbb{R}_+$ be a bounded continuous cost function. If $\{f_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$ is a solution to the related dual problem, then*

$$f_{12}(x_1,x_2) + f_{13}(x_1,x_3) + f_{23}(x_2,x_3) \geq -12\,\|c\|_\infty$$

for $\mu_1\otimes\mu_2\otimes\mu_3$-a.e. points $x \in X$.

Proof. Denote $F(x_1,x_2,x_3) = f_{12}(x_1,x_2) + f_{13}(x_1,x_3) + f_{23}(x_2,x_3)$, and denote $\mu = \mu_1\otimes\mu_2\otimes\mu_3$. For $1 \leq i \leq 3$, let $A_i \in \mathcal{B}_i$ be a measurable subset of X_i. If $\mu_i(A_i) = 0$ for some $1 \leq i \leq 3$, then $\mu(A_1\times A_2\times A_3) = 0$, and therefore $\int_{A_1\times A_2\times A_3} F\,d\mu = 0$.

Suppose otherwise that $\mu_i(A_i) > 0$ for all $1 \leq i \leq 3$. Denote $\nu_i = (\mathbb{1}[A_i]/\mu_i(A_i))\cdot\mu_i$, where $\mathbb{1}[A]$ is an indicator function of the set A. The measure ν_i is a probability measure and $\nu_i \leq (1/\mu_i(A_i))\cdot\mu_i$, and therefore $\nu_i \ll_b \mu_i$. By Lemma 6.9, we conclude that $\int_X F\,d(\nu_1\otimes\nu_2\otimes\nu_3) \geq -12\,\|c\|_\infty$. By construction,

$$\int_X F\,d(\nu_1\otimes\nu_2\otimes\nu_3) = \frac{\int_{A_1\times A_2\times A_3} F\,d\mu}{\mu_1(A_1)\mu_2(A_2)\mu_3(A_3)}.$$

Thus, we get

$$\int_{A_1\times A_2\times A_3} F\,d\mu \geq -12\,\|c\|_\infty\cdot\mu(A_1\times A_2\times A_3) \ \text{ for all } A_i \in \mathcal{B}_i. \tag{32}$$

Consider the measure $(F + 12\,\|c\|_\infty)\cdot\mu$. By equation (32) this measure is non-negative on a semialgebra $\mathcal{A}_0 = \{A_1\times A_2\times A_3\colon A_i \in \mathcal{B}_i\}$, and therefore this measure is non-negative on every element of $\sigma(\mathcal{A}_0)$, and this σ-algebra coincides with the Borel σ-algebra on the space X. Thus, the measure $(F + 12\,\|c\|_\infty)\cdot\mu$ is non-negative, and therefore $F(x_1,x_2,x_3) + 12\,\|c\|_\infty \geq 0$ for μ-a.e. points $x \in X$. $\square$

Theorem 6.11. *Let X_1, X_2, X_3 be Polish spaces, let $\mu_i \in \mathcal{P}(X_i)$ for $1 \le i \le 3$, and let $\mu_{ij} = \mu_i \otimes \mu_j$ for all $\{i,j\} \in \mathcal{I}_{3,2}$. Let $c\colon X \to \mathbb{R}_+$ be a bounded continuous cost function. Then*

(a) *there exists a solution $\{f_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$ to the relaxed dual problem (see Definition 5.16) such that*

$$-17\,\|c\|_\infty \le f_{ij}(x_i,x_j) \le 13\frac{1}{3}\,\|c\|_\infty\,;$$

(b) *there exists a solution $\{f_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$ to the standard dual problem such that*

$$-26\frac{2}{3}\,\|c\|_\infty \le f_{ij}(x_i,x_j) \le 13\frac{1}{3}\,\|c\|_\infty\,.$$

Proof. First, it follows from Theorem 5.19 that there exists an optimal (real-valued) solution $\{f_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$ to the relaxed dual problem. By Theorem 6.10 we conclude that the inequality

$$\|c\|_\infty \ge f_{12}(x_1,x_2) + f_{13}(x_1,x_3) + f_{23}(x_2,x_3) \ge -12\,\|c\|_\infty \tag{33}$$

holds for $\mu_1\otimes\mu_2\otimes\mu_3$-almost all points.

Consider the optimal finite $(3,2)$-function $F(x_1,x_2,x_3) = f_{12}(x_1,x_2) + f_{13}(x_1,x_3) + f_{23}(x_2,x_3)$. Let A be the set of points $(x_1,x_2,x_3) \in X$ such that either $F(x_1,x_2,x_3) < -12\,\|c\|_\infty$ or $F(x_1,x_2,x_3) > \|c\|_\infty$. By inequality (33) we have $\mu_1\otimes\mu_2\otimes\mu_3(A) = 0$. Applying Lemma 5.5 to the indicator function of the set A, we conclude that there exists a point $(y_1,y_2,y_3) \in X$ such that for each $\alpha \in \mathcal{I}_3$ the set $A_\alpha = \{x_\alpha \in X_\alpha\colon (x_\alpha, y_{\{1,2,3\}\setminus\alpha}) \in A\}$ have a zero measure with respect to μ_α.

For each $\alpha \in \mathcal{I}_3$ consider the function $F_\alpha\colon x_\alpha \mapsto F(x_\alpha y_{\{1,2,3\}\setminus\alpha})$. If $x_\alpha \notin A_\alpha$, then $\|c\|_\infty \ge F_\alpha(x_\alpha) \ge -12\,\|c\|_\infty$, and therefore this inequality holds for μ_α-almost all $x_\alpha \in X_\alpha$. Consider the functions

$$\begin{aligned}
\widehat{f}_{12}(x_1,x_2) &= F(x_1,x_2,y_3) - \frac{1}{2}F(x_1,y_2,y_3) - \frac{1}{2}F(y_1,x_2,y_3) + \frac{1}{3}F(y_1,y_2,y_3),\\
\widehat{f}_{13}(x_1,x_3) &= F(x_1,y_2,x_3) - \frac{1}{2}F(x_1,y_2,y_3) - \frac{1}{2}F(y_1,y_2,x_3) + \frac{1}{3}F(y_1,y_2,y_3),\\
\widehat{f}_{23}(x_2,x_3) &= F(y_1,x_2,x_3) - \frac{1}{2}F(y_1,x_2,y_3) - \frac{1}{2}F(y_1,y_2,x_3) + \frac{1}{3}F(y_1,y_2,y_3).
\end{aligned}$$

By Example 5.4 the equation $F(x_1,x_2,x_3) = \widehat{f}_{12}(x_1,x_2) + \widehat{f}_{13}(x_1,x_3) + \widehat{f}_{23}(x_2,x_3)$ holds for all $(x_1,x_2,x_3) \in X$. In addition, one can easily verify that the inequality

$$-17\,\|c\|_\infty \le \widehat{f}_{ij}(x_i,x_j) \le 13\frac{1}{3}\,\|c\|_\infty$$

holds for μ_{ij}-almost all $(x_i,x_j) \in X_{ij}$.

Thus, there exists a triple of bounded measurable functions $\{g_{ij}\}$ such that $g_{ij} = \widehat{f}_{ij}$ almost everywhere and

$$-17\,\|c\|_\infty \le g_{ij}(x_i,x_j) \le 13\frac{1}{3}\,\|c\|_\infty$$

for all $(x_i,x_j) \in X_{ij}$. The inequality

$$g_{12}(x_1,x_2) + g_{13}(x_1,x_3) + g_{23}(x_2,x_3) = F(x_1,x_2,x_3) \le c(x_1,x_2,x_3)$$

holds at all points except a zero $(3,2)$-thickness set, and therefore $\{\widehat{g}_{ij}\} \in \Psi_c(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$. Finally, we have

$$\int_{X_{12}} g_{12}\, d\mu_{12} + \int_{X_{13}} g_{13}\, d\mu_{13} + \int_{X_{23}} g_{23}\, d\mu_{23} = \int_X F\, d\mu$$
$$= \int_{X_{12}} f_{12}\, d\mu_{12} + \int_{X_{13}} f_{13}\, d\mu_{13} + \int_{X_{23}} f_{23}\, d\mu_{23},$$

and therefore $\{g_{ij}\}$ is a solution to the relaxed dual problem satisfying assertion 6.11(a).

Since $\{g_{ij}\} \in \Psi_c(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$, there exists a triple of subsets $Y_{ij} \subset X_{ij}$ such that $\mu_{ij}(Y_{ij}) = 0$ and if $(x_i, x_j) \notin Y_{ij}$ for all $\{i,j\}$, then

$$g_{12}(x_1, x_2) + g_{13}(x_1, x_3) + g_{23}(x_2, x_3) \le c(x_1, x_2, x_3).$$

Consider the triple of functions $\{\widehat{g}_{ij}\}$: $\widehat{g}_{ij}(x_i, x_j) = g(x_i, x_j)$ if $(x_i, x_j) \notin Y_{ij}$, and $\widehat{g}_{ij}(x_i, x_j) = -26\frac{2}{3}\|c\|_\infty$ otherwise. We have $\widehat{g}_{ij}(x_i, x_j) = g_{ij}(x_i, x_j)$ almost everywhere, and one can easily verify that the inequality

$$\widehat{g}_{12}(x_1, x_2) + \widehat{g}_{13}(x_1, x_3) + \widehat{g}_{23}(x_2, x_3) \le c(x_1, x_2, x_3)$$

holds for all points $(x_1, x_2, x_3) \in X$. Thus, $\{\widehat{g}_{ij}\}$ is a solution to the standard dual problem satisfying assertion 6.11(b). □

6.2. Uniqueness of a continuous dual solution for the cost function $x_1x_2x_3$

Let us recall to the reader our main example of the multistochastic $(3, 2)$-problem:

Problem 6.12. *For $1 \le i \le 3$, let $X_i = [0, 1]$, let μ_{ij} be the restriction of the Lebesgue measure to the square $[0, 1]^2$, and let $c(x_1, x_2, x_3) = x_1x_2x_3$.*

Primal problem. *Find a uniting measure $\pi \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$ such that*

$$\int x_1x_2x_3\, d\pi \to \min .$$

Dual problem. *Find a triple of functions $\{f_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$, $f_{ij} \subset L^1([0, 1]^2)$ such that*

$$\sum_{\{i,j\}\in\mathcal{I}_{3,2}} f_{ij}(x_i, x_j) \le x_1x_2x_3 \text{ for all } (x_1, x_2, x_3) \in [0, 1]^3,$$

$$\sum_{\{i,j\}\in\mathcal{I}_{3,2}} \int_0^1 \int_0^1 f_{ij}(x_i, x_j)\, dx_i dx_j \to \max .$$

In [19] the authors describe solutions to these problems. First, we define a binary operator $\oplus$ (called "bitwise exclusive or" or just "xor") on the segment $[0, 1]$.

Definition 6.13. Given x and y on $[0, 1]$, we consider their binary representations $x = \overline{0{,}x_1x_2x_3\ldots}_2$, $y = \overline{0{,}y_1y_2y_3\ldots}_2$. We agree that every dyadic rational number less than 1 has a finite numbers of units in its decomposition. The number 1 will be always decomposed as follows: $1 = \overline{0{,}111\ldots}_2$. Then we define $x \oplus y = \overline{0{,}x_1 \oplus y_1\, x_2 \oplus y_2 \ldots}_2$, where $\oplus$ is an addition in $\mathbb{F}_2$.

Using this binary operation, the solutions to the primal problem can be described as follows:

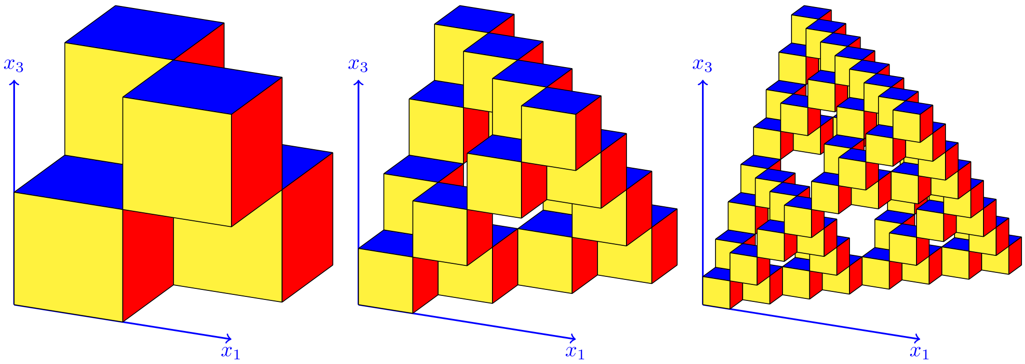

Fig. 4. The sets J_1, J_2 and J_3.

Theorem 6.14 *(Primal problem solution). Consider the mapping* $T\colon [0,1]^2 \to [0,1]^3$, $(x,y) \mapsto (x,y,x \oplus y)$. *Denote by* π *the image of the Lebesgue measure restricted to the square* $[0,1]^2$ *under the mapping* T*. Then* π *is a solution to primal Problem 6.12.*

In [19] the authors show that π is concentrated on the set

$$\{(x,y,z) \in [0,1]^3 \colon x \oplus y \oplus z = 0\},$$

and this set is a self-similar fractal, which is called "Sierpińsky tetrahedron". Let us verify for the completeness of the picture the following description of the support of π.

Definition 6.15. Denote by $J_n^{a_1,a_2,a_3}$ the image of $[0,1]^3$ under the mapping

$$(x_1,x_2,x_3) \mapsto \left(\frac{a_1+x_1}{2^n}, \frac{a_2+x_2}{2^n}, \frac{a_3+x_3}{2^n}\right).$$

Let

$$J_n = \bigcup_{\substack{0 \le a_i < 2^n \\ a_1 \oplus a_2 \oplus a_3 = 0}} J_n^{a_1,a_2,a_3}.$$

One can find images of J_1, J_2 and J_3 on Fig. 4. Denote

$$\mathfrak{S} = \bigcap_{n \ge 1} J_n.$$

The set $\mathfrak{S}$ is called *Sierpińsky tetrahedron.*

Lemma 6.16. *The set* J_n *contains a point* (x_1,x_2,x_3) *if and only if there exist binary representations of each coordinates* $x_i = \sum_{k=1}^{\infty} x_{i,k}/2^k$ *such that* $x_{1,k} \oplus x_{2,k} \oplus x_{3,k} = 0$ *for all* $1 \le k \le n$.

Proof. First, suppose that $(x_1,x_2,x_3) \in J_n$. By construction, there exist integers a_1,a_2,a_3 such that $0 \le a_i < 2^n$, bitwise xor of a_1,a_2 and a_3 is zero, and $(x_1,x_2,x_3) \in J_n^{a_1,a_2,a_3}$. Since

$$J_n^{a_1,a_2,a_3} = \left[\frac{a_1}{2^n}, \frac{a_1+1}{2^n}\right] \times \left[\frac{a_2}{2^n}, \frac{a_2+1}{2^n}\right] \times \left[\frac{a_3}{2^n}, \frac{a_3+1}{2^n}\right],$$

we conclude that $x_i = (a_i + y_i)/2^n$ for all $1 \le i \le 3$, where $0 \le y_i \le 1$.

Since $a_i < 2^n$, the binary representation of a_i contains at most n digits. Let $\overline{a_{i,1}a_{i,2}\dots a_{i,n}}_2$ be the binary representation of a_i supplemented by zeros up to length n. Since $a_1 \oplus a_2 \oplus a_3 = 0$, we have $a_{1,k} \oplus a_{2,k} \oplus a_{3,k} = 0$ for all $1 \leq k \leq n$. Hence, if $y_i = \sum_{k=1}^{\infty} y_{i,k}/2^k$, then

$$x_i = \sum_{k=1}^{n} \frac{a_{i,k}}{2^k} + \sum_{k=n+1}^{\infty} \frac{y_{i,k-n}}{2^k}$$

provided by $x_i = (a_i + y_i)/2^n$. This equation provides a binary representation of each coordinates $x_i = \sum_{k=1}^{\infty} x_{i,k}/2^k$ such that $x_{1,k} \oplus x_{2,k} \oplus x_{3,k} = 0$ for all $1 \leq k \leq n$.

Suppose that (x_1, x_2, x_3) is a point on $[0,1]^3$ and $x_i = \sum_{k=1}^{\infty} x_{i,k}/2^k$ for $1 \leq i \leq 3$, where all $x_{i,k}$ are 0 or 1, and $x_{1,k} \oplus x_{2,k} \oplus x_{3,k} = 0$ for all $1 \leq k \leq n$. Denote by a_i an integer formed by the first n digits of x_i after radix point. We have $x_i = (a_i + y_i)/2^n$ for $1 \leq i \leq 3$, where $0 \leq y_i \leq 1$, and therefore $(x_1, x_2, x_3) \in J_n^{a_1,a_2,a_3}$. In addition, $0 \leq a_i < 2^n$, and since $x_{1,k} \oplus x_{2,k} \oplus x_{3,k} = 0$ for all $1 \leq k \leq n$, we conclude that $a_1 \oplus a_2 \oplus a_3 = 0$. Thus, $(x_1, x_2, x_3) \in J_n^{a_1,a_2,a_3} \subset J_n$. □

Using that, we can describe all points of the Sierpińsky tetrahedron in terms of their binary representations.

Proposition 6.17. *The Sierpińsky tetrahedron $\mathfrak{S}$ contains a point (x_1, x_2, x_3) if and only if there exist binary representations of each coordinates $x_i = \sum_{k=1}^{\infty} x_{i,k}/2^k$ such that*

$$x_{1,k} \oplus x_{2,k} \oplus x_{3,k} = 0 \text{ for all } k;$$

Proof. Suppose that (x_1, x_2, x_3) is a point on $[0,1]^3$ and $x_i = \sum_{k=1}^{\infty} x_{i,k}/2^k$ for $1 \leq i \leq 3$, where all $x_{i,k}$ are 0 or 1, and $x_{1,k} \oplus x_{2,k} \oplus x_{3,k} = 0$ for all k. Then it follows from Lemma 6.16 that (x_1, x_2, x_3) is contained in J_n for all n. Thus,

$$(x_1, x_2, x_3) \in \bigcap_{n \geq 1} J_n = \mathfrak{S}.$$

Suppose that $(x_1, x_2, x_3) \in \mathfrak{S}$. Then $(x_1, x_2, x_3) \in J_n$ for all n, and therefore there exist binary representations of each coordinates $x_i = \sum_{k=1}^{\infty} x_{i,k}^n/2^k$ such that $x_{1,k}^n \oplus x_{2,k}^n \oplus x_{2,k}^n = 0$ for all $1 \leq k \leq n$. For any nonnegative real number, there are at most two binary representations of this number, and therefore there exist at most eight triples of binary representations of the point (x_1, x_2, x_3). Hence, there exists at least one of them $x_i = \sum_{k=1}^{\infty} x_{i,k}/2^k$ such that the property $x_{1,k} \oplus x_{2,k} \oplus x_{3,k} = 0$ for all $1 \leq k \leq n$ holds for an infinite number of n. Thus, $x_{1,k} \oplus x_{2,k} \oplus x_{3,k} = 0$ for all k. □

Proposition 6.18. *The Sierpińsky tetrahedron $\mathfrak{S}$ has the following properties:*

(a) *the set $\mathfrak{S}$ a closed subset of $[0,1]^3$;*
(b) *a point $(x, y, x \oplus y)$ is contained in S for all $x, y \in [0,1]$;*
(c) *if $\mathfrak{S}_n^{a_1,a_2,a_3}$ is the image of $\mathfrak{S}$ under a mapping*

$$(x_1, x_2, x_3) \mapsto \left(\frac{a_1 + x_1}{2^n}, \frac{a_2 + x_2}{2^n}, \frac{a_3 + x_3}{2^n} \right),$$

then

$$\mathfrak{S} = \bigcup_{\substack{0 \leq a_i < 2^n, \\ a_1 \oplus a_2 \oplus a_3 = 0}} \mathfrak{S}_n^{a_1,a_2,a_3}$$

Proof. The set J_n is closed since J_n is a finite union of closed sets. Thus, since $\mathfrak{S}$ is an intersection of the closed sets J_n, we conclude that $\mathfrak{S}$ is closed too, and this implies assertion 6.18(a).

Assertion 6.18(b) trivially holds by Proposition 6.17.

Let us verify assertion 6.18(c). Suppose that $(x_1, x_2, x_3) \in \mathfrak{S}$. By Proposition 6.17, there exist binary representations $x_i = \sum_{k=1}^{n} x_{i,k}/2^k$ such that $x_{1,k} \oplus x_{2,k} \oplus x_{3,k} = 0$ for all k. Denote by a_i an integer formed by the first n digits of x_i after radix point. We have $0 \le a_i < 2^n$, and since $x_{1,k} \oplus x_{2,k} \oplus x_{3,k} = 0$ for all k, we conclude that $a_1 \oplus a_2 \oplus a_3 = 0$. In addition, $x_i = (a_i + y_i)/2^n$, where $y_i = \sum_{k=1}^{n} x_{i,n+k}/2^k$. By Proposition 6.17 $(y_1, y_2, y_3) \in \mathfrak{S}$, and therefore $(x_1, x_2, x_3) \in \mathfrak{S}_n^{a_1,a_2,a_3}$. Thus

$$\mathfrak{S} \subseteq \bigcup_{\substack{0 \le a_i < 2^n, \\ a_1 \oplus a_2 \oplus a_3 = 0}} \mathfrak{S}_n^{a_1,a_2,a_3}$$

Suppose that $(x_1, x_2, x_3) \in \mathfrak{S}_n^{a_1,a_2,a_3}$, where $0 \le a_i < 2^n$ and $a_1 \oplus a_2 \oplus a_3 = 0$. Since $a_i < 2^n$, the binary representation of a_i contains at most n digits. Let $\overline{a_{i,1}a_{i,2} \dots a_{i,n_2}}$ be the binary representation of a_i supplemented by zeros up to length n. Since $a_1 \oplus a_2 \oplus a_3 = 0$, we have $a_{1,k} \oplus a_{2,k} \oplus a_{3,k} = 0$ for all $1 \le k \le n$.

By construction, there exists a point $(y_1, y_2, y_3) \in \mathfrak{S}$ such that $x_i = (a_i + y_i)/2^n$. By Proposition 6.17, there exist binary representations $y_i = \sum_{k=1}^{\infty} y_{i,k}/2^k$ such that $y_{1,k} \oplus y_{2,k} \oplus y_{3,k} = 0$. Hence,

$$x_i = \frac{a_i + y_i}{2^n} = \sum_{k=1}^{n} \frac{a_{i,k}}{2^k} + \sum_{k=n+1}^{\infty} \frac{y_{i,k-n}}{2^k},$$

and therefore by Proposition 6.17 $(x_1, x_2, x_3) \in \mathfrak{S}$. Thus,

$$\mathfrak{S} \supseteq \bigcup_{\substack{0 \le a_i < 2^n, \\ a_1 \oplus a_2 \oplus a_3 = 0}} \mathfrak{S}_n^{a_1,a_2,a_3},$$

and this completes the proof of assertion 6.18(c). $\square$

Following the proof of the main result in [19] the reader can extract the following statement:

Theorem 6.19. *For $1 \le i \le 3$, let $X_i = [0, 1]$, let μ_{ij} be the Lebesgue measure restricted to the square $[0, 1]^2$, and let $c(x_1, x_2, x_3) = x_1x_2x_3$. If the measure π is uniting for $\{\mu_{ij}\}_{\{i,j\} \in \mathcal{I}_{3,2}}$ and $\operatorname{supp}(\pi) \nsubseteq J_n$ for some n, then there exists a measure $\widetilde{\pi} \in \Pi(\{\mu_{ij}\}_{\{i,j\} \in \mathcal{I}_{3,2}})$ such that*

$$\int_{[0,1]^3} x_1x_2x_3 \, \widetilde{\pi}(dx_1, dx_2, dx_3) < \int_{[0,1]^3} x_1x_2x_3 \, \pi(dx_1, dx_2, dx_3).$$

If π is a solution to primal Problem 6.12, then it follows from Theorem 6.19 that $\operatorname{supp}(\pi) \subseteq J_n$ for all n. Hence, $\operatorname{supp}(\pi) \subseteq \cap_{n \ge 1} J_n$, and this implies the following proposition.

Proposition 6.20. *If π is a solution to primal Problem 6.12, then $\operatorname{supp}(\pi) \subseteq \mathfrak{S}$, where $\mathfrak{S}$ is the Sierpińsky tetrahedron.*

Using that, let us prove that there exists a unique solution to primal Problem 6.12.

Lemma 6.21. *There exists at most one measure π on $[0, 1]^3$ such that $\operatorname{supp}(\pi) \subseteq \mathfrak{S}$ and $\operatorname{Pr}_{12}(\pi)$ coincides with the Lebesgue measure μ_{12} on the square $[0, 1]^2$.*

Proof. Let $\Gamma = \{(x, y, x \oplus y) : (x, y) \in [0,1]^2\}$. It follows from assertion 6.18(b) that $\Gamma \subseteq \mathfrak{S}$. Consider the set $\mathfrak{S}_b = \mathfrak{S} \backslash \Gamma$, and consider a point $(x_1, x_2, x_3) \in \mathfrak{S}_b$. Suppose that both points x_1 and x_2 are not dyadic rationals. If x is not a dyadic rational, then there exists a unique binary representation of x. Hence, it follows from Proposition 6.17 that there exists at most one $z \in [0,1]$ such that $(x_1, x_2, z) \in \mathfrak{S}$. By assertion 6.18(b) we have $(x_1, x_2, x_1 \oplus x_2) \in \mathfrak{S}$, and therefore $x_3 = x_1 \oplus x_2$. Thus, $(x_1, x_2, x_3) \in \Gamma$, and this contradicts the point selection.

This contradiction proves that if $(x_1, x_2, x_3) \in \mathfrak{S}_b$, then at least one of x_1 and x_2 is a dyadic rational. Hence, $\mu_{12}(\mathrm{Pr}_{12}(\mathfrak{S}_b)) = 0$, and therefore $\pi(\mathfrak{S}_b) = 0$ provided by $\mathrm{Pr}_{12}(\pi) = \mu_{12}$. Thus, since $\mathrm{supp}(\pi) \subseteq \mathfrak{S}$, we get $\pi(\Gamma) = 1$.

Let A be a measurable subset of $[0,1]^3$. Since $\pi(\Gamma) = 1$, we have $\pi(A \backslash \Gamma) = 0$, and therefore

$$\pi(A) = \pi(A \cap \Gamma). \tag{34}$$

Denote $A_\Gamma = A \cap \Gamma$. The set A_Γ is a measurable subset of Γ. Since for each $(x_1, x_2) \in [0,1]^2$ there exists exactly one x_3 such that $(x_1, x_2, x_3) \in \Gamma$, we get

$$A_\Gamma = (\mathrm{Pr}_{12}(A_\Gamma) \times X_3) \cap \Gamma.$$

Applying equation (34) to the set $\mathrm{Pr}_{12}(A_\Gamma) \times X_3$, we get

$$\pi((\mathrm{Pr}_{12}(A_\Gamma) \times X_3) \cap \Gamma) = \pi(\mathrm{Pr}_{12}(A_\Gamma) \times X_3) = \mu_{12}(\mathrm{Pr}_{12}(A_\Gamma))$$

provided by $\mathrm{Pr}_{12}(\pi) = \mu_{12}$. From all equations above we get

$$\pi(A) = \pi(A_\Gamma) = \pi((\mathrm{Pr}_{12}(A_\Gamma) \times X_3) \cap \Gamma) = \mu_{12}(\mathrm{Pr}_{12}(A_\Gamma)).$$

Thus, the measure of the set A with respect to π is independent on π, and therefore there exists at most one measure π such that $\mathrm{supp}(\pi) \subseteq \mathfrak{S}$ and $\mathrm{Pr}_{12}(\pi) = \mu_{12}$. □

Theorem 6.22. *There exists a unique solution π to primal Problem 6.12.*

Proof. If π is a solution to the problem, then $\mathrm{Pr}_{12}(\pi) = \mu_{12}$, and it follows from Proposition 6.20 that $\mathrm{supp}(\pi) \subseteq \mathfrak{S}$. By Lemma 6.21, there exists at most one measure π with that properties. Thus, there exists at most one solution to primal Problem 6.12.

The existence of a solution follows from Theorem 2.8. □

Finally, let us find exactly the support of the solution to primal Problem 6.12.

Proposition 6.23. *If π is the solution to primal Problem 6.12, then* $\mathrm{supp}(\pi) = \mathfrak{S}$.

Proof. It follows from Proposition 6.20 that $\mathrm{supp}(\pi) \subseteq \mathfrak{S} \subset J_n$ for all n, and therefore $\pi(J_n) = 1$. By definition of J_n,

$$J_n = \bigcup_{\substack{0 \le a_i < 2^n, \\ a_1 \oplus a_2 \oplus a_3 = 0}} J_n^{a_1, a_2, a_3}.$$

We have

$$\mathrm{Pr}_{12}(J_n^{a_1, a_2, a_3}) = \left[\frac{a_1}{2^n}, \frac{a_1 + 1}{2^n}\right] \times \left[\frac{a_2}{2^n}, \frac{a_2 + 1}{2^n}\right].$$

For each pair a_1, a_2 such that $0 \le a_1, a_2 < 2^n$ there exists a unique a_3 such that $0 \le a_3 < 2^n$ and $a_1 \oplus a_2 \oplus a_3 = 0$. Hence, projections to $X_1 \times X_2$ of all components of J_n overlapping by the sets of measure zero with respect to μ_{12}, and therefore

$$\pi(J_n^{a_1,a_2,a_3}) = \mu_{12}(\Pr_{12}(J_n^{a_1,a_2,a_3})) = \frac{1}{4^n} \text{ if } a_1 \oplus a_2 \oplus a_3 = 0. \tag{35}$$

Suppose that $\operatorname{supp}(\pi) \neq \mathfrak{S}$. Since $\operatorname{supp}(\pi)$ is closed, there exist a point $x_0 \in \mathfrak{S}$ and a non-negative integer n such that if $|x - x_0| < 2^{1-n}$, then x is not contained in $\operatorname{supp}(\pi)$. Since $x_0 \in \mathfrak{S} \subset J_n$, there exist integers a_1, a_2, a_3 such that $0 \le a_1, a_2, a_3 < 2^n$, bitwise xor of a_1, a_2, a_3 is zero, and $x_0 \in J_n^{a_1,a_2,a_3}$. We have

$$J_n^{a_1,a_2,a_3} = \left[\frac{a_1}{2^n}, \frac{a_1+1}{2^n}\right] \times \left[\frac{a_2}{2^n}, \frac{a_2+1}{2^n}\right] \times \left[\frac{a_3}{2^n}, \frac{a_3+1}{2^n}\right];$$

hence, $\operatorname{diam}(J_n^{a_1,a_2,a_3}) < 2^{1-n}$, and therefore $\operatorname{supp}(\pi) \cap J_n^{a_1,a_2,a_3} = \varnothing$. This contradicts equation (35). $\square$

In [19] the authors also found a solution to the dual Problem 6.12.

Theorem 6.24 *(Dual problem solution). Denote*

$$f(x,y) = \int_0^x \int_0^y s \oplus t \, ds dt - \frac{1}{4} \int_0^x \int_0^x s \oplus t \, ds dt - \frac{1}{4} \int_0^y \int_0^y s \oplus t \, ds dt.$$

Then the triple of functions $f_{ij} \colon (x_i, x_j) \mapsto f(x_i, x_j)$ is a solution to dual Problem 6.12.

This solution to the dual problem is not unique. First, for $1 \le i \le 3$ let f_i be an integrable function on the segment $[0,1]$. Consider the following functions

$$\begin{aligned}
\widehat{f}_{12}(x_1, x_2) &= f_{12}(x_1, x_2) + f_1(x_1) - f_2(x_2), \\
\widehat{f}_{23}(x_2, x_3) &= f_{23}(x_2, x_3) + f_2(x_2) - f_3(x_3), \\
\widehat{f}_{13}(x_1, x_3) &= f_{13}(x_1, x_3) + f_3(x_3) - f_1(x_1).
\end{aligned}$$

Clearly

$$\sum_{\{i,j\} \in \mathcal{I}_{32}} \widehat{f}_{ij}(x_i, x_j) = \sum_{\{i,j\} \in \mathcal{I}_{32}} f_{ij}(x_i, x_j) \text{ for all } (x_1, x_2, x_3) \in [0,1]^3$$

and

$$\sum_{\{i,j\} \in \mathcal{I}_{32}} \int_0^1 \int_0^1 \widehat{f}_{ij}(x_i, x_j) \, dx_i dx_j = \sum_{\{i,j\} \in \mathcal{I}_{32}} \int_0^1 \int_0^1 f_{ij}(x_i, x_j) \, dx_i dx_j,$$

and therefore the functions $\{\widehat{f}_{ij}\}$ are also the solution to the dual problem.

In what follows, we prove that there is no other continuous solutions to the related dual problem.

Lemma 6.25. *If a triple of functions $\{f_{ij}\}_{\{i,j\} \in \mathcal{I}_{3,2}}$ is a solution to dual Problem 6.12, function f_{ij} is continuous for all $\{i,j\} \in \mathcal{I}_{3,2}$, and a_1, a_2 and a_3 are non-negative integers such that $0 \le a_1, a_2, a_3 < 2^n$ and $a_1 \oplus a_2 \oplus a_3 = 0$, then*

$$|F(x_1,x_2,x_3) - x_1x_2x_3| \le \frac{13}{2^{3n}} \textit{ for all } (x_1,x_2,x_3) \in J_n^{a_1,a_2,a_3},$$

where

$$F(x_1,x_2,x_3) = f_{12}(x_1,x_2) + f_{13}(x_1,x_3) + f_{23}(x_2,x_3)$$

and

$$J_n^{a_1,a_2,a_3} = \left[\frac{a_1}{2^n}, \frac{a_1+1}{2^n}\right] \times \left[\frac{a_2}{2^n}, \frac{a_2+1}{2^n}\right] \times \left[\frac{a_3}{2^n}, \frac{a_3+1}{2^n}\right].$$

Proof. Since $\{f_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$ is a solution to the dual problem, we have

$$F(x_1,x_2,x_3) \le x_1x_2x_2 \text{ for all } (x_1,x_2,x_3) \in [0,1]^3. \tag{36}$$

Let π be the solution to primal Problem 6.12. We have $\int F\,d\pi = \int x_1x_2x_2\,d\pi$, and therefore

$$F(x_1,x_2,x_3) = x_1x_2x_3 \text{ for } \pi\text{-a.e. } (x_1,x_2,x_3) \in [0,1]^3.$$

The function $F(x_1,x_2,x_3) - x_1x_2x_3$ is continuous; hence, the equation holds for all $(x_1,x_2,x_3) \in \operatorname{supp}(\pi)$. By Proposition 6.23, the support of π coincides with the Sierpińsky tetrahedron $\mathfrak{S}$, and therefore

$$F(x_1,x_2,x_3) = x_1x_2x_3 \text{ for all } (x_1,x_2,x_3) \in \mathfrak{S}. \tag{37}$$

Consider the following functions:

$$\begin{aligned}
\widehat{f}_{12}(x_1,x_2) &= 2^{3n} f_{12}\left(\frac{a_1+x_1}{2^n}, \frac{a_2+x_2}{2^n}\right) - \frac{a_1a_2a_3}{3} - \frac{a_2a_3x_1 + a_1a_3x_2}{2} - a_3x_1x_2,\\
\widehat{f}_{13}(x_1,x_3) &= 2^{3n} f_{13}\left(\frac{a_1+x_1}{2^n}, \frac{a_3+x_3}{2^n}\right) - \frac{a_1a_2a_3}{3} - \frac{a_2a_3x_1 + a_1a_2x_3}{2} - a_2x_1x_3,\\
\widehat{f}_{23}(x_2,x_3) &= 2^{3n} f_{23}\left(\frac{a_2+x_2}{2^n}, \frac{a_3+x_3}{2^n}\right) - \frac{a_1a_2a_3}{3} - \frac{a_1a_3x_2 + a_1a_2x_3}{2} - a_1x_2x_3,
\end{aligned}$$

where $0 \le x_i \le 1$ for $1 \le i \le 3$. We claim that $\{\widehat{f}_{ij}\}$ is a solution to the dual problem.

First, one can easily verify that

$$\begin{aligned}
&\widehat{f}_{12}(x_1,x_2) + \widehat{f}_{13}(x_1,x_3) + \widehat{f}_{23}(x_2,x_3) =\\
&\qquad 2^{3n}\left[F\left(\frac{a_1+x_1}{2^n}, \frac{a_2+x_2}{2^n}, \frac{a_3+x_3}{2^n}\right) - \frac{a_1+x_1}{2^n}\cdot\frac{a_2+x_2}{2^n}\cdot\frac{a_3+x_3}{2^n}\right] + x_1x_2x_3.
\end{aligned} \tag{38}$$

Using inequality (36), we conclude that

$$\widehat{f}_{12}(x_1,x_2) + \widehat{f}_{13}(x_1,x_3) + \widehat{f}_{23}(x_2,x_3) \le x_1x_2x_3 \text{ for all } (x_1,x_2,x_3) \in [0,1]^3. \tag{39}$$

If $(x_1,x_2,x_3) \in \mathfrak{S}$, then

$$\left(\frac{a_1+x_1}{2^n}, \frac{a_2+x_2}{2^n}, \frac{a_3+x_3}{2^n}\right) \in \mathfrak{S}_n^{a_1,a_2,a_3}.$$

By assertion 6.18(c), $\mathfrak{S}_n^{a_1,a_2,a_3} \subset \mathfrak{S}$; hence, if $(x_1,x_2,x_3) \in \mathfrak{S}$, then by (37) we get

$$F\left(\frac{a_1+x_1}{2^n}, \frac{a_2+x_2}{2^n}, \frac{a_3+x_3}{2^n}\right) = \frac{a_1+x_1}{2^n} \cdot \frac{a_2+x_2}{2^n} \cdot \frac{a_3+x_3}{2^n},$$

and therefore

$$\widehat{f}_{12}(x_1, x_2) + \widehat{f}_{13}(x_1, x_3) + \widehat{f}_{23}(x_2, x_3) = x_1x_2x_3 \text{ for all } (x_1, x_2, x_3) \in \mathfrak{S}.$$

Since $\operatorname{supp}(\pi) = \mathfrak{S}$, we have

$$\int\limits_{[0,1]^2} \widehat{f}_{12}(x_1, x_2)\, dx_1dx_2 + \int\limits_{[0,1]^2} \widehat{f}_{13}(x_1, x_3)\, dx_1dx_3 + \int\limits_{[0,1]^2} \widehat{f}_{23}(x_2, x_3)\, dx_2dx_3$$
$$= \int\limits_{[0,1]^3} \left(\widehat{f}_{12}(x_1, x_2) + \widehat{f}_{13}(x_1, x_3) + \widehat{f}_{23}(x_2, x_3)\right) d\pi = \int\limits_{[0,1]^3} x_1x_2x_3\, d\pi. \quad (40)$$

By equations (39) and (40) we conclude that $\{\widehat{f}_{ij}\}$ is a solution to dual Problem 6.12.

The cost function $x_1x_2x_3$ is non-negative and $\mu_{ij} = \mu_i \otimes \mu_j$ for all $\{i, j\} \in \mathcal{I}_{3,2}$. Thus, we are under assumptions of Theorem 6.10. We have $\|x_1x_2x_3\|_\infty = 1$, where $0 \le x_i \le 1$ for all $1 \le i \le 3$, and therefore

$$-12 \le \widehat{f}_{12}(x_1, x_2) + \widehat{f}_{13}(x_1, x_3) + \widehat{f}_{23}(x_2, x_3) \le x_1x_2x_3 \le 1$$

for almost all $(x_1, x_2, x_3) \in [0, 1]^3$. Since all $\widehat{f}_{ij}$ are continuous, we conclude that inequalities holds for all points, and therefore

$$\left|\widehat{f}_{12}(x_1, x_2) + \widehat{f}_{13}(x_1, x_3) + \widehat{f}_{23}(x_2, x_3)\right| \le 12 \text{ for all } (x_1, x_2, x_3) \in [0, 1]^3.$$

Using equation (38), we conclude that

$$\left|F\left(\frac{a_1+x_1}{2^n}, \frac{a_2+x_2}{2^n}, \frac{a_3+x_3}{2^n}\right) - \frac{a_1+x_1}{2^n} \cdot \frac{a_2+x_2}{2^n} \cdot \frac{a_3+x_3}{2^n}\right| \le \frac{12 + x_1x_2x_3}{2^{3n}} \le \frac{13}{2^{3n}}$$

for all $(x_1, x_2, x_3) \in [0, 1]^3$, and therefore

$$|F(x_1, x_2, x_3) - x_1x_2x_3| \le \frac{13}{2^{3n}} \text{ for all } (x_1, x_2, x_3) \in J_n^{a_1,a_2,a_3}. \quad \square$$

Lemma 6.26. *Let $\{f_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$ be a solution to the dual Problem 6.12. If $\{i, j\} \in \mathcal{I}_{3,2}$, a number n is a positive integer, numbers a_i and a_j are non-negative integers such that $0 \le a_i, a_j < 2^n$, and (x_i, x_j) and (y_i, y_j) are arbitrary points in the square*

$$\left[\frac{a_i}{2^n}, \frac{a_i+1}{2^n}\right] \times \left[\frac{a_j}{2^n}, \frac{a_j+1}{2^n}\right],$$

then

$$\left|f_{ij}(x_i, x_j) - f_{ij}(y_i, x_j) - f_{ij}(x_i, y_j) + f_{ij}(y_i, y_j) - \int\limits_{x_i}^{y_i}\int\limits_{x_j}^{y_j} s \oplus t\, dsdt\right| \le \frac{54}{2^{3n}}.$$

Without loss of generality it can be assumed that $\{i, j\} = \{1, 2\}$. Let $a_3 = a_1 \oplus a_2$, and let (x_1, x_2, x_3) and (y_1, y_2, y_3) be arbitrary points of the cube $J_n^{a_1,a_2,a_3}$. We have

$$F(x_1, x_2, x_3) - F(y_1, x_2, x_3) - F(x_1, y_2, x_3) + F(y_1, y_2, x_3) \\ = f_{12}(x_1, x_2) - f_{12}(y_1, x_2) - f_{12}(x_1, y_2) + f_{12}(y_1, y_2). \quad (41)$$

In addition,

$$x_1x_2x_3 - y_1x_2x_3 - x_1y_2x_3 + y_1y_2x_3 = x_3(x_1 - y_1)(x_2 - y_2). \quad (42)$$

On the other hand, it follows from Lemma 6.25 that

$$|F(x_1, x_2, x_3) + F(y_1, y_2, x_3) - x_1x_2x_3 - y_1y_2x_3 \\ - F(y_1, x_2, x_3) - F(x_1, y_2, x_3) + y_1x_2x_3 + x_1y_2x_3| \leq 4 \cdot \frac{13}{2^{3n}} = \frac{52}{2^{3n}}.$$

Thus, taking into account equations (41) and (42), we get

$$|f_{12}(x_1, x_2) - f_{12}(y_1, x_2) - f_{12}(x_1, y_2) + f_{12}(y_1, y_2) - x_3(x_1 - y_1)(x_2 - y_2)| \leq \frac{52}{2^{3n}}. \quad (43)$$

Since $(x_1, x_2, x_3) \in J_n^{a_1,a_2,a_3}$, we have $|a_3/2^n - x_3| \leq 2^{-n}$. Since $(y_1, y_2, y_3) \in J_n^{a_1,a_2,a_3}$, we also have $|x_1 - y_1| \leq 2^{-n}$ and $|x_2 - y_2| \leq 2^{-n}$. Thus,

$$\left|x_3(x_1 - y_1)(x_2 - y_2) - \frac{a_3}{2^n}(x_1 - y_1)(x_2 - y_2)\right| = \left|\frac{a_3}{2^n} - x_3\right| \cdot |x_1 - y_1| \cdot |x_2 - y_2| \leq \frac{1}{2^{3n}}. \quad (44)$$

Next, let t be a point on the interval $(a_1/2^n, (a_1 + 1)/2^n)$, and let s be a point on an interval $(a_2/2^n, (a_2 + 1)/2^n)$. One can easily verify that

$$\frac{a_1 \oplus a_2}{2^n} \leq s \oplus t \leq \frac{(a_1 \oplus a_2) + 1}{2^n},$$

and therefore, since $a_1 \oplus a_2 = a_3$, we get

$$\left|\int_{x_1}^{y_1}\int_{x_2}^{y_2} s \oplus t\, dsdt - \frac{a_3}{2^n}(x_1 - y_1)(x_2 - y_2)\right| \leq \frac{1}{2^n} \cdot |x_1 - y_1| \cdot |x_2 - y_2| \leq \frac{1}{2^{3n}}. \quad (45)$$

Summarizing inequalities (43), (44), and (45), we conclude that

$$\left|f_{12}(x_1, x_2) - f_{12}(y_1, x_2) - f_{12}(x_1, y_2) + f_{12}(y_1, y_2) - \int_{x_1}^{y_1}\int_{x_2}^{y_2} s \oplus t\, dsdt\right| \leq \frac{54}{2^{3n}}.$$

Lemma 6.27. *If a triple of functions $\{f_{ij}\}_{\{i,j\} \in \mathcal{I}_{3,2}}$ is a solution to dual Problem 6.12 and f_{ij} is continuous for all $\{i, j\} \in \mathcal{I}_{3,2}$, then*

$$f_{ij}(x_i, x_j) - f_{ij}(x_i, 0) - f_{ij}(0, x_j) + f_{ij}(0, 0) = \int_0^{x_i}\int_0^{x_j} s \oplus t\, dsdt$$

for all $(x_i, x_j) \in [0, 1]^2$.

Proof. Let $\{u_k\}_{k=0}^N$ and $\{v_l\}_{l=0}^M$ be arbitrary points on the segment $[0, 1]$. One can easily verify that

$$\begin{aligned}\sum_{k=1}^{N}\sum_{l=1}^{M}(f_{ij}(u_k, v_l) - f_{ij}(u_{k-1}, v_l) - f_{ij}(u_k, v_{l-1}) + f_{ij}(u_{k-1}, v_{l-1})) \\ = f_{ij}(u_N, v_M) - f_{ij}(u_0, v_M) - f_{ij}(u_N, v_0) + f_{ij}(u_0, v_0)\end{aligned} \tag{46}$$

and

$$\sum_{k=1}^{N}\sum_{l=1}^{M}\int_{u_{k-1}}^{u_k}\int_{v_{l-1}}^{v_l} s \oplus t\, dsdt = \int_{u_0}^{u_N}\int_{v_0}^{v_M} s \oplus t\, dsdt. \tag{47}$$

Let (x_i, x_j) be an arbitrary point on the square $[0, 1]^2$. Let $N = \lceil 2^n x_i \rceil$, and let $M = \lceil 2^n x_j \rceil$. Finally, let $u_k = k/2^n$ for all $0 \le k < N$ and $u_N = x_i$, and similarly let $v_l = l/2^n$ for all $0 \le l < M$ and $v_M = x_j$. By construction, both points (u_{k-1}, v_{l-1}) and (u_k, v_l) belong to the square

$$\left[\frac{k-1}{2^n}, \frac{k}{2^n}\right] \times \left[\frac{l-1}{2^n}, \frac{l}{2^n}\right],$$

and therefore by Lemma 6.26 we have

$$\left| f_{ij}(u_k, v_l) - f_{ij}(u_{k-1}, v_l) - f_{ij}(u_k, v_{l-1}) + f_{ij}(u_{k-1}, v_{l-1}) - \int_{u_{k-1}}^{u_k}\int_{v_{l-1}}^{v_l} s \oplus t\, dsdt \right| \le \frac{54}{2^{3n}}$$

for all $1 \le k \le N$ and for all $1 \le l \le M$.

Taking into account equations (46) and (47), we conclude that

$$\begin{aligned}&\left| f_{ij}(x_i, x_j) - f_{ij}(x_i, 0) - f_{ij}(0, x_j) + f_{ij}(0, 0) - \int_0^{x_i}\int_0^{x_j} s \oplus t\, dsdt \right| \\ &\le \sum_{k=1}^{N}\sum_{l=1}^{M}\Big| f_{ij}(u_k, v_l) - f_{ij}(u_{k-1}, v_l) - f_{ij}(u_k, v_{l-1}) \\ &\qquad\qquad + f_{ij}(u_{k-1}, v_{l-1}) - \int_{u_{k-1}}^{u_k}\int_{v_{l-1}}^{v_l} s \oplus t\, dsdt \Big| \\ &\le \sum_{k=1}^{N}\sum_{l=1}^{M}\frac{54}{2^{3n}} = \frac{54 \cdot N \cdot M}{2^{3n}}.\end{aligned}$$

Thus, since $N, M \le 2^n$, we get

$$\left| f_{ij}(x_i, x_j) - f_{ij}(x_i, 0) - f_{ij}(0, x_j) + f_{ij}(0, 0) - \int_0^{x_i}\int_0^{x_j} s \oplus t\, dsdt \right| \le \frac{54}{2^n}$$

for all $(x_i, x_j) \in [0, 1]^2$ and for every positive integer n, and therefore

$$f_{ij}(x_i, x_j) - f_{ij}(x_i, 0) - f_{ij}(0, x_j) + f_{ij}(0, 0) = \int_0^{x_i}\int_0^{x_j} s \oplus t\, dsdt. \quad \square$$

Theorem 6.28. *If a triple of functions $\{f_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$ is a solution to Problem 6.12 and f_{ij} is continuous for all $\{i,j\}\in\mathcal{I}_{3,2}$, then there exist continuous functions $f_i\colon[0,1]\to\mathbb{R}$, $1\le i\le 3$, such that*

$$f_{12}(x_1,x_2)=f(x_1,x_2)+f_1(x_1)-f_2(x_2),$$
$$f_{23}(x_2,x_3)=f(x_2,x_3)+f_2(x_2)-f_3(x_3),$$

and

$$f_{13}(x_1,x_3)=f(x_1,x_3)+f_3(x_3)-f_1(x_1),$$

where

$$f(x,y)=\int_0^x\int_0^y s\oplus t\,dsdt-\frac{1}{4}\int_0^x\int_0^x s\oplus t\,dsdt-\frac{1}{4}\int_0^y\int_0^y s\oplus t\,dsdt.$$

Proof. First, consider the function

$$F(x_1,x_2,x_3)=f_{12}(x_1,x_2)+f_{13}(x_1,x_3)+f_{23}(x_2,x_3).$$

It follows from equation (37) that

$$F(x_1,x_2,x_3)=x_1x_2x_3 \text{ for all } (x_1,x_2,x_3)\in\mathfrak{S}.$$

By assertion 6.18(b), all the points $(0,x,x)$, $(x,0,x)$ and $(x,x,0)$ are contained in $\mathfrak{S}$, and therefore

$$F(0,x,x)=F(x,0,x)=F(x,x,0)=0 \text{ for all } x\in[0,1]. \tag{48}$$

In particular, taking $x=0$, we conclude that

$$f_{12}(0,0)+f_{13}(0,0)+f_{23}(0,0)=F(0,0,0)=0. \tag{49}$$

Denote $\widehat{f}_{ij}(x_i,x_i)=f_{ij}(x_i,x_j)-f_{ij}(0,0)$. We have $\widehat{f}_{ij}(0,0)=0$; it follows from (49) that

$$F(x_1,x_2,x_3)=\widehat{f}_{12}(x_1,x_2)+\widehat{f}_{13}(x_1,x_3)+\widehat{f}_{23}(x_2,x_3).$$

By Lemma 6.27 we have

$$\widehat{f}_{ij}(x_i,x_j)=\int_0^{x_i}\int_0^{x_j} s\oplus t\,dsdt+\widehat{f}_{ij}(x_i,0)+\widehat{f}_{ij}(0,x_j), \tag{50}$$

and therefore

$$\begin{aligned}F(x_1,x_2,x_3)=&\int_0^{x_1}\int_0^{x_2} s\oplus t\,dsdt+\int_0^{x_1}\int_0^{x_3} s\oplus t\,dsdt+\int_0^{x_2}\int_0^{x_3} s\oplus t\,dsdt\\&+\varphi_1(x_1)+\varphi_2(x_2)+\varphi_3(x_3),\end{aligned} \tag{51}$$

where

$$\begin{aligned}\varphi_1(x_1) &= \widehat{f}_{12}(x_1,0)+\widehat{f}_{13}(x_1,0),\\ \varphi_2(x_2) &= \widehat{f}_{12}(0,x_2)+\widehat{f}_{23}(x_2,0),\\ \varphi_3(x_3) &= \widehat{f}_{13}(0,x_3)+\widehat{f}_{23}(0,x_3).\end{aligned} \tag{52}$$

Since $\widehat{f}_{i,j}(0,0)=0$ for all $\{i,j\}\in\mathcal{I}_{3,2}$, we have $\varphi_i(0)=0$ for all $1\le i\le 3$. Hence, using equations (48) and (51) we get

$$\begin{aligned}0 = F(0,x,x) &= \int_0^x\int_0^x s\oplus t\,dsdt+\varphi_2(x)+\varphi_3(x),\\ 0 = F(x,0,x) &= \int_0^x\int_0^x s\oplus t\,dsdt+\varphi_1(x)+\varphi_3(x),\\ 0 = F(x,x,0) &= \int_0^x\int_0^x s\oplus t\,dsdt+\varphi_1(x)+\varphi_2(x)\end{aligned}$$

for all $x\in[0,1]$. Thus, we obtain

$$\varphi_i(x) = -\frac{1}{2}\int_0^x\int_0^x s\oplus t\,dsdt \tag{53}$$

for all $x\in[0,1]$ for $1\le i\le 3$.

Consider the functions $f_i(x_i)$, $1\le i\le 3$, satisfying the following equations:

$$\begin{aligned}\widehat{f}_{12}(x_1,0) &= f_1(x_1)-\frac{1}{4}\int_0^{x_1}\int_0^{x_1} s\oplus t\,dsdt,\\ \widehat{f}_{23}(x_2,0) &= f_2(x_2)-\frac{1}{4}\int_0^{x_2}\int_0^{x_2} s\oplus t\,dsdt,\\ \widehat{f}_{13}(0,x_3) &= f_3(x_3)-\frac{1}{4}\int_0^{x_3}\int_0^{x_3} s\oplus t\,dsdt.\end{aligned} \tag{54}$$

The function f_i is continuous for $1\le i\le 3$. Combining equations (52) and (53) we get

$$\widehat{f}_{12}(0,x_2) = \varphi_2(x_2)-\widehat{f}_{23}(x_2,0) = -\frac{1}{2}\int_0^{x_2}\int_0^{x_2} s\oplus t\,dsd-\widehat{f}_{23}(x_2,0),$$

and using the representation of $\widehat{f}_{23}$ from equation (54) we get

$$\widehat{f}_{12}(0,x_2) = -f_2(x_2)-\frac{1}{4}\int_0^{x_2}\int_0^{x_2} s\oplus t\,dsdt. \tag{55}$$

Substituting equations (54) and (55) into (50) we obtain the following relation:

$$\begin{aligned}\widehat{f}_{12}(x_1,x_2) &= \int_0^{x_1}\int_0^{x_2} s\oplus t\, dsdt + \widehat{f}_{12}(x_1,0) + \widehat{f}_{12}(0,x_2)\\ &= \int_0^{x_1}\int_0^{x_2} s\oplus t\, dsdt - \frac{1}{4}\int_0^{x_1}\int_0^{x_1} s\oplus t\, dsdt - \frac{1}{4}\int_0^{x_2}\int_0^{x_2} s\oplus t\, dsdt + f_1(x_1) - f_2(x_2)\\ &= f(x_1,x_2) + f_1(x_1) - f_2(x_2).\end{aligned}$$

Similarly, we conclude that $\widehat{f}_{23}(x_2,x_3) = f(x_2,x_3) + f_2(x_2) - f_3(x_3)$ and $\widehat{f}_{13}(x_1,x_3) = f(x_1,x_3) + f_3(x_3) - f_1(x_1)$.

Finally, since $f_{12}(0,0) + f_{13}(0,0) + f_{23}(0,0) = 0$, there exist real numbers C_1, C_2 and C_3 such that $f_{12}(0,0) = C_1 - C_2$, $f_{23}(0,0) = C_2 - C_3$ and $f_{13}(0,0) = C_3 - C_1$. Thus,

$$\begin{aligned} f_{12}(x_1,x_2) &= \widehat{f}_{12}(x_1,x_2) + f_{12}(0,0) = f(x_1,x_2) + (f_1(x_1)+C_1) - (f_2(x_2)+C_2),\\ f_{23}(x_2,x_3) &= \widehat{f}_{23}(x_2,x_3) + f_{23}(0,0) = f(x_2,x_3) + (f_2(x_2)+C_2) - (f_3(x_3)+C_3),\\ f_{11}(x_1,x_3) &= \widehat{f}_{13}(x_1,x_3) + f_{13}(0,0) = f(x_1,x_3) + (f_3(x_3)+C_3) - (f_1(x_1)+C_1). \quad \square\end{aligned}$$

6.3. Example of a discontinuous solution to a dual problem

Usually, dual multimarginal problem admits a regular solution provided the cost function is regular. For instance, applying the Legendre-type transformation, one can prove (see [30, Theorem 2.2] for more details) that there exists a solution $\{\varphi_i\}_{i=1}^n$ to the dual multimarginal problem with the following property: for all $1 \le i \le n$ and for all $x_i \in X_i$,

$$\varphi_i(x_i) = \inf_{\substack{x_j\in X_j,\\ j\neq i}} \left(c(x_1,\dots,x_n) - \sum_{j=1}^{i-1}\varphi_j(x_j) - \sum_{j=i+1}^{n}\varphi_j(x_j)\right).$$

If c is a Lipschitz function with the Lipschitz constant L, then the expression inside the infimum is also a Lipschitz function in x_i with the same Lipschitz constant. Then φ_i is the infimum of Lipschitz continuous functions with common constant; therefore, φ_i is Lipschitz continuous as well.

In this section we prove that a natural solution to the dual $(3,2)$-problem with Lipschitz c can be even discontinuous and (in a sense) unique.

Consider the following $(3,2)$-problem.

Problem 6.29. *For $1 \le i \le 3$, let $X_i = [0,1]$, let μ_{ij} be the restriction of the Lebesgue measure onto the square $[0,1]^2$, and let $c(x_1,x_2,x_3) = \max(0, x_1+x_2+3x_3-3)$.*

Primal problem. *Find a uniting measure $\pi \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$ such that*

$$\int c(x_1,x_2,x_3)\, d\pi \to \min.$$

Dual problem. *Find a triple of functions $\{f_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$, $f_{ij} \subset L^1([0,1]^2)$ such that*

$$\sum_{\{i,j\}\in\mathcal{I}_{3,2}} f_{ij}(x_i,x_j) \le c(x_1,x_2,x_3) \text{ for all } (x_1,x_2,x_3)\in[0,1]^3,$$

$$\sum_{\{i,j\}\in\mathcal{I}_{3,2}} \int_0^1\int_0^1 f_{ij}(x_i,x_j)\, dx_i dx_j \to \max.$$

The cost function $c(x_1, x_2, x_3) = \max(0, x_1 + x_2 + 3x_3 - 3)$ is Lipschitz continuous, and the triple of measures $\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$ is reducible; hence, there is no duality gap, and solutions to both primal and dual problems exist.

Proposition 6.30. *Let*

$$f_{12}(x_1, x_2) = 0 \textit{ for all points } (x_1, x_2) \in [0,1]^2;$$

$$f_{13}(x_1, x_3) = \begin{cases} 0, & \textit{if } x_3 < \frac{2}{3}, \\ x_1 + \frac{3}{2}x_3 - \frac{3}{2}, & \textit{if } x_3 \geq \frac{2}{3}; \end{cases}$$

$$f_{23}(x_2, x_3) = \begin{cases} 0, & \textit{if } x_3 < \frac{2}{3}, \\ x_2 + \frac{3}{2}x_3 - \frac{3}{2}, & \textit{if } x_3 \geq \frac{2}{3}. \end{cases}$$

Denote $F(x_1, x_2, x_3) = f_{12}(x_1, x_2) + f_{13}(x_1, x_3) + f_{23}(x_2, x_3)$. *Then*

(a) $F(x_1, x_2, x_3) \leq c(x_1, x_2, x_3)$ *for all* $(x_1, x_2, x_3) \in [0,1]^3$*;*
(b) *if the value of* $x_1 + x_2 + 3x_3$ *is integer and* $(x_1, x_2, x_3) \neq (0, 0, 2/3)$*, then* $F(x_1, x_2, x_3) = c(x_1, x_2, x_3)$*.*

Proof. First, one can easily verify the following representation for the function F:

$$F(x_1, x_2, x_3) = \begin{cases} 0, & \text{if } x_3 < \frac{2}{3}, \\ x_1 + x_2 + 3x_3 - 3, & \text{if } x_3 \geq \frac{2}{3}. \end{cases} \tag{56}$$

Thus, $F(x_1, x_2, x_3) \leq \max(0, x_1 + x_2 + 3x_3 - 3) = c(x_1, x_2, x_3)$ for all $(x_1, x_2, x_3) \in [0,1]^3$, and this implies assertion 6.30(a).

Suppose that the value of $x_1 + x_2 + 3x_3$ is integer. Consider the case $x_3 < 2/3$. Equation (56) implies that $F(x_1, x_2, x_3) = 0$. Since $x_1, x_2 \leq 1$, we have $x_1 + x_2 + 3x_3 < 4$, and therefore $x_1 + x_2 + 3x_3 \leq 3$. Thus, $c(x_1, x_2, x_3) = \max(x_1 + x_2 + 3x_3 - 3, 0) = 0 = F(x_1, x_2, x_3)$.

Consider the case $x_3 \geq 2/3$. By equation (56), $F(x_1, x_2, x_3) = x_1 + x_2 + 3x_3 - 3$. If $(x_1, x_2, x_3) \neq (0, 0, 2/3)$, then $x_1 + x_2 + 3x_3 > 2$, and therefore, since $x_1 + x_2 + 3x_3$ is integer, $x_1 + x_2 + 3x_3 \geq 3$. Thus, if $(x_1, x_2, x_3) \neq (0, 0, 2/3)$, then $c(x_1, x_2, x_3) = x_1 + x_2 + 3x_3 - 3 = F(x_1, x_2, x_3)$, and this implies assertion 6.30(b). $\square$

We claim that the constructed triple of functions $\{f_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$ is a solution to the dual Problem 6.29. By Proposition 6.30 it is enough to find a measure $\pi \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$ such that π is concentrated on the set $\{(x_1, x_2, x_3) \colon \operatorname{frac}(x_1 + x_2 + 3x_3) = 0\}$ where $\operatorname{frac}(x)$ means the fractional part of x. The proof of the following lemma is easy and is left to the reader.

Lemma 6.31. *Let* $\pi_{1,1,1}$ *be the surface probability measure concentrated on two triangles that form a set*

$$\{(x_1, x_2, x_3) \colon \operatorname{frac}(x_1 + x_2 + x_3) = 0\}.$$

Then $\operatorname{Pr}_{ij}(\pi_{1,1,1})$ *coincides with the Lebesgue measure restricted to the square* $[0,1]^2$ *for all* $\{i,j\} \in \mathcal{I}_{3,2}$*.*

Using this lemma, we prove a more general statement.

Proposition 6.32. *Assume we are given positive integers* a_1*,* a_2 *and* a_3*. Then there exists a measure* $\pi_{a_1,a_2,a_3} \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$ *concentrated on the set*

$$\{(x_1, x_2, x_3) \colon \operatorname{frac}(a_1x_1 + a_2x_2 + a_3x_3) = 0\}.$$

Proof. Let t_1, t_2 and t_3 be non-negative integers such that $0 \le t_i < a_i$ for $1 \le i \le 3$. Consider the mapping

$$T\colon (x_1, x_2, x_3) \mapsto \left(\frac{x_1 + t_1}{a_1}, \frac{x_2 + t_2}{a_2}, \frac{x_3 + t_3}{a_3}\right).$$

Let $\pi^{t_1,t_2,t_3}_{a_1,a_2,a_3}$ be the image of the measure $\pi_{1,1,1}$ under the mapping T. First, if $(y_1, y_2, y_3) = T(x_1, x_2, x_3)$, then $a_1y_1 + a_2y_2 + a_3y_3 = (x_1 + x_2 + x_3) + (t_1 + t_2 + t_3)$. Hence,

$$\operatorname{frac}(x_1 + x_2 + x_3) = \operatorname{frac}(a_1y_1 + a_2y_2 + a_3y_3),$$

and therefore, since $\pi_{1,1,1}$ is concentrated on the set $\{(x_1, x_2, x_3)\colon \operatorname{frac}(x_1 + x_2 + x_3) = 0\}$, the measure $\pi^{t_1,t_2,t_3}_{a_1,a_2,a_3}$ is concentrated on the set $\{(y_1, y_2, y_3)\colon \operatorname{frac}(a_1y_1 + a_2y_2 + a_3y_3) = 0\}$.

In addition, for all $\{i, j\} \in \mathcal{I}_{3,2}$ the measure $\operatorname{Pr}_{ij}(\pi^{t_1,t_2,t_3}_{a_1,a_2,a_3})$ is the image of $\operatorname{Pr}_{ij}(\pi_{1,1,1})$ under the mapping

$$(x_i, x_j) \mapsto \left(\frac{x_i + t_i}{a_i}, \frac{x_j + t_j}{a_j}\right).$$

Thus, $\operatorname{Pr}_{ij}(\pi^{t_1,t_2,t_3}_{a_1,a_2,a_3})$ is proportional to the Lebesgue measure restricted to the square

$$\left[\frac{t_i}{a_i}, \frac{t_i + 1}{a_i}\right] \times \left[\frac{t_j}{a_j}, \frac{t_j + 1}{a_j}\right]. \tag{57}$$

Let

$$\pi_{a_1,a_2,a_3} = \frac{1}{a_1a_2a_3} \sum_{0 \le t_i < a_i} \pi^{t_1,t_2,t_3}_{a_1,a_2,a_3}.$$

The measure π_{a_1,a_2,a_3} is a probability measure concentrated on the set

$$\{(y_1, y_2, y_3)\colon \operatorname{frac}(a_1y_1 + a_2y_2 + a_3y_3) = 0\}.$$

In addition, it follows from (57) that $\operatorname{Pr}_{ij}(\pi_{a_1,a_2,a_3})$ is the Lebesgue measure restricted to the square $[0, 1]^2$. $\square$

Using this proposition, we immediately obtain the following theorem.

Theorem 6.33. *The triple of functions $\{f_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$ described in Proposition 6.30 is a solution to the dual Problem 6.29, and the measure $\pi_{1,1,3} \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$, concentrated on the set*

$$\{(x_1, x_2, x_3)\colon \operatorname{frac}(x_1 + x_2 + 3x_3) = 0\},$$

is a solution to the primal Problem 6.29.

Unlike Problem 6.12, a solution to the primal Problem 6.29 is non-unique.

Proposition 6.34. *Let π_1 be the restriction of the Lebesgue measure to the set $\{(x_1, x_2, x_3)\colon 0 \le x_1, x_2 \le 1, 0 \le x_3 \le 1/3\}$, and let $\widehat{\pi}_{1,1,2}$ be the image of the measure $\pi_{1,1,2}$ described in Proposition 6.32 under the mapping*

$$T\colon (x_1, x_2, x_3) \mapsto \left(x_1, x_2, \frac{2}{3}x_3 + \frac{1}{3}\right).$$

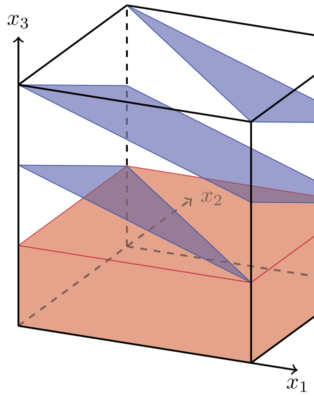

Fig. 5. The support of the solution π described in Proposition 6.34. The support of the measure π_1 is red, and support of $\widehat{\pi}_{1,1,2}$ is blue. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

Then the measure $\pi = \pi_1 + \frac{2}{3}\widehat{\pi}_{1,1,2}$ is uniting and the function $F(x_1, x_2, x_3)$ described in Proposition 6.30 satisfies: $F(x_1, x_2, x_3) = c(x_1, x_2, x_3)$ π-a.e. Consequently, the measure π is a solution to the primal Problem 6.29 (see Fig. 5).

Remark 6.35. In fact, the constructed measure $\frac{2}{3}\widehat{\pi}_{1,1,2}$ is also the restriction of $\pi_{1,1,3}$ to $[0, 1]^2 \times [1/3, 1]$; therefore, the uniting measures provided in Theorem 6.33 and Proposition 6.34 are distinct only on a bottom part $[0, 1]^2 \times [0, 1/3]$ of the space X.

Proof of Proposition 6.34. By construction, $\mathrm{Pr}_{12}(\pi_1)$ is proportional to the restriction of the Lebesgue measure to the square $[0, 1]^2$. The mapping T does not change the projection of a measure onto the space X_{12}, and therefore $\mathrm{Pr}_{12}(\widehat{\pi}_{1,1,2})$ is also proportional to the restriction of the Lebesgue measure to the square $[0, 1]^2$. Thus, $\mathrm{Pr}_{12}(\pi) = \mu_{12}$.

The measure $\mathrm{Pr}_{13}(\pi_1)$ coincides with the restriction of the Lebesgue measure to the rectangle $\{(x_1, x_3)\colon 0 \le x_1 \le 1, 0 \le x_3 \le 1/3\}$. The measure $\mathrm{Pr}_{13}(\widehat{\pi}_{1,1,2})$ is the image of $\mathrm{Pr}_{13}(\pi_{1,1,2})$ under the mapping

$$(x_1, x_3) \mapsto \left(x_1, \frac{2}{3}x_3 + \frac{1}{3}\right).$$

Thus, $\frac{2}{3}\mathrm{Pr}_{13}(\widehat{\pi}_{1,1,2})$ coincides with the restriction of the Lebesgue measure to the rectangle $\{(x_1, x_3)\colon 0 \le x_1 \le 1, 1/3 \le x_3 \le 1\}$, and therefore $\mathrm{Pr}_{13}(\pi) = \mu_{13}$. Similarly, $\mathrm{Pr}_{23}(\pi) = \mu_{23}$, and we conclude that $\pi \in \Pi(\{\mu_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}})$.

Let (x_1, x_2, x_3) be a point in $[0, 1]^3$ such that $x_3 \le 1/3$. By equation (56) we have $F(x_1, x_2, x_3) = 0$. In addition, $x_1 + x_2 + 3x_3 - 3 \le 0$, and therefore $c(x_1, x_2, x_3) = 0$. Thus, since $\mathrm{supp}(\pi_1) = \{(x_1, x_2, x_3) \in [0, 1]^3\colon 0 \le x_3 \le 1/3\}$, we conclude that $F(x_1, x_2, x_3) = c(x_1, x_2, x_3)$ π_1-a.e.

Let (x_1, x_2, x_3) be an arbitrary point in the cube $[0, 1]^2$, and let $(y_1, y_2, y_3) = T(x_1, x_2, x_3)$. We have $y_1 + y_2 + 3y_3 = x_1 + x_2 + 2x_3 + 1$, and therefore

$$\mathrm{frac}(y_1 + y_2 + 3y_3) = \mathrm{frac}(x_1 + x_2 + 2x_3).$$

Hence, we conclude that $\widehat{\pi}_{1,1,2}$ is concentrated on the set $\{(x_1, x_2, x_3)\colon \mathrm{frac}(x_1+x_2+3x_3) = 0\}$, and therefore by assertion 6.30(b) $F(x_1, x_2, x_3) = c(x_1, x_2, x_3)$ $\widehat{\pi}_{1,1,2}$-a.e.

Thus, $F(x_1, x_2, x_3) = c(x_1, x_2, x_3)$ for π-almost all points $(x_1, x_2, x_3) \in [0, 1]^3$, and the measure π is a solution to the primal Problem 6.29. $\square$

Unlike the primal problem, the dual problem admits a unique solution in the following sense.

Proposition 6.36. *Let $\{g_{ij}\}$ be a solution to the relaxed dual Problem 6.29. Then the equation*

$$g_{12}(x_1, x_2) + g_{13}(x_1, x_3) + g_{23}(x_2, x_3) = f_{12}(x_1, x_2) + f_{13}(x_1, x_3) + f_{23}(x_2, x_3)$$

holds for almost all $(x_1, x_2, x_3) \in [0, 1]^3$, where the triple of functions $\{f_{ij}\}_{\{i,j\}\in\mathcal{I}_{3,2}}$ is defined in Proposition 6.30.

First, let us verify the following statement.

Lemma 6.37. *Let $\{g_{ij}\}$ be a solution to the relaxed dual Problem 6.29 (see Definition 5.16). Then there exist integrable functions φ_1 and φ_2 such that $g_{12}(x_1, x_2) = \varphi_1(x_1) + \varphi_2(x_2)$ almost everywhere.*

Proof. Consider the finite $(3, 2)$-function

$$G(x_1, x_2, x_3) = g_{12}(x_1, x_2) + g_{13}(x_1, x_3) + g_{23}(x_2, x_3).$$

Since $\{g_{ij}\}$ is a solution to the relaxed dual problem, the equation $G(x_1, x_2, x_3) = c(x_1, x_2, x_3)$ holds π-almost everywhere, where π is a solution to the primal problem defined in Proposition 6.34. In particular, $G(x_1, x_2, x_3) = c(x_1, x_2, x_3)$ for almost all points $(x_1, x_2, x_3) \in [0, 1]^3$ such that $0 \le x_3 \le 1/3$. Since $c(x_1, x_2, x_3) = \max(x_1 + x_2 + 3x_3 - 3, 0) = 0$ if $x_3 \le 1/3$, we conclude that $G(x_1, x_2, x_3) = 0$ for almost all $(x_1, x_2, x_3) \in [0, 1]^3$ such that $0 \le x_3 \le 1/3$.

In particular, there exists a point $0 \le x_3^{(0)} \le 1/3$ such that the equation $G(x_1, x_2, x_3^{(0)}) = 0$ holds for almost all $(x_1, x_2) \in [0, 1]^2$. Hence, if we denote $\varphi_1(x_1) = -g_{13}(x_1, x_3^{(0)})$ and $\varphi_2(x_2) = -g_{23}(x_2, x_3^{(0)})$, then the equation

$$g_{12}(x_1, x_2) = -g_{13}(x_1, x_3^{(0)}) - g_{23}(x_2, x_3^{(0)}) = \varphi_1(x_1) + \varphi_2(x_2)$$

holds for almost all $(x_1, x_2) \in [0, 1]^2$.

Let us verify that φ_1 and φ_2 are integrable. Since g_{12} is integrable, it follows from the Fubini-Tonelli theorem that for almost all $x_2^{(0)} \in [0, 1]$ the function $x_1 \mapsto g_{12}(x_1, x_2^{(0)}) = \varphi(x_1) + \varphi_2(x_2^{(0)})$ is also integrable. Since $\varphi_2(x_2^{(0)})$ is a constant, we conclude that $\varphi_1(x_1)$ is integrable. The integrability of φ_2 is proven in the same manner. □

It follows from Lemma 6.37 that if $\{g_{ij}\}$ is a solution to the relaxed dual problem, then we can set $\widehat{g}_{12}(x_1, x_2) = 0$, $\widehat{g}_{13}(x_1, x_3) = g_{13}(x_1, x_3) + \varphi_1(x_1)$ and $\widehat{g}_{23}(x_2, x_3) = g_{23}(x_2, x_3) + \varphi_2(x_2)$. Then the equation

$$g_{12}(x_1, x_2) + g_{13}(x_1, x_3) + g_{23}(x_2, x_3) = \widehat{g}_{12}(x_1, x_2) + \widehat{g}_{13}(x_1, x_3) + \widehat{g}_{23}(x_2, x_3)$$

holds for all $(x_1, x_2, x_3) \in [0, 1]^3$ except a zero $(3, 2)$-thickness set, and therefore the triple of functions $\{\widehat{g}_{ij}\}$ is also a solution to the relaxed dual problem. Thus, in Proposition 6.36 we may additionally assume that $g_{12}(x_1, x_2) = 0$ for all $(x_1, x_2) \in [0, 1]^2$.

Lemma 6.38. *Let φ_1 and φ_2 be integrable functions defined on the segment $[0, 1]$. Suppose that there exists a real $\varepsilon > 0$ such that the inequality $\varphi_1(x_1) + \varphi_2(x_2) \le 0$ holds for almost all points (x_1, x_2) such that $0 \le x_1 + x_2 \le 1 + \varepsilon$. Then*

$$\int_0^1 \varphi_1(x_1)\, dx_1 + \int_0^1 \varphi_2(x_2) \le 0.$$

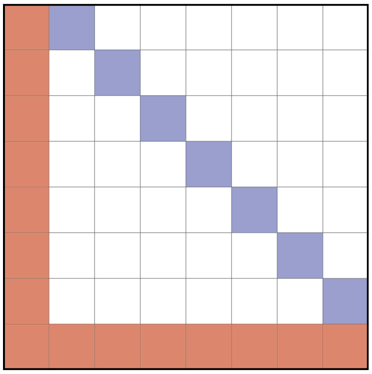

Fig. 6. The supports of the measures μ_1 and μ_2 for the case $\varepsilon = \frac{1}{4}$. The set A_1 is colored red, and the set A_2 is blue.

Moreover, if the equality is achieved, then $\varphi_1(x_1) + \varphi_2(x_2) = 0$ *for almost all* $(x_1, x_2) \in [0,1]^2$. *The same is true if we replace the inequality* $0 \le x_1 + x_2 \le 1 + \varepsilon$ *with* $1 - \varepsilon \le x_1 + x_2 \le 2$.

Proof. Without loss of generality we may assume that $\varepsilon = 1/n$ for some positive integer n. Consider the set $A_1 = \{(x_1, x_2) \in [0,1]^2 \colon \min(x_1, x_2) \le 1/(2n)\}$. Let μ_1 be the restriction of the Lebesgue measure to the set A_1. One can easily verify that if ρ is the density of the projection of μ_1 to the axis, then $\rho(x) = 1$ if $0 \le x \le 1/(2n)$ and $\rho(x) = 1/(2n)$ if $1/(2n) < x \le 1$. In addition, if $\min(x_1, x_2) \le 1/(2n)$, then $0 \le x_1 + x_2 \le 1 + 1/(2n)$, and therefore the inequality $\varphi_1(x_1) + \varphi_2(x_2) \le 0$ holds μ_1-almost everywhere.

Consider the set $A_2 = \{(x_1, x_2) \in [0,1]^2 \colon \lfloor 2nx_1 \rfloor + \lfloor 2nx_2 \rfloor = 2n\}$. Let μ_2 be the restriction of the Lebesgue measure to the set A_2. If $\lfloor 2nx_1 \rfloor + \lfloor 2nx_2 \rfloor = 2n$, then $2nx_1 + 2nx_2 < 2n + 2$, and therefore $x_1 + x_2 < 1 + 1/n$. Hence, $\varphi_1(x_1) + \varphi_2(x_2) \le 0$ for μ_2-almost all points (x_1, x_2). In addition, the projection of μ_2 to the axis is proportional to the restriction of the Lebesgue measure to the segment $[1/(2n), 1]$, and the density of this projection is equal to $1/(2n)$ on this segment. See Fig. 6 for the visualization of the sets A_1 and A_2.

Consider the measure $\mu = \mu_1 + (2n-1)\mu_2$. The projections of this measure to the axes coincides with the restriction of the Lebesgue measure to the segment $[0,1]$. In addition, $\operatorname{supp}(\mu) \subset \{(x_1, x_2) \in [0,1]^2 \colon 0 \le x_1 + x_2 \le 1 + 1/n\}$. Thus, we have

$$\int_0^1 \varphi_1(x_1)\,dx_1 + \int_0^1 \varphi_2(x_2)\,dx_2 = \int_{[0,1]^2} (\varphi_1(x_1) + \varphi_2(x_2))\,\mu(dx_1, dx_2) \le 0.$$

Assume that the equality holds. Then $\varphi_1(x_1) + \varphi_2(x_2) = 0$ μ-almost everywhere. In particular, $\varphi_1(x_1) + \varphi_2(x_2) = 0$ for almost all points $(x_1, x_2) \in A_1$, and therefore this equation holds for almost all points (x_1, x_2) such that $0 \le x_2 \le 1/(2n)$. Thus, by the Fubini-Tonelli theorem there exists a point $x_2^{(0)} \in [0, 1/(2n)]$ such that the equation $\varphi_1(x_1) + \varphi_2(x_2^{(0)}) = 0$ holds for almost all $x_1 \in [0,1]$, and therefore there exists a constant $C_1 = -\varphi_2(x_2^{(0)})$ such that $\varphi_1(x_1) = C_1$ almost everywhere.

Similarly, there exists a constant C_2 such that $\varphi_2(x_2) = C_2$ almost everywhere. Then

$$0 = \int_0^1 \varphi_1(x_1)\,dx_1 + \int_0^1 \varphi_2(x_2)\,dx_2 = C_1 + C_2,$$

and therefore $\varphi_1(x_1) + \varphi_2(x_2) = 0$. The case of the inequality $1 - \varepsilon \le x_1 + x_2 \le 2$ is proven in the same manner. $\square$

Proof of Proposition 6.36. By Lemma 6.37 we may assume that $g_{12} \equiv 0$. Consider the finite $(3,2)$-function

$$G(x_1, x_2, x_3) = g_{13}(x_1, x_3) + g_{23}(x_2, x_3).$$

The function G is integrable and the inequality $G(x_1, x_2, x_3) \leq c(x_1, x_2, x_3)$ holds for almost all points $(x_1, x_2, x_3) \in [0,1]^3$. Hence, there exists a set $A \subseteq [0,1]$ with full measure such that if $x_3^{(0)} \in A$, then the function $G(\cdot, \cdot, x_3^{(0)})$ is integrable and the inequality $G(x_1, x_2, x_3^{(0)}) \leq c(x_1, x_2, x_3^{(0)})$ holds for almost all $(x_1, x_2) \in [0,1]^2$.

Assume that $x_3^{(0)} \in A$ and that $x_3^{(0)} < 2/3$. Consider the $(3,2)$-function

$$F(x_1, x_2, x_3) = f_{12}(x_1, x_2) + f_{13}(x_1, x_3) + f_{23}(x_2, x_3) = f_{13}(x_1, x_3) + f_{23}(x_2, x_3).$$

By equation (56) we have $F(x_1, x_2, x_3^{(0)}) = 0$ for all $(x_1, x_2) \in [0,1]^2$.

Denote $\varepsilon = 2 - 3x_3^{(0)}$. We have $\varepsilon > 0$. If $x_1 + x_2 \leq 1 + \varepsilon$, then $x_1 + x_2 + 3x_3^{(0)} - 3 \leq 0$, and therefore

$$c(x_1, x_2, x_3^{(0)}) = \max(x_1 + x_2 + 3x_3^{(0)} - 3) = 0 = F(x_1, x_2, x_3^{(0)}).$$

In addition, since $G(x_1, x_2, x_3^{(0)}) \leq c(x_1, x_2, x_3^{(0)})$ for almost all points (x_1, x_2), we conclude that the inequality $G(x_1, x_2, x_3^{(0)}) \leq F(x_1, x_2, x_3^{(0)})$ holds for almost all points (x_1, x_2) such that $0 \leq x_1 + x_2 \leq 1 + \varepsilon$.

Consider the functions

$$\varphi_1(x_1) = g_{13}(x_1, x_3^{(0)}) - f_{13}(x_1, x_3^{(0)}) \text{ and } \varphi_2(x_2) = g_{23}(x_2, x_3^{(0)}) - f_{23}(x_2, x_3^{(0)}). \tag{58}$$

We have

$$\varphi_1(x_1) + \varphi_2(x_2) = G(x_1, x_2, x_3^{(0)}) - F(x_1, x_2, x_3^{(0)}).$$

Hence, the function $\varphi_1(x_1) + \varphi_2(x_2)$ is integrable on $[0,1]^2$, and therefore both functions φ_1 and φ_2 are integrable on $[0,1]$. In addition, the inequality $\varphi_1(x_1) + \varphi_2(x_2) \leq 0$ holds for almost all points (x_1, x_2) such that $0 \leq x_1 + x_2 \leq 1 + \varepsilon$. Thus, it follows from Lemma 6.38 that

$$\int\limits_{[0,1]^2} \left(G(x_1, x_2, x_3^{(0)}) - F(x_1, x_2, x_3^{(0)}) \right) dx_1 dx_2 = \int\limits_0^1 \varphi_1(x_1)\, dx_1 + \int\limits_0^1 \varphi_2(x_2)\, dx_2 \leq 0.$$

Moreover, if the equality holds, then $G(x_1, x_2, x_3^{(0)}) = F(x_1, x_2, x_3^{(0)})$ almost everywhere.

Assume that $x_3^{(0)} \in A$ and that $x_3^{(0)} > 2/3$. By equation (56) we have

$$F(x_1, x_2, x_3^{(0)}) = x_1 + x_2 + 3x_3^{(0)} - 3.$$

Denote $\varepsilon = 3x_3^{(0)} - 2 > 0$. If $x_1 + x_2 > 1 - \varepsilon$, then $x_1 + x_2 + 3x_3 - 3 > 0$, and therefore

$$c(x_1, x_2, x_3^{(0)}) = \max(x_1 + x_2 + 3x_3^{(0)} - 3, 0) = x_1 + x_2 + 3x_3^{(0)} - 3 = F(x_1, x_2, x_3^{(0)}).$$

Hence, since $G(x_1, x_2, x_3^{(0)}) \leq c(x_1, x_2, x_3^{(0)})$ for almost all (x_1, x_2), we conclude that $\varphi_1(x_1) + \varphi_2(x_2) \leq 0$ for almost all points (x_1, x_2) such that $1 - \varepsilon \leq x_1 + x_2 \leq 2$, where the functions φ_1 and φ_2 are defined in equation (58). Thus, it follows from Lemma 6.38 that

$$\int\limits_{[0,1]^2} G(x_1, x_2, x_3^{(0)})\, dx_1 dx_2 \leq \int\limits_{[0,1]^2} F(x_1, x_2, x_3^{(0)})\, dx_1 dx_2, \tag{59}$$

and if the equality holds, then $G(x_1, x_2, x_3^{(0)}) = F(x_1, x_2, x_3^{(0)})$ for almost all (x_1, x_2).

Summarizing these results, we conclude that if $x_3^{(0)} \in A$ and if $x_3^{(0)} \neq 2/3$, then inequality (59) holds, and therefore, since A is a set of full measure, we have

$$\int\limits_{[0,1]^3} G(x_1, x_2, x_3)\, dx_1 dx_2 dx_3 \leq \int\limits_{[0,1]^3} F(x_1, x_2, x_3)\, dx_1 dx_2 dx_3.$$

Since $\{g_{ij}\}$ is a solution to the relaxed dual problem, the equality holds, and therefore the equality in inequality (59) is achieved for almost all $x_3^{(0)}$. Thus, for almost all $x_3^{(0)} \in [0,1]$ the equation $F(x_1, x_2, x_3^{(0)}) = G(x_1, x_2, x_3^{(0)})$ holds for almost $(x_1, x_2) \in [0,1]^2$, and therefore

$$g_{12}(x_1, x_2) + g_{13}(x_1, x_3) + g_{23}(x_2, x_3) = f_{12}(x_1, x_2) + f_{13}(x_1, x_3) + f_{23}(x_2, x_3)$$

almost everywhere. □

Funding

The article was prepared within the framework of the HSE University Basic Research Program. The second named author was supported by RFBR project 20-01-00432.

Declaration of competing interest

None.

References

[1] L. Ambrosio, N. Gigli, G. Savaré, Gradient Flows in Metric Spaces and in the Space of Probability Measures, 2 edition, Birkhäuser, Basel, 2008.
[2] M. Beiglböck, N. Juillet, On a problem of optimal transport under marginal martingale constraints, Ann. Probab. 44 (1) (Jan. 2016) 42–106, https://doi.org/10.1214/14-AOP966.
[3] M. Beiglböck, M. Goldstern, G. Maresch, W. Schachermayer, Optimal and better transport plans, J. Funct. Anal. 256 (6) (Mar. 2009) 1907–1927, https://doi.org/10.1016/j.jfa.2009.01.013.
[4] U. Bindini, L. De Pascale, Optimal transport with Coulomb cost and the semiclassical limit of density functional theory, J. Éc. Polytech. Math. 4 (2017) 909–934, https://doi.org/10.5802/jep.59.
[5] V.I. Bogachev, Measure Theory, Springer-Verlag, Berlin Heidelberg, 2007.
[6] V.I. Bogachev, A.V. Kolesnikov, The Monge–Kantorovich problem: achievements, connections, and perspectives, Russ. Math. Surv. 67 (5) (Oct. 2012) 785–890, https://doi.org/10.1070/RM2012v067n05ABEH004808.
[7] Y. Brenier, Décomposition polaire et réarrangement monotone des champs de vecteurs, C. R. Acad. Sci., Sér. 1 Math. 305 (1987) 805–808.
[8] H. Brézis, Analyse fonctionelle. Théorie et applications, Masson, Paris, 1983.
[9] D. Chafaï, A. Hardy, M. Maïda, Concentration for Coulomb gases and Coulomb transport inequalities, J. Funct. Anal. 275 (6) (2018) 1447–1483, https://doi.org/10.1016/j.jfa.2018.06.004.
[10] M. Colombo, L. De Pascale, S. Di Marino, Multimarginal optimal transport maps for one dimensional repulsive costs, Can. J. Math. 67 (2015) 350–368, https://doi.org/10.4153/CJM-2014-011-x.
[11] C. Cotar, G. Friesecke, C. Klüppelberg, Density functional theory and optimal transportation with Coulomb cost, Commun. Pure Appl. Math. 66 (4) (2013) 548–599, https://doi.org/10.1002/cpa.21437.
[12] C. Cotar, G. Friesecke, B. Pass, Infinite-body optimal transport with Coulomb cost, Calc. Var. Partial Differ. Equ. 54 (1) (2015) 717–742, https://doi.org/10.1007/s00526-014-0803-0.
[13] M. Cuturi, G. Peyré, Computational optimal transport: with applications to data science, Found. Trends Mach. Learn. 11 (5-6) (2019) 355–607, https://doi.org/10.1561/2200000073.
[14] S. Di Marino, A. Gerolin, L. Nenna, Optimal transportation theory with repulsive costs, in: Topological Optimization and Optimal Transport: in the Applied Sciences, De Gruyter, Berlin, Boston, 2017, pp. 204–256, chapter 9.
[15] A.N. Doledenok, On a Kantorovich problem with a density constraint, Math. Notes 104 (1) (2018) 45–55, https://doi.org/10.4213/mzm11506.
[16] A. Galichon, Optimal Transport Methods in Economics, Princeton University Press, 2016.
[17] N. Ghoussoub, A. Moameni, Symmetric Monge-Kantorovich problems and polar decompositions of vector fields, Geom. Funct. Anal. 24 (4) (June 2014) 1129–1166, https://doi.org/10.1007/s00039-014-0287-2.

[18] N.A. Gladkov, A.P. Zimin, An explicit solution for a multimarginal mass transportation problem, SIAM J. Math. Anal. 52 (4) (Aug. 2020) 3666–3696, https://doi.org/10.1137/18M122707X.
[19] N.A. Gladkov, A.V. Kolesnikov, A.P. Zimin, On multistochastic Monge–Kantorovich problem, bitwise operations, and fractals, Calc. Var. Partial Differ. Equ. 58 (5) (Sept. 2019), https://doi.org/10.1007/s00526-019-1610-4.
[20] N. Gozlan, N. Juillet, On a mixture of Brenier and Strassen theorems, Proc. Lond. Math. Soc. 120 (3) (Mar. 2020) 434–463, https://doi.org/10.1112/plms.12302.
[21] N. Gozlan, C. Roberto, P.-M. Samson, P. Tetali, Kantorovich duality for general transport costs and applications, J. Funct. Anal. (ISSN 0022-1236) 273 (11) (Dec. 2017) 3327–3405, https://doi.org/10.1016/j.jfa.2017.08.015.
[22] N. Guillen, R. McCann, Five lectures on optimal transportation: geometry, regularity and applications, in: Analysis and Geometry of Metric Measure Spaces: Lecture Notes of the 50th Séminaire de Mathématiques Supérieures (SMS), Montréal, 2011, in: CRM Proceedings & Lecture Notes, 2013, chapter 6.
[23] P. Henry-Labordere, Model-Free Hedging: A Martingale Optimal Transport Viewpoint, 1 edition, Chapman and Hall/CRC, New-York, 2017.
[24] H.G. Kellerer, Verteilungsfunktionen mit gegebenen marginalverteilungen, Z. Wahrscheinlichkeitstheor. Verw. Geb. (ISSN 1432-2064) 3 (3) (Sept. 1964) 247–270, https://doi.org/10.1007/BF00534912.
[25] H.G. Kellerer, Duality theorems for marginal problems, Z. Wahrscheinlichkeitstheor. Verw. Geb. (ISSN 1432-2064) 67 (4) (Nov. 1984) 399–432, https://doi.org/10.1007/BF00532047.
[26] A.V. Kolesnikov, D.A. Zaev, Exchangeable optimal transportation and log-concavity, Theory Stoch. Process. 20 (2) (2015) 54–62.
[27] A.V. Kolesnikov, D.A. Zaev, Optimal transportation of processes with infinite Kantorovich distance: independence and symmetry, Kyoto J. Math. 57 (2) (June 2017) 293–324, https://doi.org/10.1215/21562261-3821819.
[28] J. Korman, R.J. McCann, Insights into capacity-constrained optimal transport, Proc. Natl. Acad. Sci. USA 110 (25) (June 2013) 10064–10067, https://doi.org/10.1073/pnas.1221333110.
[29] J. Korman, R.J. McCann, Optimal transportation with capacity constraints, Trans. Am. Math. Soc. 367 (3) (Nov. 2015) 1501–1521, https://doi.org/10.1090/S0002-9947-2014-06032-7.
[30] B. Pass, Uniqueness and Monge solutions in the multimarginal optimal transportation problem, SIAM J. Math. Anal. 43 (6) (2011) 2758–2775, https://doi.org/10.1137/100804917.
[31] B. Pass, Multi-marginal optimal transport: theory and applications, Modél. Math. Anal. Numér. 49 (6) (Nov. 2015) 1771–1790, https://doi.org/10.1051/m2an/2015020.
[32] S.T. Rachev, L. Rüschendorf, Mass Transportation Problems, Probab. Appl., Springer, New York, NY, 1998.
[33] F. Santambrogio, Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling (Progress in Nonlinear Differential Equations and Their Applications), Birkhäuser, 2015.
[34] W. Schachermayer, J. Teichmann, Characterization of optimal transport plans for the Monge-Kantorovich problem, Proc. Am. Math. Soc. 137 (2) (Feb. 2009) 519–529, https://doi.org/10.1090/S0002-9939-08-09419-7.
[35] A.M. Vershik, P.B. Zatitskiy, F.V. Petrov, Virtual continuity of measurable functions and its applications, Russ. Math. Surv. 69 (6) (Dec. 2014) 1031–1063, https://doi.org/10.1070/RM2014v069n06ABEH004927.
[36] C. Villani, Topics in Optimal Transportation Theory, Grad. Stud. Math., vol. 58, Amer. Math. Soc., Providence, RI, 2003.
[37] C. Villani, Optimal Transport: Old and New, Grundlehren Math. Wiss., vol. 338, Springer-Verlag, Berlin, Heidelberg, 2009.
[38] D.A. Zaev, On the Monge–Kantorovich problem with additional linear constraints, Math. Notes 98 (5–6) (Nov. 2015) 725–741, https://doi.org/10.1134/S0001434615110036.
[39] A.P. Zimin, The Multistochastic Monge-Kantorovich Problem, Master's thesis, HSE University, Moscow, Russia, 2021, https://www.hse.ru/en/edu/vkr/474657068.