Monge solutions and uniqueness in multi-marginal optimal transport: costs associated to graphs and a general condition

by

Adolfo Vargas-Jiménez

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Mathematics

Department of Mathematical and Statistical Sciences
University of Alberta

Abstract

This thesis is devoted to the proof of several results on the existence and uniqueness of Monge solutions to the multi-marginal optimal transportation problem. These results are found in Chapters 3, 4 and 5, and represent joint work with Brendan Pass. The Chapters 1 and 2 are devoted to the introduction and preliminaries respectively.

In Chapter 3 we study a multi-marginal optimal transportation problem with a cost function of the form $c(x_1, \ldots, x_m) = \sum_{k=1}^{m-1} |x_k - x_{k+1}|^2 + |x_m - F(x_1)|^2$, where $F : \mathbb{R}^n \to \mathbb{R}^n$ is a given map. When $m = 4$, F is a positive multiple of the identity mapping, and the first and last marginals are absolutely continuous with respect to Lebesgue measure, we establish that any solution of the Kantorovich problem is induced by a map; the solution is therefore unique. We go on to show that this result is sharp in a certain sense. Precisely, we exhibit examples showing that Kantorovich solutions may concentrate on higher dimensional sets if any of the following hold: 1) F is any linear mapping other than a positive scalar multiple of the identity, 2) the last marginal is not absolutely continuous with respect to Lebesgue measure, or 3) the number of marginals $m \geq 5$, even when F is the identity mapping. In the fourth chapter we study a multi-marginal optimal transport problem with cost $c(x_1, \ldots, x_m) = \sum_{\{i,j\} \in P} |x_i - x_j|^2$, where $P \subseteq Q := \{\{i, j\} : i, j \in \{1, 2, ...m\}, i \neq j\}$. We reformulate this problem by associating each cost of this type with a graph with m vertices whose set of edges is indexed by P. We then establish uniqueness and Monge solution results for two general classes of cost functions. Among many other examples, these classes encapsulate the Gangbo and Święch cost [27] and the cost $c(x_1, \ldots, x_m) = \sum_{k=1}^{m-1} |x_k - x_{k+1}|^2 + |x_m - x_1|^2$ when $m \leq 4$. In the final chapter we establish a general condition on the cost function to obtain uniqueness and Monge solutions in the multi-marginal optimal transport problem, under the assumption that a given collection of the marginals are absolutely continuous with respect to Lebesgue measure. When only the first marginal is assumed to be absolutely continuous, our condition is equivalent to the twist on splitting sets condition found in [35]. In addition, it is satisfied by the special cost functions of Chapter 3 and 4 (found also in [48, 49]), when absolute continuity is imposed on certain other collections of marginals. We also present several new examples of cost functions which violate the twist on splitting sets condition but satisfy the new condition introduced here; we therefore obtain Monge solution and uniqueness results for these cost functions, under regularity conditions on an appropriate subset of the marginals.

Preface

This thesis is an original work by Adolfo Vargas-Jiménez and Brendan Pass.

Chapter 3 has been published as Brendan Pass and Adolfo Vargas-Jiménez, (2021) ”Multi-marginal optimal transportation problem for cyclic costs”, SIAM J. Math. Anal. 53, 4386–4400.

Chapter 4 has been submitted for publication as Brendan Pass and Adolfo Vargas-Jiménez, ”Monge solutions and uniqueness in multi-marginal optimal transport via graph theory”.

Chapter 5 has been submitted for publication as Brendan Pass and Adolfo Vargas-Jiménez, ”Monge solutions and uniqueness in multi-marginal optimal transport: weaker conditions on the cost, stronger conditions on the marginals”.

I was responsible for the proofs of results. Brendan Pass was a supervisory author and contributed to the ideas of the manuscripts as well as manuscript edits.

Dedication

Para Vane y Vale

Acknowledgements

First, I want to express my deepest gratitude to my supervisor, Professor Brendan Pass, for his generous support and excellent guidance. Our weekly meetings have been enriching, expanding my horizons in mathematics and making my Ph.D. journey a smooth path. He has been especially courageous and patient in facing my shortcomings in English. Without his guidance and support beyond the limits of mathematics, I could not have completed this thesis.

I thank the members of my supervisory committee, Professors Xinwei Yu and Eric Woolgar. I am grateful with their valuable feedback on this work during some stages of Chapters 3 and 4, as well as in my thesis defense. I am also grateful for the positive comments e insightful suggestions received from my external examiner Guillaume Carlier.

I thank the staff in the Department of Mathematical and Statistical Sciences for assisting me in a variety of aspects. I would like to especially thank to Tara Schuetz and Professor Kuttler for helping me at crucial moments during the program.

I am indebted to many people outside of the academia. Their support has been crucial, especially in difficult moments where no mathematical technique can help you. I will not make a list here, but all of you are in my heart. Of course, among them, I include my family: my beautiful wife Vanessa and my sweet daughter Valentina. Thank you for your unconditional love. I am definitely a lucky guy.

Table of Contents

List of Figures

Chapter 1

Introduction

1.A Background

Multi-marginal optimal transport is the problem of correlating a finite number of mass distributions to minimize a notion of total cost. This problem is a natural extension of the well-known classical optimal transport where the correlation is done over two mass distributions (think of using a pile of dirt to fill a hole as efficiently as possible relative to a cost function c modeling the cost of transportation). The classical problem was initiated by Monge in 1781 [40]; later in 1942, Kantorovich established a relaxation by allowing the mass to be split into different target points [33].

There are two formulations of the multi-marginal transportation problem: the Kantorovich formulation and the Monge formulation. In the Kantorovich formulation, given Borel probability measures μ_i on open bounded sets $X_i \subseteq \mathbb{R}^n$, with $i = 1, \ldots, m$, and c a real-valued cost function on the product space $X_1 \times \ldots \times X_m$, the goal is to minimize

$$\int_{X_1 \times \ldots \times X_m} c(x_1, \ldots, x_m) d\mu, \tag{KP}$$

among all Borel probability measures μ on the product space $X_1 \times \ldots \times X_m$ whose marginals are the μ_i; that is, for each fixed $i \in \{1, \ldots, m\}$, $\mu(X_1 \times \ldots \times X_{i-1} \times A \times X_{i+1} \times \ldots \times X_m) = \mu_i(A)$ for any Borel set $A \subseteq X_i$.

In the Monge formulation, one seeks to minimize

$$\int_{X_1} c(x_1, T_2x_1, \ldots, T_mx_1)d\mu_1, \tag{MP}$$

among all $(m-1)$-tuples of maps $(T_2, \ldots, T_m)$ such that $(T_i)_\sharp\mu_1 = \mu_i$ for all $i = 2, \ldots, m$, where $(T_i)_\sharp\mu_1$ denotes the *image measure* of μ_1 through T_i, defined by $(T_i)_\sharp\mu_1(A) = \mu_1(T_i^{-1}(A))$, for any Borel set $A \subseteq X_i$. It is well known that problem (KP) is a relaxation of problem (MP), as for any $(m-1)$-tuple of maps $(T_2, \ldots, T_m)$ satisfying the image measure constraint in (MP), we can define $\mu = (Id, T_2, \ldots, T_m)_\sharp\mu_1$, which satisfies the constraint in (KP) and

$$\int_{X_1 \times \ldots \times X_m} c(x_1, \ldots, x_m)d\mu = \int_{X_1} c(x_1, T_2x_1, \ldots, T_mx_1)d\mu_1.$$

Under very general conditions (for instance, compactness of the spaces and continuity of the cost is more than enough) there exists a solution for (KP)[51].

When $m = 2$, the classical optimal transport problems of Monge and Kantorovich arise in (MP) and (KP), respectively. This case has been widely studied and it is reasonably well understood; in particular, under a twist condition on c (the map $x_2 \mapsto D_{x_1}c(x_1, x_2)$ is injective, for each fixed $x_1 \in X_1$, where $D_{x_1}c$ denotes the derivative of c with respecto to x_1) and assuming μ_1 is absolutely continuous with respect to Lebesgue measure $\mathcal{L}^n$, there exists a unique solution to (KP) and it is induced by a map [11, 25, 26, 37]. The classical optimal transport has profound connections with many different areas of mathematics, including analysis, probability, PDE and geometry, and an extremely wide range of applications in other fields, surveyed in, for example, [51, 52, 53] (see also [2] for an overview). For the case $m \geq 3$, a wide variety of applications has also recently emerged, including, for example, matching in economics [18, 19, 44], density functional theory in computation [9, 10, 14, 15, 21], and interpolating among distributions in machine learning and statistics [7, 54] (see also [45] for an overview and additional references). However, determining whether solutions to the multi-marginal Kantorovich problem (KP) are unique and of Monge form has proven much more challenging, as the answer depends on the form of the cost c in subtle ways which are still not understood.

One of the best known cost functions in the multi-marginal setting is the Gangbo and Święch cost function [27]:

$$\sum_{1\leq i<j\leq m} |x_i - x_j|^2. \tag{1.1}$$

In their seminal work, they prove that every solution to (KP) is concentrated on a graph of a measurable map (with μ_1 absolutely continuous with respect to $\mathcal{L}^n$), thus obtaining a unique solution to the Monge-Kantorovich problem (in the subsequently developed terminology of [35], (1.1) is twisted on splitting sets; see definition 2.5 below). In [1], Agueh and Carlier proved that solving the multi-marginal Kantorovich problem with a weighted version of (1.1) is equivalent to finding the barycenter of the marginals $\mu_1, \ldots, \mu_m$.

A fundamental characteristic of the Gangbo and Święch cost is that the first variable x_1 exhibits a direct interaction with all the other variables; that is, the sum $\sum_{1<j\leq m} |x_1 - x_j|^2$ is a term of the sum in (1.1). As we will see in Chapter 4, if this interaction is not given and only μ_1 is absolutely continuous, we can show that uniqueness is not obtained via simple examples; in particular, the twist on splitting sets condition does not hold. An example of a cost function where such interaction is not given is the Euler cost with $m \geq 4$,

$$\sum_{i=1}^{m-1} |x_i - x_{i+1}|^2 + |x_m - x_1|^2. \tag{1.2}$$

As highlighted in Section 1.7.4 of [51] this cost measures the discrete time kinetic energy of a cloud of particles whose density at timestep k is μ_k, such that the initial and final position of a particle is x_1. A more general framework is given if the final position of the particle initially at x_1 is fixed to be $F(x_1)$; that is,

$$\sum_{i=1}^{m-1} |x_i - x_{i+1}|^2 + |x_m - F(x_1)|^2, \tag{1.3}$$

where $F : \mathbb{R}^n \to \mathbb{R}^n$ is a given map. In particular, when each $\mu_k = \mathcal{L}^n|_D$ is (normalized) Lebesgue measure on a common bounded domain $X_k = D \subset \mathbb{R}^n$ and $F : D \to D$ is measure preserving, $F_{\#}\mu_k = \mu_k$, the Monge problem with this cost corresponds to the time discretization of Arnold's variational interpretation of the incompressible Euler equation [3]; the Kantorovich formulation corresponds to a discretization of Brenier's generalization [12]. If $m = 2$ and $I + DF(x)$ is invertible (alternatively it corresponds to the quadratic cost up to a change of variables) where I denotes the identity matrix, the cost is twisted; while for $m = 3$, it is twisted on splitting sets as long as $DF(x) + DF(x)^T > 0$. On the other hand, for $m \geq 4$, little is known about the structure of solutions, although the problems has received a fair bit of attention from a numerical perspective [6, 13, 24, 38]. In Chapter 3, we will establish new results on this structure (available also in [48]).

In Chapter 4 we encapsulate the cost functions (1.1) and (1.2) by studying a more general form in which arbitrary interaction structures between the variables are permitted. More precisely, we consider

$$\sum_{\{i,j\} \in P} |x_i - x_j|^2, \tag{1.4}$$

where $P \subseteq Q := \{\{i, j\} : i, j \in \{1, 2, ...m\}, i \neq j\}$ (note that (1.4) takes the form (1.1) when $P = Q$, and form (1.2) when $P = \big\{\{i, i + 1\} : i = 1, \ldots, m - 1\big\} \cup \big\{\{1, m\}\big\}$; our main goal is then to identify conditions on P which lead to Monge solutions. For this, we exploit a natural connection to graph theory; in particular, we associate the cost function (1.4) with a graph whose vertices we label $\{v_1, ..., v_m\}$ and whose set of edges is indexed by P. For instance, it is evident that cost (1.1) is associated to a complete graph with vertices $\{v_1, \ldots, v_m\}$, denoted by K_m. See Figure (1.1) for the case $m = 7$.

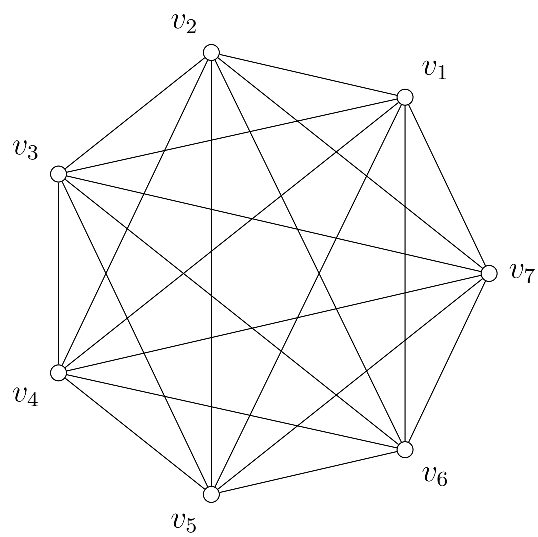


Figure 1.1: K_m with $m = 7$.

In this setting every subgraph with m vertices G of K_m is associated to a cost $\sum_{\{v_i,v_j\}\in E(G)} |x_i - x_j|^2$, where

$$E(G) = \{\{v_i, v_j\} : \ G \text{ has an edge between } v_i \text{ and } v_j, v_i \neq v_j\}.$$

For instance, the "border" of K_m, that is, the cycle graph with vertex sequence $(v_1, \ldots, v_m, v_1)$ (see definition in Section 4.A and figure below for the case $m = 7$), is associated to cost (1.2).

The connection between multi-marginal costs functions and graphs described above recently appeared in a computational setting in [30], where a regularized (through an entropy term) multi-marginal optimal transport problem with cost associated to a tree was studied. Although the scope of that work is restricted to

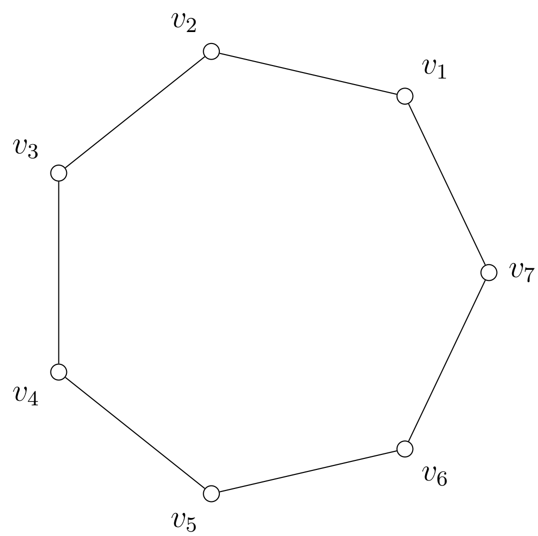


Figure 1.2: Cycle graph with $m = 7$.

a more basic graph structure (only trees were considered), the edges $\{v_i, v_j\}$ are associated to more general symmetric costs $c_{ij}(x_i, x_j)$. Also, [31] established an equivalence of the regularized multi-marginal optimal transport and the inference problem for a probabilistic graphical model when both problems are associated to a common graph structure. On the other hand, the same relationship was noted in [23] when $m = 3$, where connectedness of the graph played an important role in solving a one dimensional multi-marginal martingale optimal transport problem under various assumptions; see Theorem 5.3 in [23].

The final chapter of this thesis focuses on the existence and uniqueness of solutions to the multi-marginal Monge-Kantorovich problems (MP) and (KP) for more general cost functions. This question is in general quite delicate, as the structure of solutions depends subtly on c. In this setting, a condition playing an analogous role to the twist condition was discovered in [35]; this condition was called *twist on c-splitting sets* and states that for every $x_1 \in X_1$ fixed, the map $(x_2, \ldots, x_m) \mapsto D_{x_1}c(x_1, x_2, \ldots, x_m)$ is injective on c-splitting sets (see definition 2.5). The main result in [35] is then that whenever μ_1 is absolutely continuous with respect to $\mathcal{L}^n$, and c is twisted on c-splitting sets, the solution μ to (KP) is unique and induced by a graph. This encapsulates the results for specific costs, or costs satisfying certain conditions, found in [16, 27, 32, 36, 44, 46, 47]. Unlike its two marginal analogue (the classical twist condition), the twist on c-splitting sets is very strong; there are many examples of cost functions for which it fails, and for which non-unique, non-Monge type solutions exist [17, 19, 20, 28, 39, 42, 43]. It is, however, the most general known condition guaranteeing the unique Monge structure of solutions, and it seems unlikely that there is a significantly weaker condition on c under uniqueness is obtained for *all* choices of marginals $\mu_1, ..., \mu_m$ with μ_1 absolutely continuous.

1.B Summary of results

In Chapter 3 we show that the cost function (1.3) is not twisted on splitting sets for $m \geq 4$. Nevertheless, when $m = 4$ and F is a positive scalar multiple of the identity mapping, we are able to prove that all solutions are of Monge type, and therefore unique, under an additional regularity condition on the marginals (in addition to μ_1, either μ_2 or μ_4 must be absolutely continuous). This result is very special; indeed, as we show later on, it is in some sense impossible to go further. A simple example shows that the extra regularity condition on μ_4 or μ_2 is required. When $m = 4$, and F is any linear mapping other than a positive scalar multiple of the identity, we demonstrate that solutions may not be of Monge type, even for diffuse marginals. Similarly, when $m \geq 5$, we prove that solutions may not be of Monge type, even for $F(x) = x$.

To offer some perspective on these results, we note that generalized incompressible flows (ie, solutions to the infinite marginal version of the Kantorovich problem, when each marginal is uniform and F is measure preserving) are not generally unique in dimension $n \geq 2$ [8]; however, unique Monge-type solutions exist when F is close to the identity mapping [22]. It seems reasonable to expect the same to hold for the time discretized problem. Our counterexamples essentially show that this is not the case for $m \geq 4$, at least when the marginals are allowed to differ.

As we shall see in sections 4.B and 4.C, our main results in Chapter 4 (Theorem 4.9 and Theorem 4.12, as well as the related Propositions 4.13 and 4.15), provide a broad class of graphs providing unique Monge solutions to (KP) with cost (1.4); some of these are classical, well known graphs, whereas others are less standard and more exotic. In particular, we highlight in Corollary 4.10 a special subclass of graphs encompassed by our theory, offering a generalization of the Gangbo and Święch result which we find conceptually appealing: the class in which each vertex is connected to all, except at most one, of the other vertices. Generally speaking, the graphs for which we establish Monge solution results come in two complementary classes; one (see Section 4.B) results from the extraction from the complete graph of subgraphs with a particular structure, while the other (see Section 4.C) is obtained by joining complete graphs in a special way.

We would like to emphasize that, in addition to the regularity assumption on μ_1, which is standard in optimal transport, many of our results in Chapter 4 require extra regularity conditions on certain other marginals; these assumptions are not typical in optimal transport theory, but are necessary in our setting, since many counterexamples to Monge solutions and uniqueness exist in their absence (see the second assertion of Proposition 4.6). Note that these examples confirm that the framework developed here reaches well beyond the twist on splitting sets theory, the most general currently known condition implying Monge solution and uniqueness results for multi-marginal problems; indeed, Proposition 4.6 verifies that the twist on splitting sets condition is violated by a wide variety of cost functions, many of which fall within the scope of either Theorem 4.9 or the results in Section 4.C (Theorem 4.12 and Propositions 4.13 and 4.15). The trade-off is that we had to assume regularity of certain subsets of the marginals, rather than only μ_1. This naturally motivates the pursuit of a general condition on c, under which solutions to (KP) are of Monge type and unique, for any collection of marginals $\mu_1, ..., \mu_m$ with μ_i absolutely continuous for all i in a given subset of $\{1, 2, ..., m\}$. The purpose of Chapter 5 is to develop such a condition.

Our condition is formulated in terms of c-splitting functions (see Definition 2.3) and the points where some of them are differentiable (the ones corresponding to the marginals different than μ_1 where regularity is needed). More specifically, we require the mapping $(x_2, \ldots, x_m) \mapsto D_{x_1}c(x_1, x_2, \ldots, x_m)$ to be injective on special subsets generated by c-splitting sets and their associated Borel functions (see Definition 5.2). This condition ensures Monge structure and uniqueness of the optimal elements in $\Pi(\mu_1, \ldots, \mu_m)$, as we shall see in our main result (Theorem 5.6), and it reduces to the twist on splitting sets condition in the special case when only regularity of μ_1 is assumed, but reaches substantially beyond it in general. Aside from including the cost functions in [48] and [49], our condition applies to a wide variety of new costs, as we illustrate with several examples.

One essential aspect of the version of the twist condition presented in this work is the dependence on c-splitting functions of the sets where the map $(x_2, \ldots, x_m) \mapsto D_{x_1}c(x_1, x_2, \ldots, x_m)$ is injective (unlike the twist on c-splitting sets condition where such map is injective on splitting sets with no dependency on c-splitting functions). The involvement of c-splitting functions allows us to naturally generate several differential equations as presented in Lemma 2.8, which are key to naturally exploit the structure of a variety of cost functions. This type of approach is possible, in particular, by the incorporation of additional regularity conditions on the marginals. We also establish an equivalent condition to the twist on c-splitting sets condition that facilitates the proof of some of the results; this condition focuses on every m-tuple of c-splitting functions and an associated largest c-splitting set, instead of every c-splitting set and its associated c-splitting functions (see Lemma 5.1).

Chapter 2

Preliminaries and definitions

In this chapter, we recall some preliminary results and definitions.

2.A The dual problem

For a cost c and $X_i \subseteq \mathbb{R}^n$ for each i, set

$$\mathcal{U} = \left\{ (u_1, u_2, \ldots, u_m) \in \prod_{i=1}^{m} L^1(\mu_i) : c(x_1, \ldots, x_m) \geq \sum_{i=1}^{m} u_i(x_i), \forall (x_1, \ldots, x_m) \in X_1 \times \ldots \times X_m \right\}.$$

The dual of (KP) asks to maximize on $\mathcal{U}$ the map:

$$(u_1, u_2, \ldots, u_m) \mapsto \sum_{i=1}^{m} \int_{X_i} u_i(x_i) d\mu_i(x_i). \tag{DP}$$

The following subclass of $\mathcal{U}$ plays a key role in multi-marginal optimal transport theory.

Definition 2.1. *An m-tuple of functions $(u_1, u_2, \ldots, u_m)$ is c-conjugate if for all i,*

$$u_i(x_i) = \inf_{x_j \in X_j, j \neq i} \Big(c(x_1, \ldots, x_m) - \sum_{j \neq i} u_j(x_j) \Big). \tag{2.1}$$

It is well known that if $(u_1, u_2, \dots, u_m)$ is c-conjugate, then each u_k inherits local Lipschitz and semi-convexity properties from c [37].

The following well known duality result captures the connection between (DP) and (KP). Most of the assertions can be traced back to Kellerer [34]; a proof that the solutions to (DP) can be taken to be c-conjugate can be found in [27] or [46].

Theorem 2.2. *Assume X_k is compact for every k. Then, there exists a solution μ to the Kantorovich problem and a c-conjugate solution $(u_1, u_2, \dots, u_m)$ to its dual. The minimum and maximum values in (DP) and (KP) respectively are the same and $\sum_{k=1}^m u_k(x_k) = c(x_1, \dots, x_m)$ for all $(x_1, \dots, x_m) \in spt(\mu)$, where $spt(\mu)$ denotes the support of μ.*

2.B The twist on c-splitting sets condition

Let us recall some main concepts from [35].

Definition 2.3. *A set $S \subseteq \prod_{i=1}^m X_i$ is called a c-splitting set if there are Borel functions $u_i : X_i \mapsto \mathbb{R}$ such that*

$$\sum_{i=1}^m u_i(x_i) \le c(x_1, \dots, x_m) \tag{2.2}$$

for every $(x_1, \dots, x_m) \in \prod_{i=1}^m X_i$, and whenever $(x_1, \dots, x_m) \in S$ equality holds. The functions $u_1(x_1), \dots, u_m(x_m)$ are called c-splitting functions for S.

Definition 2.4. *A set $S \subseteq \prod_{i=1}^m X_i$ is called c-cyclically monotone if for any finite collection $\{(x_1^k, \dots, x_m^k)\}_{k=1}^p \subseteq S$ we get*

$$\sum_{k=1}^p c(x_1^k, \dots, x_m^k) \le \sum_{k=1}^p c(x_1^{\sigma_1(k)}, \dots, x_m^{\sigma_m(k)}),$$

for every $\sigma_1, \dots, \sigma_m \in S_P$, where S_P denotes the set of permutations of $P := \{1, \dots, p\}$.

It is straightforward to prove that any c-splitting set is c-cyclically monotone. When $m = 2$, the converse is true by Rüschendorf's Theorem [50]. The converse for $m \geq 3$, remained an open question until Griessler proved that in fact, every c-cyclically monotone set is c-splitting [29]. In this work, we shall find it convenient to use both definitions interchangeably.

Definition 2.5. *Let c be a continuous semi-concave cost function. It is called twisted on c-splitting sets, whenever for each fixed $x_1^0 \in X_1$ and c-splitting set $S \subseteq \{x_1^0\} \times X_2 \times \ldots X_m$, the map*

$$(x_2, \ldots, x_m) \mapsto D_{x_1} c(x_1^0, x_2, \ldots, x_m)$$

is injective on the subset of S where $D_{x_1} c(x_1^0, x_2, \ldots, x_m)$ exists.

Remark 2.6. *The main result in [35] establishes that if c is twisted on c-splitting sets, then every solution to (KP) is induced by a map, whenever μ_1 is absolutely continuous with respect to $\mathcal{L}^n$.*

The classical duality theorem of Kellerer (Theorem 2.2) automatically connects Definitions 2.3 and 2.4 with the optimal measures μ in (KP). From now on, $spt(\mu)$ denotes the support of μ.

Lemma 2.7. *A measure $\mu \in \Pi(\mu_1, \ldots, \mu_m)$, solves* (KP) *if and only if $spt(\mu)$ is a c-splitting set.*

Let us finish this section with a convenient lemma, which will reduce some of the technical details of the results in this work. For this, recall that given an open set D and a semi-concave function $f : D \subseteq \mathbb{R}^n \mapsto \mathbb{R}$, the superdifferential of f with respect to a given $x \in A$ fixed is defined as the set

$$\partial f(x) = \{z \in \mathbb{R}^n : f(y) - f(x) \leq z \cdot (y - x) \ \ \forall y \in D\}.$$

It can be proved that $\partial f(x)$ is nonempty for every $x \in D$ and $Df(x)$ exists if and only if $\partial f(x)$ is a singleton.

Lemma 2.8. *Let c be a continuous semi-concave cost function, and $u_i : X_i \mapsto \mathbb{R}$ Borel functions, $i \in \{1, \dots, m\}$, satisfying the inequality condition in (2.2). Let $(x_1^0, \dots, x_m^0) \in \prod_{i=1}^m X_i$ such that*

$$\sum_{i=1}^{m} u_i(x_i^0) = c(x_1^0, \dots, x_m^0). \tag{2.3}$$

If there exists $k \in \{1, \dots, m\}$ such that $Du_k(x_k^0)$ exists, then $D_{x_k} c(x_1^0, \dots, x_m^0)$ exists and

$$Du_k(x_k^0) = D_{x_k} c(x_1^0, \dots, x_m^0).$$

Proof. Since c is semi-concave, the map $x_k \mapsto c(x_1^0, \dots, x_{k-1}^0, x_k, x_{k+1}^0, \dots, x_m^0)$ is semi-concave. Then $\partial_{x_k} c(x_1^0, \dots, x_{k-1}^0, x_k, x_{k+1}^0, \dots, x_m^0)$ is nonempty for every $x_k \in X_k$ fixed, where $\partial_{x_k} c(x_1^0, \dots, x_{k-1}^0, x_k, x_{k+1}^0, \dots, x_m^0)$ denotes the superdifferential of c with respect to x_k. Using (2.3), it follows that

$$\partial_{x_k} c(x_1^0, \dots, x_m^0) \subseteq \partial u_k(x_k^0) = \{Du_k(x_k^0)\}.$$

Thus, $\partial_{x_k} c(x_1^0, \dots, x_m^0)$ is a singleton, which implies that $D_{x_k} c(x_1^0, \dots, x_m^0)$ exists and $Du_k(x_k^0) = D_{x_k} c(x_1^0, \dots, x_m^0)$, completing the proof. □

Chapter 3

Multi-marginal optimal transportation problem for cyclic costs

In this chapter we focus in a multi-marginal optimal transportation problem with cost

$$c(x_1, \dots, x_m) = \sum_{i=1}^{m-1} |x_i - x_{i+1}|^2 + |x_m - F(x_1)|^2, \tag{3.1}$$

where $F : \mathbb{R}^n \to \mathbb{R}^n$ and $m \geq 4$. We will approach the problem of minimizing (KP), by the equivalent problem of maximizing:

$$\mathcal{F}_b[\mu] = \int_X b(x_1, \dots, x_m) d\mu \tag{KPb}$$

where $b(x_1, \dots, x_m) = \sum_{i=1}^{m-1} x_i \cdot x_{i+1} + x_m \cdot F(x_1)$, over the same admissible class of (KP).

3.A Monge Solutions

We now show that under regularity conditions on the first and fourth marginal, we obtain a unique Monge solution for the case $m = 4$ and $F(x) = x$.

Theorem 3.1. *Let μ_i be probability measures over open bounded sets $X_i \subseteq \mathbb{R}^n$, $i = 1, 2, 3, 4$. Take $b(x, y, z, w) = x \cdot y + y \cdot z + z \cdot w + w \cdot x$ and assume μ_1, μ_4 are absolutely continuous with respect to $\mathcal{L}^n$. Then any solution of the Kantorovich problem* (KPb) *is induced by a map.*

The strategy of our proof is based on the following observation: given a solution μ to (KPb) and a b-conjugate solution (u_1, u_2, u_3, u_4) to its dual, the coupling between μ_1 and μ_2 induced by μ (that is, $(\pi_{xy})_{\#}\mu$, where $\pi_{xy} : X_1 \times X_2 \times X_3 \times X_4 \to X_1 \times X_2$ is the canonical projection, $\pi_{xy}(x, y, z, w) = (x, y)$) solves a two marginal optimal transport problem with an *effective* surplus given by:

$$f(x, y) = x \cdot y + \sup_z [y \cdot z - u_3(z) + h(x + z)], \tag{3.2}$$

where

$$h(x + z) = \sup_w [(x + z) \cdot w - u_4(w)]. \tag{3.3}$$

The key to our argument is essentially the verification that f is twisted; that is, $y \mapsto D_x f(x_0, y)$ is injective for any fixed $x_0 \in X_1$ (this condition is well known to ensure Monge solution for two marginal problems [11, 25, 26, 37]). Although the reduction to a two marginal problem can be applied more generally, this strategy to obtain Monge solution depends strongly on the form of the surplus function, as one needs to be able to prove that the effective surplus function f (defined using the Kantorovich potentials) is twisted for an *arbitrary* b-conjugate m-tuple $(u_1, ..., u_m)$ of functions. A similar strategy is applied successfully in the one dimensional case, $n = 1$, in [16].

Proof. Let μ be a solution to (KPb) and (u_1, u_2, u_3, u_4) a b-conjugate solution to its dual. Consider the set

$$S = \Big\{(x, y, z, w) : Du_1(x) \text{ and } Du_4(w) \text{ exist and } b(x, y, z, w) = u_1(x)+u_2(y)+u_3(z)+u_4(w)\Big\}.$$

Since the functions $u_1(x)$ and $u_4(w)$ are Lipschitz, they are differentiable $\mathcal{L}^n$-a.e., and therefore μ_1 and μ_4 a.e. by hypothesis. Hence, $\mu(S) = 1$. Note that

$$b(x, y, z, w) - u_3(z) - u_4(w) \leq f(x, y) \leq u_1(x) + u_2(y),$$

for all x, y, z, w and in particular equality holds on S, where f and h are defined by (3.2) and (3.3), respectively.

Now, for any fixed x_0, we will show that there is only one y_0, z_0, w_0 such that $(x_0, y_0, z_0, w_0) \in S$. Since the function $x \mapsto f(x, y_0)$ is convex and $f(x, y_0) \leqslant u_1(x) + u_2(y_0)$ for every x, it is subdifferentiable everywhere. For $(x_0, y_0, z_0, w_0) \in S$ the equality $f(x_0, y_0) = u_1(x_0)+u_2(y_0)$ implies that the subdifferential of $f(x, y_0)$ at x_0 is contained in the subdifferential of $u_1(x)$ at x_0, which is $\{Du_1(x_0)\}$; that is, $D_x f(x_0, y_0)$ exists and equals $Du_1(x_0)$. By a similar argument $Dh(x_0 + z_0)$ exists, $Dh(x_0 + z_0) = D_x f(x_0, y_0) - y_0 = w_0$, and clearly, $z_0 \in \text{argmax}[y_0 \cdot z - u_3(z) + h(x_0 + z)]$. We claim that the map $(y, z, w) \mapsto D_x f(x_0, y)$ with domain $R := \{(y, z, w) : (x_0, y, z, w) \in S\}$ is injective; this will imply the desired result.

Assume $D_x f(x_0, y_1) = D_x f(x_0, y_2)$ for some $(y_1, z_1, w_1), (y_2, z_2, w_2) \in R$. Note that

$$y_1+w_1 = y_1+Dh(x_0+z_1) = D_x f(x_0, y_1) = D_x f(x_0, y_2) = y_2+Dh(x_0+z_2) = y_2+w_2, \tag{3.4}$$

and $z_i \in \text{argmax}[y_i \cdot z - u_3(z) + h(x_0 + z)]$, $i = 1, 2$. Therefore

$$y_1 \cdot z_2 - u_3(z_2) + h(x_0 + z_2) \leq y_1 \cdot z_1 - u_3(z_1) + h(x_0 + z_1) \tag{3.5}$$

$$y_2 \cdot z_1 - u_3(z_1) + h(x_0 + z_1) \leq y_2 \cdot z_2 - u_3(z_2) + h(x_0 + z_2); \tag{3.6}$$

adding these inequalities gives $(y_1 - y_2) \cdot (z_2 - z_1) \leq 0$, then by (3.4),

$$(w_2 - w_1) \cdot (z_2 - z_1) \leq 0. \tag{3.7}$$

Furthermore, since $w_i \in \text{argmax}[(x_0 + z_i) \cdot w - u_4(w)]$,

$$(x_0 + z_1) \cdot w_2 - u_4(w_2) \leq h(x_0 + z_1) = (x_0 + z_1) \cdot w_1 - u_4(w_1) \tag{3.8}$$

$$(x_0 + z_2) \cdot w_1 - u_4(w_1) \leq h(x_0 + z_2) = (x_0 + z_2) \cdot w_2 - u_4(w_2); \tag{3.9}$$

after adding and canceling similar terms we obtain

$$(w_2 - w_1) \cdot (z_1 - z_2) \leq 0. \tag{3.10}$$

Therefore, by (3.7) and (3.10), $(w_2 - w_1) \cdot (z_1 - z_2) = 0$ and we must have equality in (3.5), (3.6), (3.8) and (3.9). This implies that $w_2 \in \text{argmax}[(x_0 + z_2) \cdot w - u_4(w)] \bigcap \text{argmax}[(x_0 + z_1) \cdot w - u_4(w)]$; additionally, $(y_2, z_2, w_2) \in R$ implies $u_4(w)$ is differentiable at w_2, and so

$$x_0 + z_1 = Du_4(w_2) = x_0 + z_2; \tag{3.11}$$

that is, $z_1 = z_2$. The equality $w_i = Dh(x_0 + z_i)$ for $i = 1, 2$ then implies that $w_1 = w_2$, and so $y_1 = y_2$ by (3.4).
In summary, the equation $D_x f(x_0, y_0) = Du_1(x_0)$, which holds on S and therefore μ almost everywhere, implies that (y_0, z_0, w_0) is uniquely defined from x_0; therefore, the 3-tuple (T_2, T_3, T_4) where T_i is the map associating each x_0 to y_0, z_0 and w_0 respectively, induces μ. □

Remark 3.2. *In a similar way, we can prove Theorem 3.1 if we replace $F = I$ by $F = \lambda I$, where $\lambda > 0$ is a scalar.*

A standard argument now implies uniqueness of solutions to (KPb).

Corollary 3.3. *Assume the same conditions as Theorem 3.1. Then the solution to the Kantorovich problem* (KPb) *is unique.*

Proof. Let μ^1 and μ^2 be distinct solutions of (KPb). By Theorem 3.1, $\mu^1 = (Id, T_2^1, T_3^1, T_4^1)$ and $\mu^2 = (Id, T_2^2, T_3^2, T_4^2)$ for some 3-tuples of measurable maps $(T_2^1, T_3^1, T_4^1) \neq (T_2^2, T_3^2, T_4^2)$. Since the set of solutions of (KPb) is convex, $\mu = \frac{1}{2}\mu^1 + \frac{1}{2}\mu^2$ is also a solution. Hence, applying one more time Theorem 3.1, we conclude that μ is concentrated on a graph. This is clearly not possible, completing the proof. □

3.B Non-Monge Solutions

We now illustrate why the conditions on the marginal μ_4, the number of variables m and the map F in the definition of b of Theorem 3.1 are necessary.

3.B.1 The regularity condition on μ_4.

Assuming m and F as in Theorem 3.1, the next example will show that if μ_4 is not absolutely continuous, we can find a solution for (KP) not induced by a map. Furthermore, the uniqueness result of Corollary 3.3 fails.

Example 3.4. *Let $X_i = B(0, r) \subseteq \mathbb{R}^n$ be an open ball, $r > 0$. Consider $c(x, y, z, w) = \frac{1}{2}\left(|x-y|^2 + |y-z|^2 + |z-w|^2 + |w-x|^2\right)$ and the following measures on X_i: The Dirac measure at the origin $\mu_2 = \mu_4 = \delta_0$ and the normalized n-dimensional Lebesgue measure $\mu_1 = \mu_3 = \frac{\mathcal{L}^n}{k_n r^n}$, where k_n is the volume of the n-dimensional ball of radius 1. Take any μ in $\Pi(\mu_1, \mu_2, \mu_3, \mu_4)$. Since $(x, y, z, w) \in spt(\mu)$ implies $y = w = 0$, we obtain*

$$\int_{X_1\times X_2\times X_3\times X_4} c(x,y,z,w)d\mu = \int_{X_1\times X_2\times X_3\times X_4} \left(|x|^2 + |z|^2\right) d\mu$$
$$= \int_{B(0,r)} |x|^2 d\mu_1(x) + \int_{B(0,r)} |z|^2 d\mu_3(z);$$

that is, $\mathcal{F}[\mu]$ is independent of μ, hence any element in $\Pi(\mu_1, \mu_2, \mu_3, \mu_4)$ is a minimizer. Therefore, we can find optimal measures μ to (KP) not concentrated on a graph of a measurable map; for instance, the product measure $\mu = \mu_1 \otimes \mu_2 \otimes \mu_3 \otimes \mu_4$. On the other hand, if we set $\mu = (Id, F, T, F)_\sharp \mu_1$ where $T_\sharp \mu_1 = \mu_3$ and $F = 0$, we get solutions for the Monge problem.

Remark 3.5. *Theorem 3.1 in [35] (see Remark 2.6) and the previous example imply that the cost function* (3.1) is not *twisted on splitting sets when $m = 4$ and $F = Id$ is the identity mapping. Indeed, if the cost* were *twisted on splitting sets, the result in [35] would imply that the solution to the problem considered in the example would be unique and of Monge type; as this is not the case, the twist on splitting sets condition must fail.*

Nearly identical examples can be constructed to show that (3.1) *is not twisted on splitting sets for any $m \geq 4$ and any choice of F.*

3.B.2 The condition $F = I$.

In this subsection, by assuming F is not a positive multiple of the identity mapping, and that $m = 4$ and $n = 2$, we will find absolutely continuous marginals in $\mathbb{R}^2$ such that a solution of (KPb) is concentrated in a 3-dimensional set. Therefore, this solution will not be induced by a map. For this purpose, we use the next theorem established in [4, 5] and Lemma 3.7.

In what follows, we denote by $\Re^d$ the set of all 2×2 real matrices that can be expressed as the product of d positive definite real matrices.

Theorem 3.6. *Assume that $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is a 2×2 matrix and $|M| > 0$, where $|M|$ denotes the determinant of M, then:*

1. *$M \in \Re^2$ iff M is diagonalizable and its eigenvalues are both positive.*

2. *$M \in \Re^3$ iff $tr(M) > 0$ or $(c - b)^2 > 4|M|$.*

We now recall a couple of well known formulas which will be useful in the construction of counterexamples for the rest of this chapter.

For any 2×2 matrices A and B we have:

$$|A + B| = |A| + |B| + tr\big(Adj(A)B\big), \tag{3.12}$$

where $Adj(A)$ denote the adjugate of A.

Given a convex function $f : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{\infty\}$, its Legendre-transform will be denoted by f^*; that is, $f^*(y) = \sup_x(x \cdot y - f(x))$. We have special interest in the Legendre-transform of $f(x) = \frac{1}{2}x^T Ax + b \cdot x$ for a given positive definite $n \times n$ matrix A and $b \in \mathbb{R}^n$. For this function, we have:

$$f^*(y) = \frac{1}{2}(y - b)^T A^{-1}(y - b). \tag{3.13}$$

Lemma 3.7. *For each 2×2 real matrix F such that $F \neq \lambda I$ for some $\lambda > 0$, there exists $M \in \Re^2$ such that $F + M$ is singular.*

Proof. Let $F = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be 2×2 real matrix such that $F \neq \lambda I$ for any $\lambda > 0$. We want to show that $|F + M| = 0$ for some $M = \begin{pmatrix} e & f \\ g & h \end{pmatrix} \in \Re^2$. First, note that by (3.12),

$$|F + M| = |F| + |M| + (de - gb) + (ah - fc).$$

We divide the proof into 3 cases:

1. If $c \neq 0$, take any $e, h \in \mathbb{R}$ with $e \neq h$ and $e, h > 0$. By setting $f = \frac{|F|+eh+de+ah}{c}$ and $g = 0$ we obtain $|M| > 0$ and $|F + M| = 0$. Furthermore, M is triangular with distinct eigenvalues e and h, hence M is diagonalizable. Since $e, h > 0$, we get $M \in \Re^2$ by Theorem 3.6.

2. If $c = 0$ and $b \neq 0$, using a similar argument as in 1. we obtain the same

result by taking $f = 0$, $g = \frac{|F|+eh+de+ah}{b}$ and any $e, h > 0$, $e \neq h$.

3. If $b = c = 0$, note that by hypothesis $a \neq d$. Also, we have $a, d \geq 0$, $a, d < 0$ or without loss of generality $a < 0$ and $d \geq 0$. For the second and third case, we can make M diagonalizable with positive determinant and satisfying $|F + M| = 0$, by taking $e = h = -a$ and $f = g = 0$. Hence, $M \in \Re^2$ by Theorem 3.6. For the first case, assume without loss of generality $a > d \geq 0$ and consider the matrix

$$M = \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} \frac{ad}{a-d} + \lambda & \frac{d^2}{a-d} + \frac{a+d}{2a}\lambda \\ -\frac{a^2}{a-d} - \lambda & \frac{-ad}{a-d} - \frac{a+d}{2a}\lambda \end{pmatrix},$$

with $\lambda > 0$. Clearly,

$$|F + M| = \begin{vmatrix} \frac{a^2}{a-d} + \lambda & \frac{d^2}{a-d} + \frac{a+d}{2a}\lambda \\ -\frac{a^2}{a-d} - \lambda & -\frac{d^2}{a-d} - \frac{a+d}{2a}\lambda \end{vmatrix} = 0.$$

Since $\frac{d^2}{a-d} + \frac{a+d}{2a}\lambda = -d + \frac{ad}{a-d} + \frac{a+d}{2a}\lambda$ and $\frac{ad}{a-d} + \lambda = -a + \frac{a^2}{a-d} + \lambda$, we have

$$\begin{aligned} |M| &= -\Big(-a + \frac{a^2}{a-d} + \lambda\Big)\Big(\frac{ad}{a-d} + \frac{a+d}{2a}\lambda\Big) + \Big(\frac{a^2}{a-d} + \lambda\Big)\Big(-d + \frac{ad}{a-d} + \frac{a+d}{2a}\lambda\Big) \\ &= a\Big(\frac{ad}{a-d} + \frac{a+d}{2a}\lambda\Big) - d\Big(\frac{a^2}{a-d} + \lambda\Big) \\ &= \Big(\frac{a+d}{2} - d\Big)\lambda \\ &= \Big(\frac{a-d}{2}\Big)\lambda > 0. \end{aligned}$$

Furthermore, $tr(M) = e + h = \frac{a-d}{2a}\lambda > 0$. Hence, $tr(M)^2 - 4|M| > 0$ for big enough λ; that is, the eigenvalues of M given by

$$\frac{tr(M) \pm \sqrt{tr(M)^2 - 4|M|}}{2}$$

are both positive and different. Then M is diagonalizable and belongs to $\Re^2$, by Theorem 3.6.

□

Proposition 3.8. *For* $b(x,y,z,w) = x\cdot y + y\cdot z + z\cdot w + w\cdot F(x)$*, with* $(x,y,z,w) \in (\mathbb{R}^2)^4$ *and* F *a linear map such that* $F \neq \lambda I$ *for any* $\lambda > 0$*, there are absolutely continuous marginals* $\mu_1, \mu_2, \mu_3, \mu_4$ *with respect to* $\mathcal{L}^2$*, such that a solution of* (KPb) *is not concentrated on a graph of a measurable map.*

Proof. Let $F = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be the matrix representation of $F(x)$. By the previous lemma we can choose $M = \begin{pmatrix} e & f \\ g & h \end{pmatrix} \in \Re^2$ such that $F + M$ is singular. Note that $A := M^{-1}F + I$ is also singular and $M^{-1} \in \Re^2$. Let $M_1, M_2 > 0$ be such that $M^{-1} = M_1 M_2$. Decompose each vector $x \in \mathbb{R}^2$ into orthogonal components $x = x_\perp + x_\parallel$ with $x_\parallel$ in the null space of A and $x_\perp$ in the orthogonal complement of the null space of A; that is, the range of A^T. For all x, y, z, w define:

i. $u_1(x) = \frac{|x_\perp|^2}{2} + g_1(x) + g_2(x), \quad u_2(y) = \frac{|A^T y|^2}{2} + g(y), \quad$ where $g_1(x) = \frac{1}{2}(M_2 F x)^T M_1 (M_2 F x), \quad g_2(x) = \frac{1}{2}(Fx)^T M_2 (Fx)$ and $g(y) = \frac{1}{2} y^T M_1 y$.

ii. $u_3(z) = \frac{1}{2} z^T (M_1^{-1} + M_2) z, \quad u_4(w) = \frac{1}{2} w^T M_2^{-1} w.$

iii. $\rho(x,y) = \sup_{z,w}[b(x,y,z,w) - u_3(z) - u_4(w)]$.

Consider the set:

$$W = \Big\{(x,y,z,w) : x_\perp = A^T y, z = M_1(y + M_2 F x) \quad \text{and} \quad w = M_2(z + Fx)\Big\}.$$

We claim

$$b(x,y,z,w) - u_1(x) - u_2(y) - u_3(z) - u_4(w) \leq 0, \tag{3.14}$$

for all $(x,y,z,w) \in (\mathbb{R}^2)^4$ and equality holds on W. For the inequality, it suffices to prove $\rho(x,y) \leq u_1(x) + u_2(y)$.

$$\begin{aligned}
\rho(x,y) &= x \cdot y + \sup_{z,w}[y \cdot z + z \cdot w + w \cdot (Fx) - u_3(z) - u_4(w)] \\
&= x \cdot y + \sup_{z}\left[y \cdot z - u_3(z) + \sup_{w}\left[(z + Fx) \cdot w - u_4(w)\right]\right] \\
&= x \cdot y + \sup_{z}\left[y \cdot z - u_3(z) + u_4^*(z + Fx)\right] \\
&= x \cdot y + \sup_{z}\left[y \cdot z - u_3(z) + \frac{1}{2}(z + Fx)^T M_2 (z + Fx)\right] && \text{by (3.13)} \\
&= x \cdot y + \sup_{z}\left[y \cdot z - \frac{1}{2} z^T M_1^{-1} z + z^T M_2 F x\right] + g_2(x) \\
&= x \cdot y + \frac{1}{2}(y + M_2 F x)^T M_1 (y + M_2 F x) + g_2(x) && \text{by (3.13)} \\
&= x \cdot y + y^T M_1 M_2 F x + g_1(x) + g_2(x) + g(y) \\
&= y^T \cdot Ax + g_1(x) + g_2(x) + g(y) \\
&= y^T \cdot Ax_\perp + g_1(x) + g_2(x) + g(y) \\
&= A^T y \cdot x_\perp + g_1(x) + g_2(x) + g(y) \\
&\leq \frac{|A^T y|^2}{2} + \frac{|x_\perp|^2}{2} + g_1(x) + g_2(x) + g(y) && \text{by the Cauchy-Schwarz Inequality} \\
&= u_1(x) + u_2(y),
\end{aligned}$$

with equality when $x_\perp = A^T y$. Hence, for any element (x_0, y_0, z_0, w_0) in W, $\rho(x_0, y_0) = u_1(x_0) + u_2(y_0)$. Furthermore, by tracing the cases of equality in the preceding string of inequalities, it is not hard to show that (z_0, w_0) maximizes the map $(z,w) \mapsto y_0 \cdot z + z \cdot w + w \cdot (Fx_0) - u_3(z) - u_4(w)$. Then $b(x_0, y_0, z_0, w_0) - u_3(z_0) - u_4(w_0) = \rho(x_0, y_0)$; that is $b(x_0, y_0, z_0, w_0) - u_3(z_0) - u_4(w_0) = u_1(x_0) + u_2(y_0)$ on W, proving the claim. Since $x_\parallel$ and y can be chosen freely, $dim(W) = 2 + dim(null(A)) \geq 3$. Then, if we take any probability measure μ, concentrated on W and absolutely continuous with respect to $dim(W)$-dimensional Hausdorff measure, $spt(\mu)$ will not be concentrated on the graph of a measurable map. Now, take the projections of μ as marginals; that is, set $\mu_1 = (\pi_x)_\sharp\mu$, $\mu_2 = (\pi_y)_\sharp\mu$, $\mu_3 = (\pi_z)_\sharp\mu$ and $\mu_4 = (\pi_w)_\sharp\mu$. Inequality (3.14) together with the fact that equality holds μ almost everywhere, implies that μ is a solution to (KPb). It remains to show that these marginals are absolutely continuous with respect to $\mathcal{L}^n$. Since μ is absolutely continuous with respect to $dim(W)$-dimensional Hausdorff measure, it will suffice to show that the canonical projections π_x, π_y, π_z, π_w from the linear subspace W are surjective. We claim that M_1 and M_2 can be chosen so that this is the case. Indeed, since y can be chosen freely in the definition of W, it is immediate that π_y is surjective. Given any $x = x_\parallel + x_\perp$, we can find y such that $x_\perp = A^T y$, since A^T maps $\mathbb{R}^2$ onto its range, which is the orthogonal complement of the null space of A. Therefore, $(x, y, z = M_1(y + M_2 Fx), w = M_2(z + Fx)) \in W$, which implies π_x is surjective. Turning to π_z and π_w, note that $M_1^\lambda := \lambda M_1 > 0$ and $M_2^\lambda := \frac{1}{\lambda} M_2 > 0$, and that replacing M_1 and M_2 by M_1^λ and M_2^λ respectively does not change M (or therefore A). Set $Q_1^\lambda = M_1^\lambda + AA^T - A^T$ and $Q_2^\lambda = (M^{-1})^T + M_2^\lambda(AA^T - A^T + FA^T)$ and note that, for λ sufficiently large, both Q_1^λ and Q_2^λ are invertible. Now, let $z \in \mathbb{R}^2$. Choose y such that $Q_1^\lambda y = z$ and $x = x_\parallel + x_\perp = 0 + A^T y = A^T y$. Then $(x, y, M_1^\lambda(y + M_2^\lambda Fx), M_2^\lambda(M_1^\lambda(y + M_2^\lambda Fx) + Fx)) \in W$, and

$$\begin{aligned}
M_1^\lambda(y + M_2^\lambda Fx) &= M_1^\lambda(y + M_2^\lambda FA^T y) \\
&= (M_1^\lambda + M_1^\lambda M_2^\lambda FA^T)y \\
&= (M_1^\lambda + (A - I)A^T)y \\
&= (M_1^\lambda + AA^T - A^T)y \\
&= Q_1^\lambda y = z.
\end{aligned}$$

This establishes surjectivity of π_z. Similarly, for $w \in \mathbb{R}^2$, choosing y such that $w = Q_2^\lambda y$ and $x = x_\parallel + x_\perp = 0 + A^T y = A^T y$, we have that $(x, y, M_1^\lambda(y + M_2^\lambda Fx), M_2^\lambda(M_1^\lambda(y + M_2^\lambda Fx) + Fx)) \in W$, and

$$
\begin{aligned}
M_2^\lambda(M_1^\lambda(y + M_2^\lambda Fx) + Fx)) &= M_2^\lambda(Q_1^\lambda y + FA^T y) \\
&= M_2^\lambda(M_1^\lambda + AA^T - A^T + FA^T)y \\
&= ((M^{-1})^T + M_2^\lambda(AA^T - A^T + FA^T))y \\
&= Q_2^\lambda y = w.
\end{aligned}
$$

Therefore, π_w is surjective, completing the proof. □

3.B.3 The condition $m = 4$.

In this subsection we show that the hypothesis on the numbers of variables in Theorem 3.1 is necessary. We will follow the ideas behind the proof of Proposition 3.8.

In what follows, the presented variables are in $\mathbb{R}^2$. For a given $x_i \in \mathbb{R}^2$ its coordinates will be denoted by x_i^1 and x_i^2 respectively.

Proposition 3.9. *For $b(x_1, \ldots, x_m) = \sum_{i=1}^{m-1} x_i \cdot x_{i+1} + x_m \cdot x_1$, $m \geq 5$, there are absolutely continuous marginals μ_i with respect to $\mathcal{L}^2$, such that a solution of (KPb) is not concentrated on a graph.*

Proof. By part 2 of Theorem 3.6, $M = \begin{bmatrix} -1 & 3 \\ 0 & -1 \end{bmatrix} \in \Re^3$. Hence, we can choose positive definite matrices $M_1, M_2, M_3 > 0$ such that $M = M_1 M_2 M_3$. For all $x_1, \ldots, x_m$ define:

i. $u_1^m(x_1) = \frac{3(x_1^2)^2}{2} + g_1(x_1) + g_2^m(x_1)$, where $g_1(x_1) = \frac{1}{2}(M_3x_1)^T M_2(M_3x_1) + \frac{1}{2}(M_2M_3x_1)^T M_1(M_2M_3x_1)$ and $g_2^m(x_1) = \frac{1}{2}x_1^T M_3 x_1 + \frac{m-5}{2}|x_1|^2$ for all $m \geq 5$.

ii. $u_2(x_2) = \frac{3(x_2^1)^2}{2} + g(x_2)$ with $g(x_2) = \frac{1}{2}x_2^T M_1 x_2$, $u_3(x_3) = \frac{1}{2}x_3^T(M_1^{-1} + M_2)x_3$ and $u_4(x_4) = \frac{1}{2}x_4^T(M_2^{-1} + M_3)x_4$ for all $m \geq 5$.

iii. $u_5(x_5) = \frac{1}{2}x_5^T M_3^{-1} x_5$ for $m = 5$.

iv. For $m > 5$, $u_i(x_i) = \begin{cases} \frac{1}{2}x_5^T(M_3^{-1} + I)x_5 & \text{if } i = 5 \\ |x_i|^2 & \text{if } 5 < i < m \\ \frac{1}{2}|x_m|^2 & \text{if } i = m. \end{cases}$

v. $\rho^m(x_1, x_2) = \sup_{x_3,\dots,x_m} \left[b(x_1, \dots, x_m) - \sum_{i=3}^m u_i(x_i)\right]$ for all $m \geq 5$.

Consider the set:

$$W = \Big\{(x_1, \dots, x_m) : x_1^2 = x_2^1, \quad x_3 = M_1(x_2 + M_2M_3x_1), \quad x_4 = M_2(x_3 + M_3x_1),$$
$$x_5 = M_3(x_4 + x_1) \quad \text{and} \quad x_i = x_1 + x_{i-1}, \quad \text{for} \quad i \geq 6\Big\}.$$

We claim that for $m \geq 5$, $b(x_1, \dots, x_m) - \sum_{i=3}^m u_i(x_i) \leq u_1^m(x_1) + u_2(x_2)$, for all $(x_1, \dots, x_m) \in (\mathbb{R}^2)^m$, and equality holds on W. For the inequality, it suffices to prove $\rho^m(x_1, x_2) \leq u_1^m(x_1) + u_2(x_2)$ for all $m \geq 5$. We divide the proof of the claim into two cases:

1. For m=5

$$\begin{aligned}
\rho^5(x_1, x_2) &= x_1 \cdot x_2 + \sup_{x_3,x_4,x_5} [x_2 \cdot x_3 + x_3 \cdot x_4 + x_4 \cdot x_5 + x_5 \cdot x_1 - u_3(x_3) - u_4(x_4) - u_5(x_5)] \\
&= x_1 \cdot x_2 + \sup_{x_3} \Big[x_2 \cdot x_3 - u_3(x_3) + \sup_{x_4} \big[x_3 \cdot x_4 - u_4(x_4) \\
&\quad + \sup_{x_5}[(x_4 + x_1)x_5 - u_5(x_5)]\big]\Big] \\
&= x_1 \cdot x_2 + \sup_{x_3} \big[x_2 \cdot x_3 - u_3(x_3) + \sup_{x_4}[x_3 \cdot x_4 - u_4(x_4) + u_5^*(x_4 + x_1)]\big] \\
&= x_1 \cdot x_2 + \sup_{x_3} \big[x_2 \cdot x_3 - u_3(x_3) + \sup_{x_4}[x_3 \cdot x_4 - u_4(x_4) \\
&\quad + \frac{1}{2}(x_4 + x_1)^T M_3(x_4 + x_1)]\big] && \text{by (3.13)} \\
&= x_1 \cdot x_2 + \sup_{x_3} \big[x_2 \cdot x_3 - u_3(x_3) + \sup_{x_4}[x_3 \cdot x_4 - \frac{1}{2}x_4^T M_2^{-1} x_4 \\
&\quad + x_4^T M_3 x_1 + g_2^5(x_1)]\big]
\end{aligned}$$

$$
\begin{aligned}
&= x_1 \cdot x_2 + \sup_{x_3} \left[x_2 \cdot x_3 - u_3(x_3) + \frac{1}{2}(x_3 + M_3 x_1)^T M_2 (x_3 + M_3 x_1) \right] \\
&\quad + g_2^5(x_1) && \text{by (3.13)} \\
&= x_1 \cdot x_2 + \sup_{x_3} \left[x_2 \cdot x_3 - \frac{1}{2} x_3^T M_1^{-1} x_3 + (M_3 x_1)^T M_2 x_3 \right] + g_2^5(x_1) \\
&\quad + \frac{1}{2}(M_3 x_1)^T M_2 (M_3 x_1) \\
&= x_1 \cdot x_2 + \sup_{x_3} \left[x_2 \cdot x_3 - u_3(x_3) + \frac{1}{2}(x_3 + M_3 x_1)^T M_2 (x_3 + M_3 x_1) \right] + g_2^5(x_1) \\
&= x_1 \cdot x_2 + \frac{1}{2}(x_2 + M_2 M_3 x_1)^T M_1 (x_2 + M_2 M_3 x_1) + g_2^5(x_1) \\
&\quad + \frac{1}{2}(M_3 x_1)^T M_2 (M_3 x_1) && \text{by (3.13)} \\
&= x_1 \cdot x_2 + x_2^T M x_1 + g_1(x_1) + g_2^5(x_1) + g(x_2) \\
&= 3 x_1^2 x_2^1 + g_1(x_1) + g_2^5(x_1) + g(x_2) \\
&\leq \frac{3(x_1^2)^2}{2} + \frac{3(x_2^1)^2}{2} + g_1(x_1) + g_2^5(x_1) + g(x_2) && \text{by the Cauchy-Schwarz Ineq.} \\
&= u_1^5(x_1) + u_2(x_2).
\end{aligned}
$$

2. The case $m \geq 6$ will be proved using induction. For $m = 6$, note that:

$$
x_4 \cdot x_5 - u_5(x_5) + \sup_{x_6}[(x_5 + x_1) x_6 - u_6(x_6)] = (x_1 + x_4) x_5 - \frac{1}{2} x_5^T M_3^{-1} x_5 + \frac{1}{2}|x_1|^2
$$

and

$$
\sup_{x_5}[(x_1 + x_4) x_5 - \frac{1}{2} x_5^T M_3^{-1} x_5 + \frac{1}{2}|x_1|^2] = \frac{1}{2}(x_4 + x_1)^T M_3 (x_4 + x_1) + \frac{1}{2}|x_1|^2.
$$

Then

$$\rho^6(x_1,x_2) = \rho^5(x_1,x_2) + \frac{1}{2}|x_1|^2$$
$$\leq u_1^5(x_1) + u_2(x_2) + \frac{1}{2}|x_1|^2$$
$$= u_1^6(x_1) + u_2(x_2). \tag{3.15}$$

Assume the statement is true for $m-1$. Then

$$\rho^m(x_1,x_2) = \sup_{x_3,\dots,x_m}\Big[\sum_{i=1}^{m-1} x_i\cdot x_{i+1} + x_m\cdot x_1 - \sum_{i=3}^{m} u_i(x_i)\Big]$$
$$= \sup_{x_3,\dots,x_{m-1}}\Big[\sum_{i=1}^{m-2} x_i\cdot x_{i+1} - \sum_{i=3}^{m-1} u_i(x_i) + \sup_{x_m}\big[(x_1+x_{m-1})x_m - u_m(x_m)\big]\Big]$$
$$= \sup_{x_3,\dots,x_{m-1}}\Big[\sum_{i=1}^{m-2} x_i\cdot x_{i+1} + \frac{|x_1+x_{m-1}|^2}{2} - \sum_{i=3}^{m-1} u_i(x_i)\Big] \qquad \text{by (3.13)}$$
$$= \sup_{x_3,\dots,x_{m-1}}\Big[b(x_1,\dots,x_{m-1}) - \sum_{i=3}^{m-2} u_i(x_i) - \frac{|x_{m-1}|^2}{2} + \frac{|x_1|^2}{2}\Big]$$
$$= \rho^{m-1}(x_1,x_2) + \frac{|x_1|^2}{2} \tag{3.16}$$
$$\leq u_1^{m-1}(x_1) + u_2(x_2) + \frac{|x_1|^2}{2} \qquad \text{by induction hypothesis}$$
$$= u_1^m(x_1) + u_2(x_2).$$

If $x_1^2 = x_2^1$, we obtain $\rho^5(x_1,x_2) = u_1^5(x_1)+u_2(x_2)$ and by (3.15), $\rho^6(x_1,x_2) = u_1^6(x_1)+u_2(x_2)$. Furthermore, by (3.16), $\rho^m(x_1,x_2) = \rho^{m-1}(x_1,x_2) + \frac{|x_1|^2}{2}$. Hence, using induction we can easily prove that $\rho^m(x_1,x_2) = u_1^m(x_1)+u_2(x_2)$ for all $m \geq 5$, when $x_1^2 = x_2^1$.

On the other hand, for any element $(\bar{x}_1,\dots,\bar{x}_m)$ in W, $(\bar{x}_3,\dots,\bar{x}_m)$ maximizes the map:

$$(x_3,\dots,x_m) \mapsto \bar{x}_2\cdot x_3 + \sum_{i=3}^{m-1} x_i\cdot x_{i+1} + x_m\cdot\bar{x}_1 - \sum_{i=3}^{m} u_i(x_i).$$

Hence $b(\bar{x}_1, \ldots, \bar{x}_m) - \sum_{i=3}^m u_i(\bar{x}_i) = \rho^m(\bar{x}_1, \bar{x}_2) = u_1^m(\bar{x}_1) + u_2(\bar{x}_2)$. This proves the claim.

Since x_1^1 and $x_2 = (x_2^1, x_2^2)$ can be chosen freely, W is three dimensional, and the claim implies that any probability measure μ supported on W is optimal for its marginals in (KPb); such solutions are manifestly not of Monge type. It remains to show that the marginals $\mu_i = (\pi_{x_i})_\sharp \mu$ of μ can be taken to be absolutely continuous. As in the proof of Proposition 3.8, this will be the case when μ is absolutely continuous with respect to 3-dimensional Hausdorff measure on W, provided that the canonical projections π_{x_i} from W are surjective; we prove this now.

It is clear that π_{x_1} and π_{x_2} are surjective. Turning to the other projections, set $x_1 = \begin{bmatrix} x_1^1 \\ 0 \end{bmatrix}$, $x_2 = \begin{bmatrix} 0 \\ x_2^2 \end{bmatrix}$ and write $M_k = \begin{bmatrix} m_k^1 & m_k^2 \\ m_k^2 & m_k^3 \end{bmatrix}$, $k = 1, 2, 3$.

Then $(x_1, \ldots, x_m) \in W$ where $x_3 = M_1(x_2 + M_2 M_3 x_1)$, $x_4 = M_2(x_3 + M_3 x_1)$, $x_5 = M_3(x_4 + x_1)$ and $x_i = x_1 + x_{i-1}$ for $i \geq 6$. Note that:

$$
\begin{aligned}
x_4 &= M_2 x_3 + M_2 M_3 x_1 \\
&= M_2 M_1 (x_2 + M_2 M_3 x_1) + M_2 M_3 x_1 \\
&= M_2 M_1 x_2 + (M_2 M_1 M_2 M_3 + M_2 M_3)\, x_1,
\end{aligned}
$$

or equivalently,

$$
x_4 = M_3^{-1}\Big(M^T x_2 + \begin{bmatrix} a & b \\ b & d \end{bmatrix} x_1\Big) = M_3^{-1} \begin{bmatrix} a x_1^1 \\ -x_2^2 + b x_1^1 \end{bmatrix}, \tag{3.17}
$$

where $\begin{bmatrix} a & b \\ b & d \end{bmatrix} := M_3 M_2 M_1 M_2 M_3 + M_3 M_2 M_3$ is symmetric and positive definite.

$$
x_5 = M_3 x_4 + M_3 x_1 = \begin{bmatrix} a x_1^1 \\ -x_2^2 + b x_1^1 \end{bmatrix} + \begin{bmatrix} m_3^1 x_1^1 \\ m_3^2 x_1^1 \end{bmatrix} = \begin{bmatrix} (a + m_3^1) x_1^1 \\ -x_2^2 + (b + m_3^2) x_1^1 \end{bmatrix} \tag{3.18}
$$

Finally, for $i \geq 6$ we get

$$x_i = x_1 + x_{i-1} = (i-5)x_1 + x_5 = \begin{bmatrix} (i-5+a+m_3^1)x_1^1 \\ -x_2^2 + (b+m_3^2)x_1^1 \end{bmatrix}. \tag{3.19}$$

Clearly, since $M_3 > 0$ and $\begin{bmatrix} a & b \\ b & d \end{bmatrix} > 0$ we have $a, a+m_3^1, i-5+a+m_3^1 > 0$ (when $i \geq 6$ for the last inequality), implying that the equations (3.17), (3.18) and (3.19) each have a solution for arbitrary choice of x_4, x_5 and x_i, $i \geq 6$, respectively, and so the projections π_{x_i}, $i \geq 4$ are surjective. Finally, we turn to the equation for x_3:

$$x_3 = M_1 x_2 + M x_1 = \begin{bmatrix} m_1^2 x_2^2 - x_1^1 \\ m_1^3 x_2^2 \end{bmatrix},$$

which has a solution for every fixed $x_3 \in \mathbb{R}^2$, as $m_1^3 > 0$. Thus, π_{x_3} is surjective. This completes the proof of the proposition. □

Chapter 4

Multi-marginal optimal transport via graph theory

In this chapter we study a multi-marginal optimal transport problem with cost

$$\sum_{\{i,j\}\in P} |x_i - x_j|^2, \tag{4.1}$$

where $P \subseteq Q := \{\{i,j\} : i,j \in \{1,2,...m\}, i \neq j\}$ (cost (4.1) reduces to (1.1) when $P = Q$, and (1.2) with $F = I$, if $P = \big\{\{i, i+1\} : i = 1, \dots, m-1\big\} \cup \big\{\{1, m\}\big\}$). We reformulate this problem using graph theory; in particular, we associate each cost of the form (4.1) to a graph with m vertices whose set of edges is indexed by P. For instance, the Gangbo-Święch cost (1.1) is associated to the complete graph K_m and the cyclic cost (1.2) to the cycle graph. See Figures (1.1) and (1.2) for the case $m = 7$.

Instead of minimizing (KP), we will trent the equivalent problem of maximizing:

$$\int_X b(x_1, \dots, x_m) d\mu \tag{KPG}$$

where

$$b(x_1, \ldots, x_m) = \sum_{\{i,j\} \in P} x_i \cdot x_j, \tag{4.2}$$

over the same admissible class of (KP).

4.A Some graph theory and preliminary results

First, let us recall some definitions from Graph Theory. An *undirected simple graph* G is an ordered pair $(V(G), E(G))$, consisting of a finite *set of vertices* $V(G)$ and a *set of edges* $E(G) \subseteq \{\{v, w\} : v, w \in V(G) \text{ and } v \neq w\}$. Throughout this work, every graph G is an undirected simple graph. A *trail* is a finite sequence $\{\{v_{i_1}, v_{i_2}\}, \{v_{i_2}, v_{i_3}\}, \ldots, \{v_{i_l}, v_{i_{l+1}}\}\}$ of pairwise distinct edges which joins a sequence of vertices. A *path* is a trail in which all vertices are distinct: $v_{i_j} \neq v_{i_k}$ for all $j \neq k$. A *cycle graph* is a trail in which the first and last vertex are the only one repeated: $\{\{v_{i_1}, v_{i_2}\}, \{v_{i_2}, v_{i_3}\}, \ldots, \{v_{i_l}, v_{i_{l+1}}\}\}$, where $\{\{v_{i_1}, v_{i_2}\}, \{v_{i_2}, v_{i_3}\}, \ldots, \{v_{i_{l-1}}, v_{i_l}\}\}$ is a path and $v_{i_{l+1}} = v_{i_1}$. A *tree* is a graph where any two distinct vertices are connected by a unique path. A graph G is *connected* if for every $v, w \in V(G)$ there exists a path in the graph joining them. We will denote by $I(V(G))$ the *set of indices* of $V(G)$ (that is, for $V(G) = \{v_1, ..., v_m\}$, $I(V(G)) = \{1, 2, ..., m\}$) and $|V(G)|$ the cardinality of $V(G)$.

A *subgraph* S of a graph G is a graph whose sets of vertices and edges are subsets of $V(G)$ and $E(G)$ respectively. In this case, we call the graph $G \setminus S := (V(G), E(G) \setminus E(S))$ the *extraction* of S from G. Note that if G is complete and $V(G) = V(S)$, $G \setminus S$ coincides with the complement of S; that is, $G \setminus S = (V(S), E(S^c))$, where $E(S^c) := \{\{v, w\} : v, w \in V(S) \ \ and \ \ \{v, w\} \notin E(S)\}$.

Given $v, w \in V(G)$, v and w are called *adjacent* if $\{v, w\} \in E(G)$. The *open neighborhood* of a vertex v, denoted $N_G(v)$ (or simply $N(v)$ if there is not danger of confusion), is the set of vertices that are adjacent to v; that is,

$$N(v) = \{w \in V(G) : \{v, w\} \in E(G)\}.$$

The *closed neighborhood* of a vertex v, denoted $\overline{N}_G(v)$(or simply $\overline{N}(v)$), is the set $N(v) \cup \{v\}$.

A graph G is *complete* if $N(v) = V(G) \setminus \{v\}$ for every $v \in V(G)$. A *clique* $\tilde{G} = (V(\tilde{G}), E(\tilde{G}))$ of a graph $G = (V(G), E(G))$ is a complete subgraph of G; that is, $\tilde{G}$ is a subgraph of G and satisfies $N_{\tilde{G}}(v) = V(\tilde{G}) \setminus \{v\}$ for every $v \in V(\tilde{G})$. A clique is *maximal* if it is not a proper subgraph of any other clique of G.

The *union* of two given graphs G_1 and G_2 (denoted by $G_1 \cup G_2$) is the graph with set of vertices $V(G_1) \cup V(G_2)$ and edges $E(G_1) \cup E(G_2)$. A *complete k-partite graph* G is a graph whose set of vertices $V(G)$, can be partitioned into k subsets $V_1, V_2, \ldots, V_k$ such that for every $v \in V_j$, $N(v) = \bigcup_{\substack{\alpha=1 \\ \alpha\neq j}}^{k} V_\alpha$ for any fixed $j \in \{1, \ldots, k\}$. A complete k-partite graph G is denoted as $K_{m_1,\ldots,m_k}$, where $|V_j| = m_j$ for every $j \in \{1, \ldots, k\}$.

Remark 4.1. *Note that the definition of a graph implies $v \notin N(v)$ for every $v \in V(G)$. Furthermore, from now on, if S is a subgraph of G and $v \in V(S)$, we will write $N(v)$ for the open neighborhood of v in G; that is, $N(v) = N_G(v)$. Similarly, $\overline{N}(v) = \overline{N}_G(v)$.*

The next two concepts are used to facilitate the description of our main results, established in Theorem 4.9 and 4.12.

Definition 4.2. *We say that a subset A of $V(G)$ is the* inner hub *of G, if $A = V(S_1) \bigcap V(S_2)$ for any two maximal cliques S_1 and S_2 of G.*

Example 4.3. *The picture below shows the graph $G := S_1 \cup S_2 \cup S_3$, where S_1, S_2 and S_3 are complete graphs with $V(S_1) = \{v_6, v_7, v_8, v_9\}$, $V(S_2) = \{v_6, v_7, v_8, v_4, v_5\}$ and $V(S_3) = \{v_6, v_7, v_8, v_1, v_2, v_3, v_{10}\}$. Clearly, $\{S_1, S_2, S_3\}$ is the collection of* *maximal cliques of G, and the inner hub of G is the set formed by the vertices of the triangle colored blue; that is, $A = \{v_6, v_7, v_8\}$.*

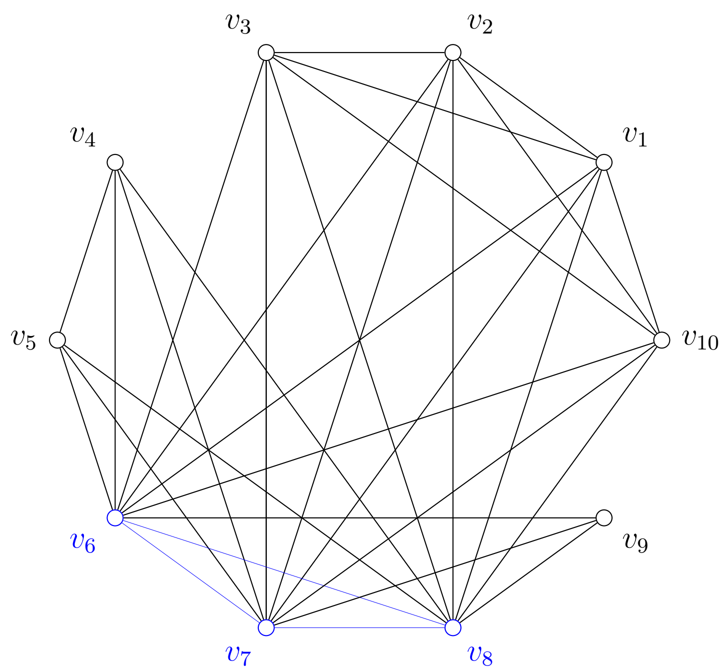


Figure 4.1: Graph $G = S_1 \cup S_2 \cup S_3$.

Not all graphs have an inner hub. Letting $\{S_j\}_{j=1}^{l}$ be the maximal cliques of a graph G, it is clear that $G = \bigcup_{j=1}^{l} S_j$; A is the inner hub of G if $V(S_j) \cap V(S_k) = A$ for all $j \neq k$, $j, k \in \{1, \ldots, l\}$. Note that we allow the inner hub A to be the empty set, which is the case when G is disconnected and each connected component is complete. At the other extreme, we could have $A = G$, which is the case when G is complete.

Definition 4.4. *Let $\{S_{1j}\}_{j=1}^{l_1}$ and $\{S_{2j}\}_{j=1}^{l_2}$ be the collection of maximal cliques of given graphs G_1 and G_2, respectively. Assume that G_1 and G_2 have inner hubs A_1 and A_2, respectively. We say the graphs G_1 and G_2 are glued on a clique if:*

1. *There are $j \in \{1, \ldots, l_1\}$ and $k \in \{1, \ldots, l_2\}$ such that $S_{1j} = S_{2k}$.*

2. $V(G_1) \cap V(G_2) = V(S_{1j})$.

Example 4.5. *The picture below shows the graph $S_1 \cup S_2 \cup S_3$, where $\{S_1, S_2, S_3\}$ is the collection of maximal cliques in Example 4.3, glued with the graph $S'_1 \cup S'_2 \cup S'_3 \cup S'_4$ on S_2, where S'_1, S'_2, S'_3 and S'_4 are complete graphs with $V(S'_1) = \{v_4, v_5, v_{14}\}$, $V(S'_2) = \{v_6, v_7, v_8, v_4, v_5\} = V(S_2)$, $V(S'_3) = \{v_4, v_5, v_{15}, v_{16}\}$ and $V(S'_4) = \{v_4, v_5, v_{11}, v_{12}, v_{13}\}$. Note that the inner hub of $S'_1 \cup S'_2 \cup S'_3 \cup S'_4$, whose collection of maximal cliques is $\{S'_j\}_{j=1}^4$, is formed by the vertices of the edges colored red; that is, $A' = \{v_4, v_5\}$.*

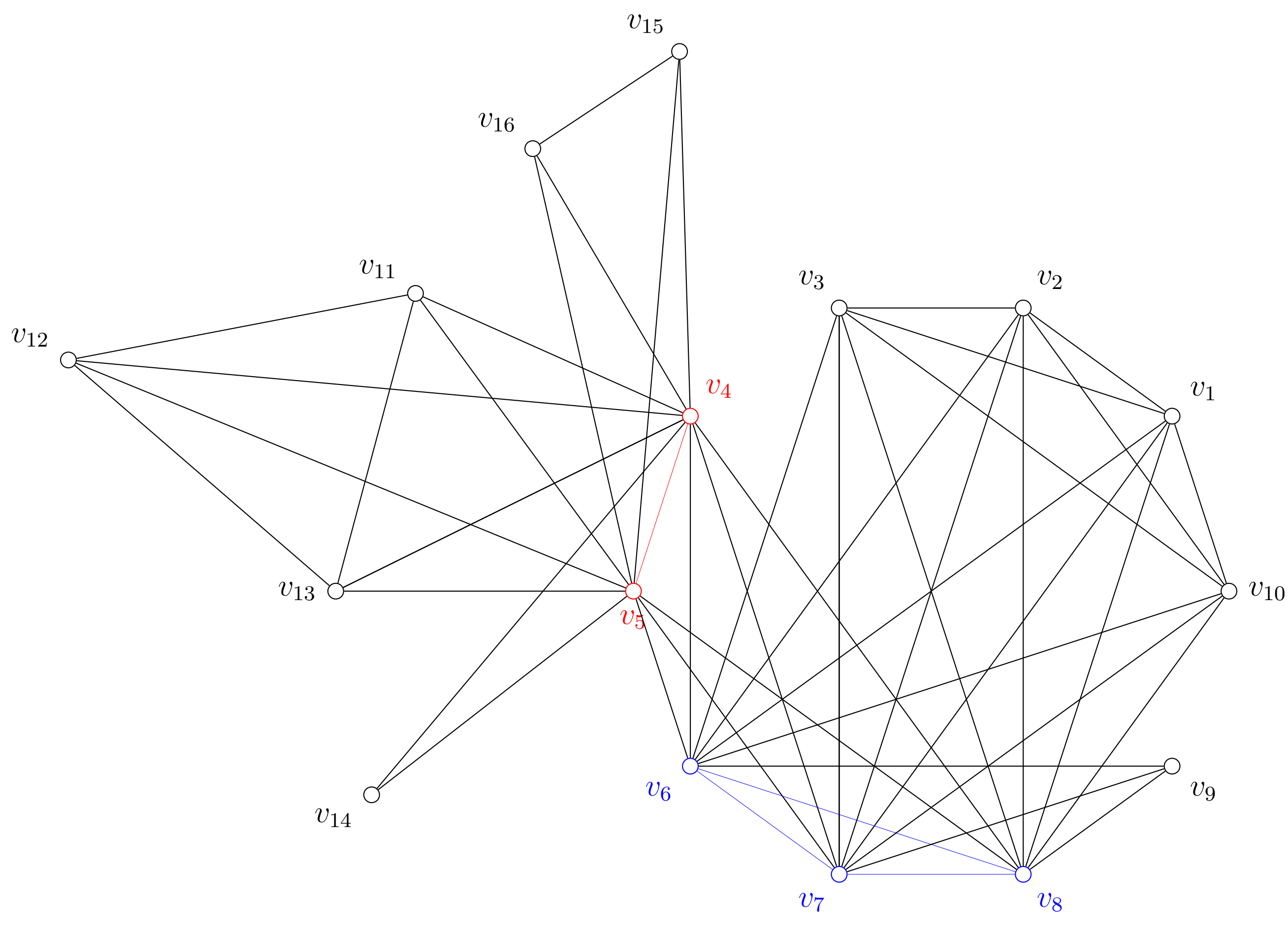


Figure 4.2: Graph $\left(\bigcup_{j=1}^3 S_j\right) \cup \left(\bigcup_{j=1}^4 S'_j\right)$.

4.A.1 Preliminary results connecting graph theory and multi-marginal optimal transport

In this subsection, we establish some initial results connecting solutions of the multi-marginal optimal transport problem (KPG) and the structure of the corresponding graph. These include a couple of very basic observations (Proposition 4.6), as well as a technical lemma which will be used throughout the paper (Lemma 4.8).

Proposition 4.6. *Let G be the graph corresponding to some $P \subseteq Q$ and b the suplus* (4.2).

1. *Assume G is not connected and let x_i be any vertex such that there is no path between x_1 and x_i, and assume that μ_i is not a dirac mass. Then there exist non Monge solutions to* (KPG)*, and, if μ_1 is not a dirac mass, the solution to* (KPG) *is not unique.*

2. *Assume $\{v_1, v_i\}$ is not an edge of G for some i, and all the marginals are dirac measures except μ_1 and μ_i, with μ_1 absolutely continuous with respect to $\mathcal{L}^n$. Then, there exist solutions of non-Monge form to* (KPG) *and the solution to* (KPG) *is not unique.*

Proof. Consider the first assertion. Let G_1 be the connected component of G satisfying $v_1 \in Z := V(G_1)$, and G_2 the graph union of the other components of G, with $W := V(G_2)$. Then the surplus (4.2) takes the separable form:

$$b(x_1, \ldots, x_m) = b_Z(x_Z) + b_W(x_W),$$

where we decompose $x = (x_Z, x_W)$ into components x_Z and x_W whose indices of their coordinates lie in $I(Z)$ and $I(W)$, respectively, and $b_Z(x_Z) = \sum_{\{v_s,v_t\}\in E(G_1)} x_s \cdot x_t$, $b_W(x_W) = \sum_{\{v_s,v_t\}\in E(G_2)} x_s \cdot x_t$. Solutions to (KPG) are then exactly measures μ whose projections μ_Z and μ_W onto the appropriate subspaces are optimal for the multi-marginal optimal transport problem with costs b_Z

and b_W, respectively, and the appropriate marginals. In particular, the dependence structure between μ_Z and μ_W is completely arbitrary, and so, if μ_i is not a dirac mass for some $v_i \in W$, we immediately get the existence of non-Monge solutions (for instance, the product measure $\mu_Z \otimes \mu_W$), and if in addition μ_1 is not a dirac mass, solutions are non-unique.

Turning to assertion 2, without loss of generality, assume $\{v_1, v_2\}$ is not an edge of G. Take μ_1 be absolutely continuous with respect to $\mathcal{L}^n$, μ_2 be any measure other than a Dirac mass (so that μ_2 charges at least two points) and let all other marginals be Dirac masses, $\mu_i = \delta_{\bar{x}_i}$. In this case, measures μ whose marginals are the μ_i all take the form $\mu = \sigma(x_1, x_2) \otimes \delta_{\bar{x}_3} \otimes \ldots \otimes \delta_{\bar{x}_m}$, where $\sigma \in P(X_1 \times X_2)$ has marginals μ_1 and μ_2. For any such μ, we have

$$\begin{aligned} \int_{X_1 \times X_2 \times \ldots \times X_m} b(x_1, \ldots, x_m) d\mu(x_1, \ldots x_m) &= \int_{X_1 \times X_2} b(x_1, x_2, \bar{x}_3, \ldots, \bar{x}_m) d\sigma(x_1, x_2) \\ &= \int_{X_1} b_1(x_1, \bar{x}_3, \ldots, \bar{x}_m) d\mu_1(x_1) \\ &\quad + \int_{X_2} b_2(x_2, \bar{x}_3, \ldots, \bar{x}_m) d\mu_2(x_2) \end{aligned}$$

where $b_1(x_1, x_3, \ldots, x_m) = \sum_{\{v_s, v_t\} \in E(G)\ s,t \neq 2} x_s \cdot x_t$ and $b_2(x_2, x_3, \ldots, x_m) = \sum_{s \in I(N(v_2))} x_2 \cdot x_s$. Thus, the Kantorovich functional is independent of σ, and so any σ with marginals μ_1 and μ_2 is optimal. We conclude that solutions are non-unique, and can be of non-Monge form (as is the case when, for example, $\sigma = \mu_1 \times \mu_2$ is the product measure). □

Corollary 4.7. *Under the hypothesis in any of assertion 1 or assertion 2 of Proposition 4.6, the surplus b is not twisted on b-splitting sets.*

Proof. Suppose b is twisted on b-splitting sets. From Remark 2.6 every solution to (KP) is induced by map; that is, every solution to (KP) is of Monge type. This clearly contradicts Proposition 4.6, completing the proof of the corollary. □

Clearly, in light of the first assertion there is no hope of obtaining Monge solution results for disconnected graphs (except in the trivial case when each μ_i with x_i not connected to μ_1 is a dirac mass, in which case the problem reduces to a problem on the connected component containing x_1.) We therefore will focus on connected graphs throughout this paper. On the other hand, our work in [48] suggests that at least for some surplus functions where $\{v_1, v_i\}$ is not an edge of G for some i, unique Monge solutions may exist when extra regularity conditions on the marginals are imposed (even though the twist on splitting sets condition fails). Our results in the following sections confirm that this is indeed the case.

The proofs of our main results will require the following technical lemma.

Lemma 4.8. *Let G be a graph, with $V(G) = \{v_1, \dots, v_m\}$, and $b(x_1, \dots, x_m) = \sum_{\{v_s,v_t\}\in E(G)} x_s \cdot x_t$ be the surplus associated to G. Let $(u_1, \dots, u_m)$ a b-conjugate m-tuple. Set*

$$W := \left\{ (x_1, \dots, x_m) \in X_1 \times \dots \times X_m : \sum_{i=1}^{m} u_i(x_i) = b(x_1, \dots, x_m) \right\}.$$

Fix $x_1^0 \in X_1$ and for convenience of notation set $x_1^1 = x_1^2 = x_1^0$. Let $(x_1^1, x_2^1, \dots, x_m^1), (x_1^2, x_2^2, \dots, x_m^2) \in W$.

1. *Assume there are sets $V_1, V_2 \subseteq V(G)$ such that $N(v_s) = V_2$ for every $s \in I(V_1)$, and set*

$$y_s = \begin{cases} x_s^1 & \text{if} \quad s \in \{1, \dots, m\} \setminus I(V_1) \\ x_s^2 & \text{if} \quad s \in I(V_1). \end{cases}$$

   If

$$\sum_{s\in I(V_2)} x_s^1 = \sum_{s\in I(V_2)} x_s^2, \tag{4.3}$$

   then $y := (y_1, y_2, \dots, y_m) \in W$.

2. *For all* $t \in \{1, \ldots, m\}$ *we have*

$$\big(x_t^2 - x_t^1\big) \cdot \sum_{s \in I(N(v_t))} \big(x_s^1 - x_s^2\big) \leq 0. \tag{4.4}$$

3. *If there exists* $t \in \{1, \ldots, m\}$ *such that*

$$\sum_{s \in I(\overline{N}(v_t))} x_s^1 = \sum_{s \in I(\overline{N}(v_t))} x_s^2, \tag{4.5}$$

   then $x_t^1 = x_t^2$.

4. *Assume* $x_p^1 = x_p^2$ *and* $Du_p(x_p^1)$ *exists for some* $p \in \{1, \ldots, m\}$.

   (a) *For every* $t \in \{2, \ldots, m\} \setminus \{p\}$ *satisfying*

$$\overline{N}(v_p) = \overline{N}(v_t), \tag{4.6}$$

   we have $x_t^1 = x_t^2$.

   (b) *Assume there are sets* F_1, F_2, F_3 *such that* $F_1, F_2 \subseteq N(v_p)$ *and* $N(v_s) = F_2 \cup F_3$ *for every* $s \in I(F_1)$. *If* $x_s^1 = x_s^2$ *for every* $s \in I(N(v_p) \setminus F_1 \cup F_2) \cup I(F_3)$, *then* $x_s^1 = x_s^2$ *for every* $s \in I(F_1)$.

Proof. Since for every $s \in I(V_1)$ we have $N(v_s) = V_2$, and $v \notin N(v)$ for all $v \in V(G)$, we get $V_1 \cap V_2 = \emptyset$. Hence, we can write

$$b(x_1, \ldots, x_m) = g(x_1, \ldots, x_m) + \Big(\sum_{s \in I(V_2)} x_s\Big) \cdot \Big(\sum_{s \in I(V_1)} x_s\Big),$$

where $g(x_1, \ldots, x_m)$ does not depend on $\{x_s\}_{s \in I(V_1)}$. Hence

$$\{x_s^2\}_{s \in I(V_1)} \in \text{Argmax}\Big\{\{x_s\}_{s \in I(V_1)} \mapsto \Big(\sum_{s \in I(V_2)} x_s^2\Big) \cdot \Big(\sum_{s \in I(V_1)} x_s\Big) - \sum_{s \in I(V_1)} u_s(x_s)\Big\},$$

as $(x_1^2, x_2^2, \ldots, x_m^2) \in W$. Then, If (4.3) holds we get

$$\{x_s^2\}_{s\in I(V_1)} \in \text{Argmax}\Big\{\{x_s\}_{s\in I(V_1)} \mapsto \Big(\sum_{s\in I(V_2)} x_s^1\Big)\cdot\Big(\sum_{s\in I(V_1)} x_s\Big) - \sum_{s\in I(V_1)} u_s(x_s)\Big\},$$

which implies $y \in W$, as $(x_1^1, x_2^1, \ldots, x_m^1) \in W$. This completes the proof of part 1. Using the arguments of the previous proof, and taking $V_1 = \{v_t\}$ and $V_2 = N(v_t)$ for any fixed $t \in \{2, \ldots, m\}$, we deduce

$$x_t^2 \in \text{Argmax}\Big\{x_t \mapsto \Big(\sum_{s\in I(N(v_t))} x_s^2\Big)\cdot x_t - u_t(x_t)\Big\}.$$

Similarly,

$$x_t^1 \in \text{Argmax}\Big\{x_t \mapsto \Big(\sum_{s\in I(N(v_t))} x_s^1\Big)\cdot x_t - u_t(x_t)\Big\}.$$

Then,

$$\Big(\sum_{s\in I(N(v_t))} x_s^2\Big)\cdot x_t^1 - u_t(x_t^1) \le \Big(\sum_{s\in I(N(v_t))} x_s^2\Big)\cdot x_t^2 - u_t(x_t^2) \tag{4.7}$$

and

$$\Big(\sum_{s\in I(N(v_t))} x_s^1\Big)\cdot x_t^2 - u_t(x_t^2) \le \Big(\sum_{s\in I(N(v_t))} x_s^1\Big)\cdot x_t^1 - u_t(x_t^1). \tag{4.8}$$

Adding (4.7) and (4.8) (and eliminating terms) we obtain inequality (4.4), completing the proof of the second part.

The proof of part 3 follows immediately from part 2, as if there exists $t \in \{2, \ldots, m\}$ satisfying (4.5) we get

$$x_t^1 + \sum_{s\in I(N(v_t))} x_s^1 = x_t^2 + \sum_{s\in I(N(v_t))} x_s^2,$$

hence, $\sum_{s\in I(N(v_t))} (x_s^1 - x_s^2) = x_t^2 - x_t^1$. Substituting it into inequality (4.4) we get $\|x_t^2 - x_t^1\|^2 \le 0$; that is, $x_t^2 = x_t^1$. To prove part 4, first note that

$$\sum_{s\in I(N(v_p))} x_s^1 = D_{x_p}b(x_1^1,\ldots,x_m^1) = Du_p(x_p^1) = Du_p(x_p^2) = D_{x_p}b(x_1^2,\ldots,x_m^2) = \sum_{s\in I(N(v_p))} x_s^2. \tag{4.9}$$

for any $t\in\{2,\ldots,m\}\setminus\{p\}$ satisfying (4.6) we obtain

$$\begin{aligned}\sum_{s\in I(\overline{N}(v_t))} x_s^1 &= \sum_{s\in I(\overline{N}(v_p))} x_s^1\\ &= \sum_{s\in I(\overline{N}(v_p))} x_s^2 && \text{by (4.9) and the equality } x_p^1 = x_p^2.\\ &= \sum_{s\in I(\overline{N}(v_t))} x_s^2\end{aligned}$$

Then, by part 3 we conclude $x_t^1 = x_t^2$, completing the proof of part 4a. To prove part 4b observe that $F_1\cap F_2=\emptyset$, as $N(v_s)=F_2\cup F_3$ and $v_s\notin N(v_s)$ for every $s\in I(F_1)$. Then, from (4.9) we get

$$\sum_{s\in I(F_1)} x_s^1 + \sum_{s\in I(F_2)} x_s^1 + \sum_{s\in I(N(v_p)\setminus F_1\cup F_2)} x_s^1 = \sum_{s\in I(F_1)} x_s^2 + \sum_{s\in I(F_2)} x_s^2 + \sum_{s\in I(N(v_p)\setminus F_1\cup F_2)} x_s^2,$$

as $F_1,F_2\subseteq N(v_p)$. Since $x_s^1 = x_s^2$ for every $s\in I(N(v_p)\setminus F_1\cup F_2)$, the above equality reduces to

$$\sum_{s\in I(F_1)} x_s^1 + \sum_{s\in I(F_2)} x_s^1 = \sum_{s\in I(F_1)} x_s^2 + \sum_{s\in I(F_2)} x_s^2, \tag{4.10}$$

and applying part 2 we get

$$\left(x_t^2 - x_t^1\right)\cdot \sum_{s\in I(F_2\cup F_3)} \left(x_s^1 - x_s^2\right) \le 0,$$

for any $t\in I(F_1)$. Summing over $t\in I(F_1)$ and using the equalities $x_s^1 = x_s^2$ on

$I(F_3)$, we obtain

$$\sum_{t\in I(F_1)} \left(x_t^2 - x_t^1\right) \cdot \sum_{s\in I(F_2)} \left(x_s^1 - x_s^2\right) \leq 0.$$

Furthermore, by (4.10) we get $\sum_{t\in I(F_1)} (x_t^2 - x_t^1) = \sum_{s\in I(F_2)} (x_s^1 - x_s^2)$. Substituting it into the above inequality we get

$$\| \sum_{s\in I(F_2)} \left(x_s^1 - x_s^2\right) \|^2 \leq 0; \quad \text{that is,}$$

$$\sum_{s\in I(F_2)} x_s^1 = \sum_{s\in I(F_2)} x_s^2. \tag{4.11}$$

Now, fix $t \in I(F_1)$ and set $V_1 = F_1 \setminus \{v_t\}$, $V_2 = F_2 \cup F_3$ and $y = (y_1, y_2, \ldots, y_m)$ such that

$$y_s = \begin{cases} x_s^1 & \text{if} \quad s \in \{1, \ldots, m\} \setminus I(V_1) \\ x_s^2 & \text{if} \quad s \in I(V_1). \end{cases}$$

Since $x_s^1 = x_s^2$ on $I(F_3)$, (4.11) can be written as $\sum_{s\in I(F_2\cup F_3)} x_s^1 = \sum_{s\in I(F_2\cup F_3)} x_s^2$. Therefore, by part 1 we get $y \in W$, as $N(v_s) = V_2$ for every $s \in I(V_1)$. Hence,

$$\sum_{s\in I(N(v_p))} y_s = Du_p(y_p) = Du_p(x_p^1) = Du_p(x_p^2) = \sum_{s\in I(N(v_p))} x_s^2,$$

or equivalently,

$$y_t + \sum_{s\in I(F_1)\setminus\{t\}} y_s + \sum_{s\in I(F_2)} y_s + \sum_{s\in I(N(v_p)\setminus F_1\cup F_2)} y_s = x_t^2 + \sum_{s\in I(F_1)\setminus\{t\}} x_s^2 + \sum_{s\in I(F_2)} x_s^2 + \sum_{s\in I(N(v_p)\setminus F_1\cup F_2)} x_s^2.$$

From (4.11), construction of y and the equalities $x_s^1 = x_s^2$ on $I(N(v_p) \setminus F_1 \cup F_2)$, we get $x_t^1 = x_t^2$, completing the proof of part 4b. $\square$

4.B Monge solutions under extraction of graphs

The main theorem of this section establishes that, roughly speaking, the extraction from K_m of a subgraph with an inner hub provides a unique Monge solution, possibly under an additional regularity condition on one of the marginals.

We will present several examples of graphs obtained in this way later on, but for now we mention that the graph below (Figure 4.3) is obtained by extracting the edges $\{x_1, x_3\}$ and $\{x_2, x_4\}$ from the complete graph with four vertices K_4, which can be interpreted as maximal cliques of the graph with edges $\{x_1, x_3\}$ and $\{x_2, x_4\}$, and inner hub $A = \emptyset$.

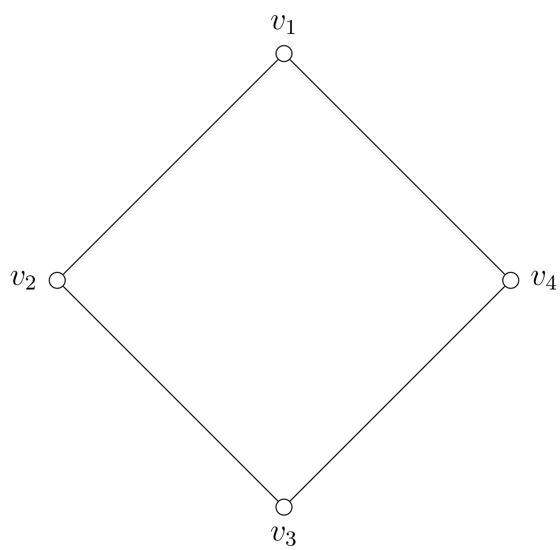


Figure 4.3: Cycle graph with $m = 4$.

As was shown in Chapter 3, this is the only cycle graph that provides unique Monge type solutions.

4.B.1 Monge solutions

We now state and prove our first main result.

Theorem 4.9. *Let $\{S_j\}_{j=1}^{l}$ be the collection of maximal cliques of a given subgraph S of C_m with inner hub A, for some $m \in \mathbb{N}$. Let $G := K_m \setminus S$ be connected, b the surplus function associated to G and μ_i be probability measures over X_i, $i = 1, \dots, m$, with μ_1 absolutely continuous with respect to $\mathcal{L}^n$. Assume that one of the following conditions is met:*

(i) $v_1 \in V(G) \setminus V(S)$,

(ii) There exists $p \in I(N_G(v_1))$ such that $A \subseteq N_G(v_p)$, with μ_p is absolutely continuous with respect to $\mathcal{L}^n$, and, if S is not complete, $v_1 \notin A$.

Then every solution to the Kantorovich problem (KP) is induced by a map.

Proof. Let μ be a solution to the Kantorovich problem with surplus b and $(u_1, \dots, u_m)$ a b-conjugate solution to its dual. Consider:

$$\widetilde{W} = \Big\{(x_1, \dots, x_m) : Du_1(x_1) \quad \text{exists,} \quad \text{and} \quad \sum_{i=1}^{m} u_i(x_i) = b(x_1, \dots, x_m)\Big\}.$$

The function u_1 is differentiable $\mathcal{L}^n$-a.e, as it is Lipschitz continuous. Hence, it is differentiable μ_1 a.e, as μ_1 is absolutely continuous. It follows that $\mu(\widetilde{W}) = 1$. Similarly, if in addition there exists $p \in \{2, \dots, m\}$ such that μ_p is absolutely continuous with respect to $\mathcal{L}^n$, we get u_p is differentiable μ_p a.e and $\mu(\widetilde{W_p}) = 1$, where

$$\widetilde{W_p} = \Big\{(x_1, \dots, x_m) : Du_1(x_1) \quad \text{and} \quad Du_p(x_p) \quad \text{exist,} \quad \text{and} \quad \sum_{i=1}^{m} u_i(x_i) = b(x_1, \dots, x_m)\Big\}.$$

Fix $x_1^0 \in spt(\mu_1)$, where $u_1(x_1)$ is differentiable, and $(x_2^0, \dots, x_m^0)$ such that $(x_1^0, \dots, x_m^0) \in \widetilde{W}$. Note that b is differentiable with respect to x_1 at $(x_1^0, \dots, x_m^0)$ and it satisfies

$$Du_1(x_1^0) = D_{x_1} b(x_1^0, \dots, x_m^0). \tag{4.12}$$

We will show that the map

$$(x_2,\dots,x_m)\mapsto D_{x_1}b(x_1^0,x_2\dots,x_m)$$

is injective on $\widetilde{W}_{x_1^0} := \left\{(x_2,\dots,x_m) : (x_1^0,x_2,\dots,x_m)\in\widetilde{W}\right\}$, if $v_1\in V(G)\setminus V(S)$, or on $\widetilde{W}_{x_{1p}^0} := \left\{(x_2,\dots,x_m) : (x_1^0,x_2,\dots,x_m)\in\widetilde{W}_p\right\}$, if there exists $p\in I\left(N(v_1)\right)$ such that μ_p is absolutely continuous with respect to $\mathcal{L}^n$ and $A\subseteq N(v_p)$; this will imply that the equation (4.12) defines $(x_2^0,\dots,x_m^0)$ uniquely from x_1^0, which will complete the proof. Let $(x_1^0,x_2^1,\dots,x_m^1),(x_1^0,x_2^2,\dots,x_m^2)\in\widetilde{W}$ and assume

$$D_{x_1}b(x_1^0,x_2^1,\dots,x_m^1)=\sum_{s\in I(N(v_1))}x_s^1=\sum_{s\in I(N(v_1))}x_s^2=D_{x_1}b(x_1^0,x_2^2,\dots,x_m^2). \tag{4.13}$$

We want to prove $x_s^1=x_s^2$ for every $s\in\{2,\dots,m\}$.

Set $x_1^0:=x_1^1:=x_1^2$ and $B_j=V(S_j)\setminus A$, with $j\in\{1,\dots,l\}$. First, note that $S=\bigcup_{j=1}^{l}S_j$ and

$$N(v_s)=V(G)\setminus\{v_s\}\quad\text{for any } s\in I(V(G)\setminus V(S)), \tag{4.14}$$

$$N(v_s)=V(G)\setminus V(S_j)\quad\text{for any } s\in I(B_j),\quad j\in\{1,\dots,l\}, \tag{4.15}$$

$$N(v_s)=V(G)\setminus V(S),\quad\text{for any } s\in I(A). \tag{4.16}$$

Let us consider two cases:

Case 1. Assume $v_1\in V(G)\setminus V(S)=\{v_1,\dots,v_m\}\setminus V(S)$, then by (4.14) we get $\overline{N}(v_1)=V(G)=\overline{N}(v_s)$ for any $s\in I(V(G)\setminus V(S))\setminus\{1\}$. It follows from part 4a of Lemma 4.8 that

$$x_s^1=x_s^2\quad\text{for all}\quad s\in I(V(G)\setminus V(S))\setminus\{1\}. \tag{4.17}$$

Fix $j\in\{1,\dots,l\}$, and let us consider two sub-cases:

(a) If $A = \emptyset$, then $V(S_j) = B_j$. By defining $F_1 = B_j$ and $F_2 = (V(G) \setminus B_j) \setminus \{v_1\}$, we get $F_1 \cup F_2 = V(G) \setminus \{v_1\} = N(v_1)$. Also, from (4.15) we have $N(v_s) = F_2 \cup F_3$ for all $s \in I(F_1)$, where $F_3 = \{v_1\}$. Then, we can apply part 4b of Lemma 4.8 to get $x_s^1 = x_s^2$ for every $s \in I(B_j)$; that is, $x_s^1 = x_s^2$ on $\bigcup_{j=1}^{l} I(B_j) = \bigcup_{j=1}^{l} I(V(S_j)) = I(V(S))$. Combining this result with (4.17) we get $x_s^1 = x_s^2$ on $I(V(G)) \setminus \{1\} = \{2, \ldots, m\}$. This completes the proof of sub-case (a).

(b) Assume $A \neq \emptyset$. By setting $V_2 = V(G) \setminus V(S)$ and $V_1 = A$ we can use (4.17) to get equality (4.3), and then, by (4.16) and part 1 of Lemma 4.8 we get $y := (y_1, y_2, \ldots, y_m) \in \widetilde{W}$, where

$$y_s = \begin{cases} x_s^1 & \text{if} \quad s \in \{1, \ldots, m\} \setminus I(A) \\ x_s^2 & \text{if} \quad s \in I(A). \end{cases}$$

Now, set $F_1 = B_j$, $F_2 = (V(G) \setminus V(S_j)) \setminus \{v_1\}$ and $F_3 = \{v_1\}$. Clearly, $F_1, F_2 \subseteq N(v_1) = V(G) \setminus \{v_1\}$ and $F_1 \cup F_2 = V(G) \setminus (A \cup \{v_1\})$, then $N(v_1) \setminus (F_1 \cup F_2) = A$. Furthermore, y and $(x_1^0, x_2^2, \ldots, x_m^2)$ trivially satisfies $y_s = x_s^2$ on $I(A)$, and by (4.15), $N(v_s) = F_2 \cup F_3$ for every $s \in I(F_1)$. Hence, by part 4b of Lemma 4.8 we get $x_s^1 = y_s = x_s^2$ on $I(B_j)$, which proves that, using (4.17) and the equality $\bigcup_{j=1}^{l} B_j = V(S) \setminus A$, $x_s^1 = x_s^2$ on $I(V(G) \setminus A) \setminus \{1\}$. We combine this result with (4.16) and part 4b of Lemma 4.8 to get $x_s^1 = x_s^2$ on $I(A)$; all the conditions needed to apply this part of the lemma are trivially satisfied by setting $F_1 = A$, $F_2 = (V(G) \setminus V(S)) \setminus \{v_1\}$, $F_3 = \{v_1\}$ and $p = 1$. We conclude that $x_s^1 = x_s^2$ on $I(V(G) \setminus \{v_1\}) = \{2, \ldots, m\}$, completing the proof of sub-case (b).

This completes the proof of case 1.

Case 2. Assume $v_1 \in V(S)$ and let $p \in I(N(v_1))$ be such that μ_p is absolutely continuous with respect to $\mathcal{L}^n$ and $A \subseteq N(v_p)$. Assume $Du_p(x_p^1)$ and $Du_p(x_p^2)$ exist. If S is complete, $S_j = S$ for all $j = 1, \ldots, l$, and so, $A = V(S)$. Using (4.13) and (4.16) we obtain

$$\sum_{s \in I(V(G) \setminus A)} x_s^1 = \sum_{s \in I(V(G) \setminus A)} x_s^2. \tag{4.18}$$

Also, using (4.16) and part 1 of Lemma 2.1 we get $z := (z_1, \ldots, z_m) \in \widetilde{W}_p$, where

$$z_s = \begin{cases} x_s^1 & \text{if} \quad s \in \{1, \ldots, m\} \setminus I(A) \\ x_s^2 & \text{if} \quad s \in I(A). \end{cases} \tag{4.19}$$

Fix $t \in I(V(G) \setminus A)$. Then

$$\begin{aligned}
\sum_{s \in I\left(\overline{N}(v_t)\right)} z_s &= z_t + \sum_{s \in I(N(v_t))} z_s \\
&= z_t + \sum_{s \in I(V(G) \setminus \{v_t\})} z_s && \text{by (4.14)} \\
&= \sum_{s \in I(V(G))} z_s \\
&= \sum_{s \in I(A)} z_s + \sum_{s \in I(V(G) \setminus A)} z_s \\
&= \sum_{s \in I(A)} x_s^2 + \sum_{s \in I(V(G) \setminus A)} x_s^1 && \text{by construction of } z \\
&= \sum_{s \in I(A)} x_s^2 + \sum_{s \in I(V(G) \setminus A)} x_s^2 && \text{by (4.18)} \\
&= \sum_{s \in I(V(G))} x_s^2 \\
&= x_t^2 + \sum_{s \in I(V(G) \setminus \{v_t\})} x_s^2 = x_t^2 + \sum_{s \in I(N(v_t))} x_s^2 = \sum_{s \in I\left(\overline{N}(v_t)\right)} x_s^2.
\end{aligned}$$

It follows that $z_s = x_s^2$ on $I(V(G) \setminus A)$, by part 3 of Lemma 4.8; that is,

$$x_s^1 = x_s^2 \ \text{ on } \ I(V(G) \setminus A), \tag{4.20}$$

by construction of z. Now, to prove that $x_s^1 = x_s^2$ on $I(A)$ we use part 4b of Lemma 4.8. For this, set $F_1 = A$, $F_2 = V(G) \setminus (A \cup \{v_p\})$ and $F_3 = \{v_p\}$. Since $v_p \in N(v_1) = V(G) \setminus V(S)$, hence $N(v_p) = V(G) \setminus \{v_p\}$ by (4.14), and so $F_1, F_2 \subseteq N(v_p)$. Note that $F_1 \cup F_2 = N(v_p)$ and by (4.16), $F_2 \cup F_3 = V(G) \setminus A = N(v_s)$ for every $s \in I(F_1)$. Also, from (4.20), $x_p^1 = x_p^2$. This allow us to apply part 4b of Lemma 4.8 to get $x_s^1 = x_s^2$ on $I(A)$. Hence, $x_s^1 = x_s^2$ on $I(V(G)) = \{1, \dots, m\}$.

Let us know assume that S is not complete, then $v_1 \notin A$ by assumption, which implies that $v_1 \in B_k$ for some $k \in \{1, \dots, l\}$. We first claim that $x_s^1 = x_s^2$ on $\bigcup_{\substack{j=1 \\ j \neq k}}^{l} I(B_j) = I(V(S) \setminus V(S_k))$. Indeed, from (4.15) and (4.13) we get

$$\sum_{s \in I(V(G) \setminus V(S_k))} x_s^1 = \sum_{s \in I(V(G) \setminus V(S_k))} x_s^2. \tag{4.21}$$

It follows that, by setting $V_1 = B_k$ and $V_2 = V(G) \setminus V(S_k)$, we can use (4.15) and part 1 of Lemma 4.8 to get $y := (y_1, \dots, y_m) \in \widetilde{W}_p$, where

$$y_s = \begin{cases} x_s^1 & \text{if} \quad s \in \{1, \dots, m\} \setminus I(B_k) \\ x_s^2 & \text{if} \quad s \in I(B_k). \end{cases} \tag{4.22}$$

Fix $j \in \{1, \dots, l\}$, with $j \neq k$. Set $F_1 = B_j$, $F_2 = V(G) \setminus (V(S_k) \cup B_j)$ and $F_3 = B_k$. Note that from (4.15) we get $F_1 \cup F_2 = V(G) \setminus V(S_k) = N(v_1)$ and $F_2 \cup F_3 = V(G) \setminus V(S_j) = N(v_s)$, for any $s \in I(F_1)$. Since y and $(x_1^0, x_2^2, \dots, x_m^2)$ satisfies $y_s = x_s^2$ on $I(F_3)$, we can apply part 4b of Lemma 4.8 to get $x_s^1 = y_s = x_s^2$ on $I(B_j)$; that is,

$$x_s^1 = x_s^2 \quad \text{on} \quad \bigcup_{\substack{j=1 \\ j \neq k}}^{l} I(B_j) = I(V(S) \setminus V(S_k)). \tag{4.23}$$

Next, as in Case 1, let us consider the following sub-cases:

(a) If $A = \emptyset$, then $V(S_k) = B_k$. Also, for any $t \in I(V(G) \setminus V(S))$ we get

$$\begin{aligned}
\sum_{s \in I(\overline{N}(v_t))} y_s &= y_t + \sum_{s \in I(N(v_t))} y_s \\
&= y_t + \sum_{s \in I(V(G) \setminus \{v_t\})} y_s && \text{by (4.14)} \\
&= \sum_{s \in I(V(G))} y_s \\
&= \sum_{s \in I(B_k)} y_s + \sum_{s \in I(V(G) \setminus B_k)} y_s \\
&= \sum_{s \in I(B_k)} x_s^2 + \sum_{s \in I(V(G) \setminus B_k)} x_s^1 && \text{by construction of } y \\
&= \sum_{s \in I(B_k)} x_s^2 + \sum_{s \in I(V(G) \setminus B_k)} x_s^2 && \text{by (4.21)} \\
&= \sum_{s \in I(V(G))} x_s^2 \\
&= x_t^2 + \sum_{s \in I(V(G) \setminus \{v_t\})} x_s^2 \\
&= x_t^2 + \sum_{s \in I(N(v_t))} x_s^2 \\
&= \sum_{s \in I(\overline{N}(v_t))} x_s^2
\end{aligned}$$

Thus, by part 3 of Lemma 4.8 we obtain $x_s^1 = x_s^2$ on $I(V(G) \setminus V(S))$, as $y_s = x_s^1$ on $I(V(G) \setminus V(S))$. Combining this with (4.23) we deduce

$$x_s^1 = x_s^2 \quad \text{on} \quad I(V(G) \setminus V(S_k)) = I(V(G) \setminus B_k) = I(N(v_1)). \tag{4.24}$$

To prove that $x_s^1 = x_s^2$ on $I(B_k)$ we use part 4b of Lemma 4.8. Let us first recall that $p \in I(N(v_1))$, and then, the above result tell us that $x_p^1 = x_p^2$. Furthermore, $p \in I(V(G) \setminus V(S))$ or $p \in I(B_j)$ for some $j \in \{1, \ldots, l\}$, $k \neq j$. Thus, from (4.14), (4.15) and the disjointness of B_k and B_j, we deduce $B_k \subseteq N(v_p)$. Now, set $F_1 = B_k$, $F_2 = N(v_p) \setminus B_k$ and $F_3 = V(G) \setminus N(v_p)$. Then, $F_1, F_2 \subseteq N(v_p)$, $F_1 \cup F_2 = N(v_p)$ and $F_2 \cup F_3 = V(G) \setminus B_k = N(v_s)$, for every $s \in I(F_1)$. Also, from (4.24) we get $x_s^1 = x_s^2$ on $I(F_3)$, as it is evident that $B_k \cap F_3 = \emptyset$. We can then apply part 4b of Lemma 4.8 to obtain $x_s^1 = x_s^2$ on $I\left(B_k\right)$, which combined with (4.24) allow us to have $x_s^1 = x_s^2$ on $\{2, \ldots, m\}$. This completes the proof of sub-case (a).

(b) Assume $A \neq \emptyset$. Let us first prove that $x_s^1 = x_s^2$ on $I(V(G) \setminus V(S))$; this will be achieved via part 3 of Lemma 4.8.

Using (4.21) and (4.23), we can write

$$\sum_{s \in I(V(G) \setminus V(S))} x_s^1 = \sum_{s \in I(V(G) \setminus V(S))} x_s^2,$$

and defining y as in (4.22) we can equivalently write

$$\sum_{s \in I(V(G) \setminus V(S))} y_s = \sum_{s \in I(V(G) \setminus V(S))} x_s^2.$$

Hence, from (4.16) and part 1 of Lemma 4.8 we get $y' = (y_1', \ldots, y_m') \in \widetilde{W}_p$, where

$$y_s' = \begin{cases} y_s & \text{if} \quad s \in I(V(G) \setminus A) \\ x_s^2 & \text{if} \quad s \in I(A) \end{cases}$$

$$= \begin{cases} x_s^1 & \text{if} \quad s \in I(V(G) \setminus V(S_k)) \\ x_s^2 & \text{if} \quad s \in I(V(S_k)). \end{cases}$$

Then,

$$\sum_{s\in I(V(G)\setminus V(S_k))} y'_s = \sum_{s\in I(N(v_1))} y'_s = Du_1(x_1^0) = \sum_{s\in I(N(v_1))} x_s^2 = \sum_{s\in I(V(G)\setminus V(S_k))} x_s^2.$$

By construction of y', one has,

$$\sum_{s\in I(V(G))} y'_s = \sum_{s\in I(V(G))} x_s^2.$$

Hence, using (4.14) we clearly might express it as

$$\sum_{s\in I\left(\overline{N}(v_t)\right)} y'_s = \sum_{s\in I\left(\overline{N}(v_t)\right)} x_s^2,$$

for every fixed $t \in I(V(G) \setminus V(S))$. We can now apply part 3 of Lemma 4.8 and get $y'_t = x_t^2$ on $I(V(G) \setminus V(S))$, which implies $x_t^1 = x_t^2$ on $I(V(G) \setminus V(S))$, since $I(V(G) \setminus V(S)) \subseteq I(V(G) \setminus V(S_k))$ and $y'_t = x_t^1$ on $I(V(G) \setminus V(S_k))$. Thus, from (4.23),

$$x_s^1 = x_s^2 \quad \text{on} \quad I(V(G) \setminus V(S_k)) = I(N(v_1)). \tag{4.25}$$

It only remains to prove that $x_s^1 = x_s^2$ on $I(V(S_k))$. Since $p \in I(N(v_1))$, from the above equalities $p \in I\left(V(G) \setminus V(S)\right)$ or $p \in I(B_j)$ for some $j \neq k$, and $x_p^1 = x_p^2$. Then $N(v_p) = V(G) \setminus \{v_p\}$ or $N(v_p) = V(G) \setminus V(S_j)$. It follows that $N(v_p) = V(G) \setminus \{v_p\}$, as $A \subseteq N(v_p)$ and $A \cap (V(G) \setminus V(S_j)) = \emptyset$. Now, set $F_1 = A$, $F_2 = V(G) \setminus (V(S) \cup \{v_p\})$ and $F_3 = \{v_p\}$. It is clear that $F_1, F_2 \subseteq N(v_p)$ and $F_2 \cup F_3 = V(G) \setminus V(S) = N(v_s)$ for every $s \in I(A)$. Furthermore, for y defined as in (4.22) and $(x_1^0, x_2^2, \ldots, x_m^2)$, we have $y_s = x_2^2$ on $I(B_k)$. It follows from (4.25) and construction of y that $y_s = x_2^2$ on $I(V(G) \setminus A)$; in particular, $y_s = x_2^2$ on $I\left(N(v_p) \setminus F_1 \cup F_2\right) \subseteq I(V(G) \setminus A)$. Hence,

$x^1_s = y_s = x^2_2$ on $I(A)$ by part 4b of Lemma 4.8, and then, we can easily obtain $x^1_s = x^2_s$ on $I(B_k)$, by applying again part 4b of Lemma 4.8; this time we set $F_1 = B_k$, $F_2 = (V(G) \setminus V(S_k)) \setminus \{v_p\}$ and $F_3 = \{v_p\}$. This completes the proof of sub-case (b).

This completes the proof of the theorem. □

Before presenting some examples, we note the following consequence of the preceding theorem.

Corollary 4.10. *Let G be a subgraph of K_m with $|G| = m$, and satisfying $|N(v)| \in \{m-1, m-2\}$ for every $v \in V(G)$. Assume that μ_1 is absolutely continuous with respect to $\mathcal{L}^n$, and that* either $|N(v_1)| = m-1$ or *that μ_i is absolutely continuous with respect to Lesbesgue measure for some $i \in I(N(v_1))$. Then every solution to the Kantorovich problem* (KPG) *is induced by a map.*

Proof. Note that if $G \neq K_m$, then $G = K_m \setminus \bigcup S^l_{j=1}$, for some disjoint collection of complete graphs $\{S_j\}^l_{j=1}$, where $|V(S_j)| = 2$ for every j (that is, every S_j consists on a single edge). Clearly, the graph $\bigcup^l_{j=1} S_j$ has inner hub $A = \emptyset$ and maximal cliques $S_1, \dots, S_l$. The result then follows from Theorem 4.9. □

Note that the Gangbo-Swiech surplus corresponds to a complete graph, or, equivalently, to the graph G satisfying $|N(v)| = m-1$ for each $v \in V(G)$; the Corollary is then a generalization to the case where each vertex can be missing at most one edge connecting it to the other vertices.

4.B.2 Examples

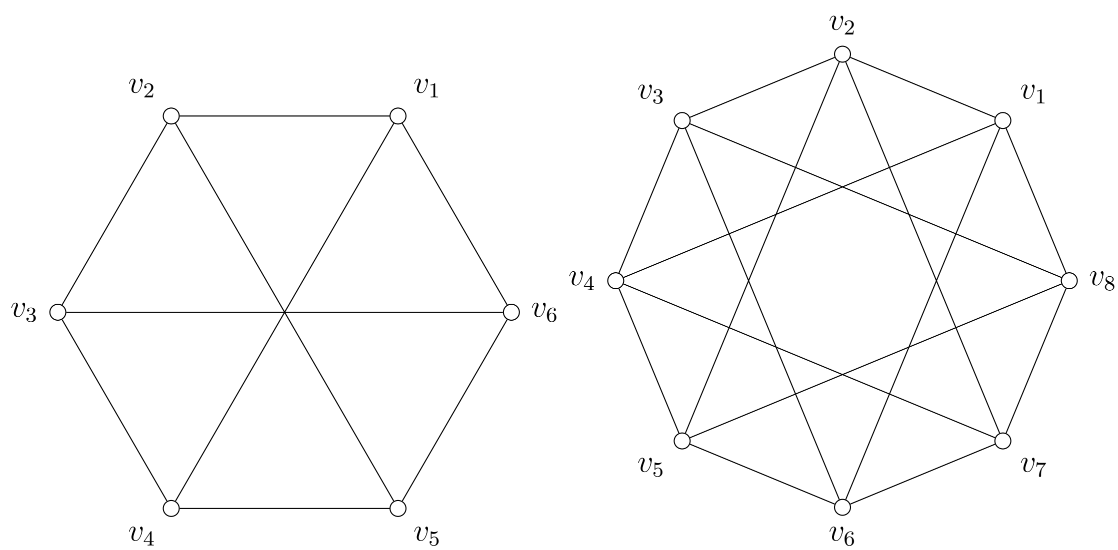
Here, we illustrate the result obtained in Theorem 4.9 through several examples.

(i) Let G be a complete k-partite graph with set partition $\{V_1, \dots, V_k\}$ and $m := |V(G)| = |\bigcup^k_{j=1} V_j|$. Write $\bigcup^k_{j=1} V_j = \{v_1, \dots, v_m\}$, and let $S_1, \dots, S_k$

be k complete graphs with sets of vertices $V_1, \ldots, V_k$, respectively. Note that $G := K_m \setminus \bigcup_{j=1}^{k} S_j$ and $N(v_1) = \bigcup_{\substack{j=1 \\ j\neq\alpha}}^{k} V_j$, for some $\alpha \in \{1, \ldots, k\}$. Hence, by assuming μ_1 and μ_p absolutely continuous, for some $p \in N(v_1)$, we can conclude by Theorem 4.9 that the graph G gives a unique Monge solution, as we can interpret $\{S_j\}_{j=1}^{k}$ as the collection of maximal cliques of the graph $\bigcup_{j=1}^{k} S_j$. Here, $A = \emptyset$ is clearly the inner hub of $\bigcup_{j=1}^{k} S_j$.

A special case is the complete graph C_k; several other examples of k-partite graphs are below.

- Complete bipartite graphs $K_{m,n}$:



(a) Graph $K_{3,3}$. Known as the *Utility graph*.

(b) Graph $K_{4,4}$. Known as the *Cayley graph*.

Figure 4.4: Graphs $K_{3,3}$ and $K_{4,4}$.

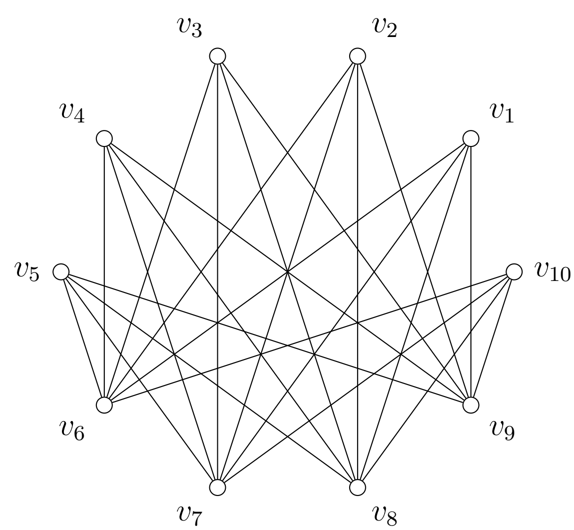
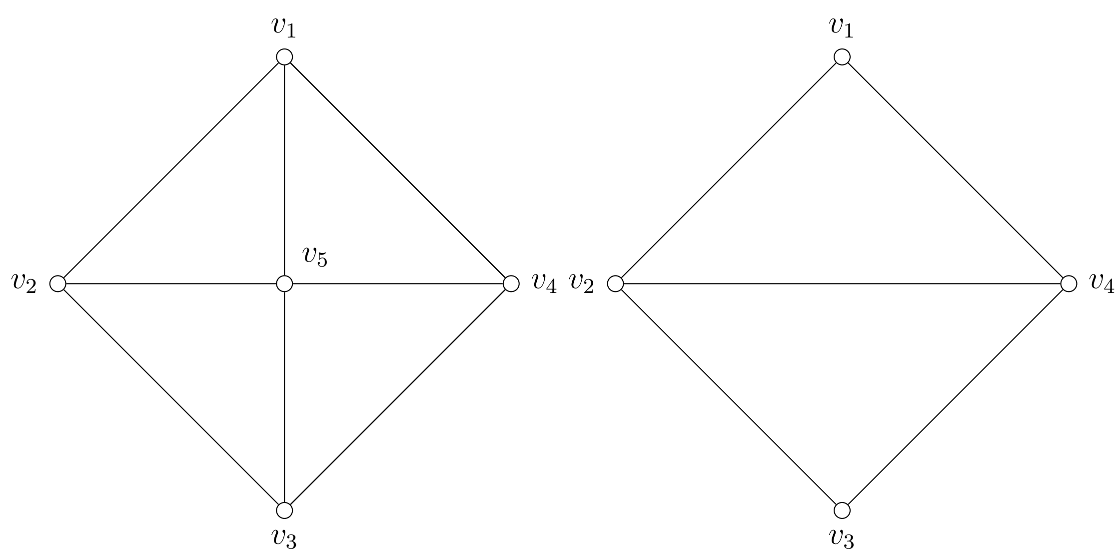


Figure 4.5: Bipartite graph with set partition $\{V_1, V_2\}$, where $V_1 = \{v_1, v_2, v_3, v_4, v_5, v_{10}\}$ and $V_2 = \{v_6, v_7, v_8, v_9\}$.

- <u>Complete Tripartite graphs $K_{m,n,p}$:</u>



(a) Graph $K_{1,2,2}$. Known as the *5-wheel graph*.

(b) Graph $K_{1,1,2}$. Known as the *Diamond graph*.

Figure 4.6: Graphs $K_{1,2,2}$ and $K_{1,1,2}$.

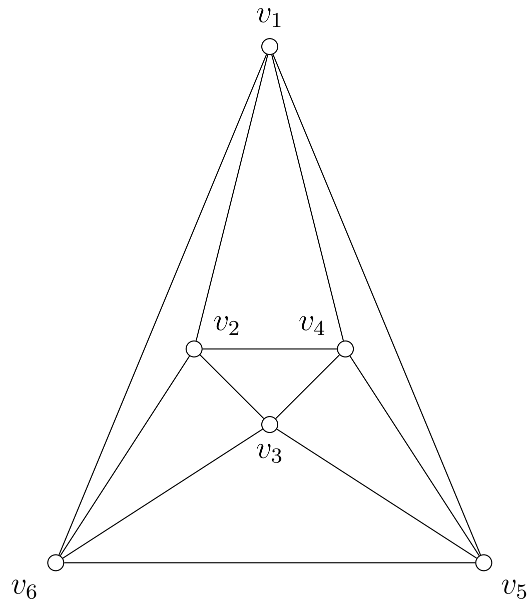


Figure 4.7: Graph $K_{2,2,2}$. Known as the *Octahedral graph*.

- A notable special case of Corollary 4.10 occurs when m is even and $|N(v)| = m - 2$ for all $v \in V(G)$, in which case $G = K_{2,\ldots,2(\frac{m}{2}\text{times})}$. This graph is known as the *Cocktail Party Graph*. See example below.

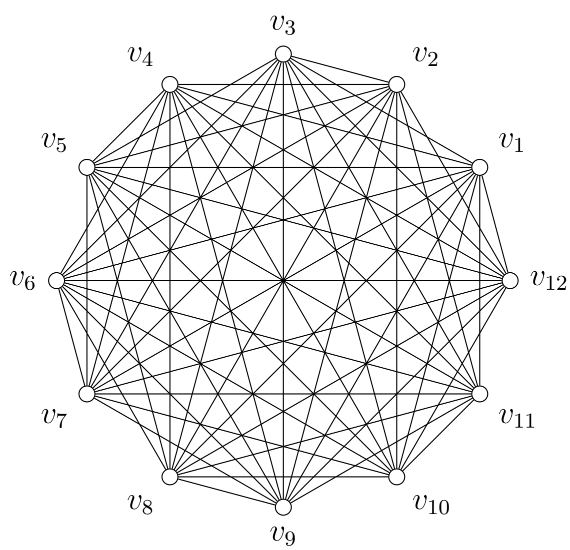


Figure 4.8: A Cocktail Party Graph with $m = 12$.

(ii) Theorem 4.9 can be used easily to construct many other, more obscure, graphs leading to Monge solutions. We construct one of such examples here; set

$$V_1 = \{v_1, v_{14}, v_{15}, v_{16}, v_{17}, v_{18}, v_{19}, v_{20}\},$$
$$V_2 = \{v_9, v_{10}, v_{11}, v_{12}, v_{13}, v_{14}, v_{15}, v_{16}\},$$
$$V_3 = \{v_6, v_{14}, v_{15}, v_{16}\},$$
$$V_4 = \{v_4, v_{14}, v_{15}, v_{16}\}.$$

Consider S_1, S_2, S_3, S_4 complete graphs with $V(S_j) = V_j$, $j = 1, 2, 3, 4$. Then, the graph $S = S_1 \cup S_2 \cup S_3 \cup S_4$ has inner hub $A = \{v_{14}, v_{15}, v_{16}\}$, with maximal cliques S_1, S_2, S_3, S_4. See Figure below.

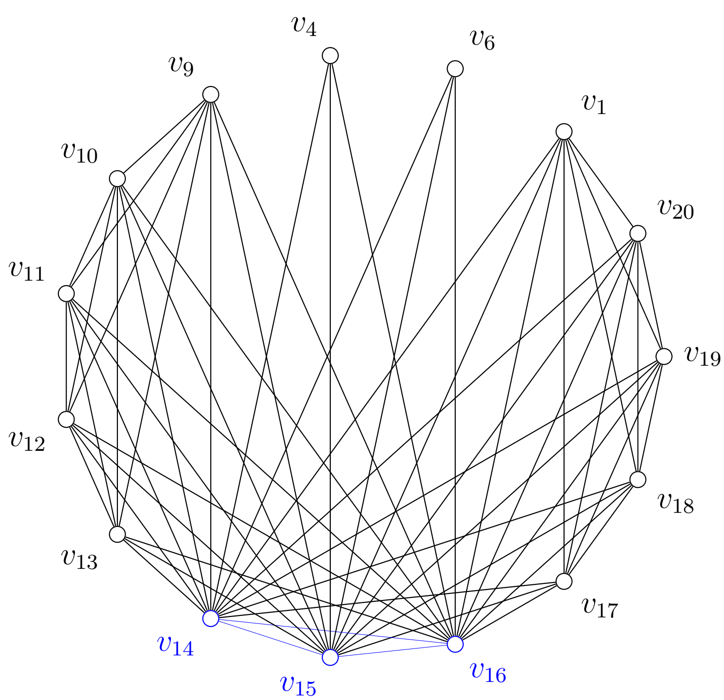


Figure 4.9: Graph $S = S_1 \cup S_2 \cup S_3 \cup S_4$.

Then, $G = C_{20} \setminus S$ provides a solution of Monge type. See graph below.

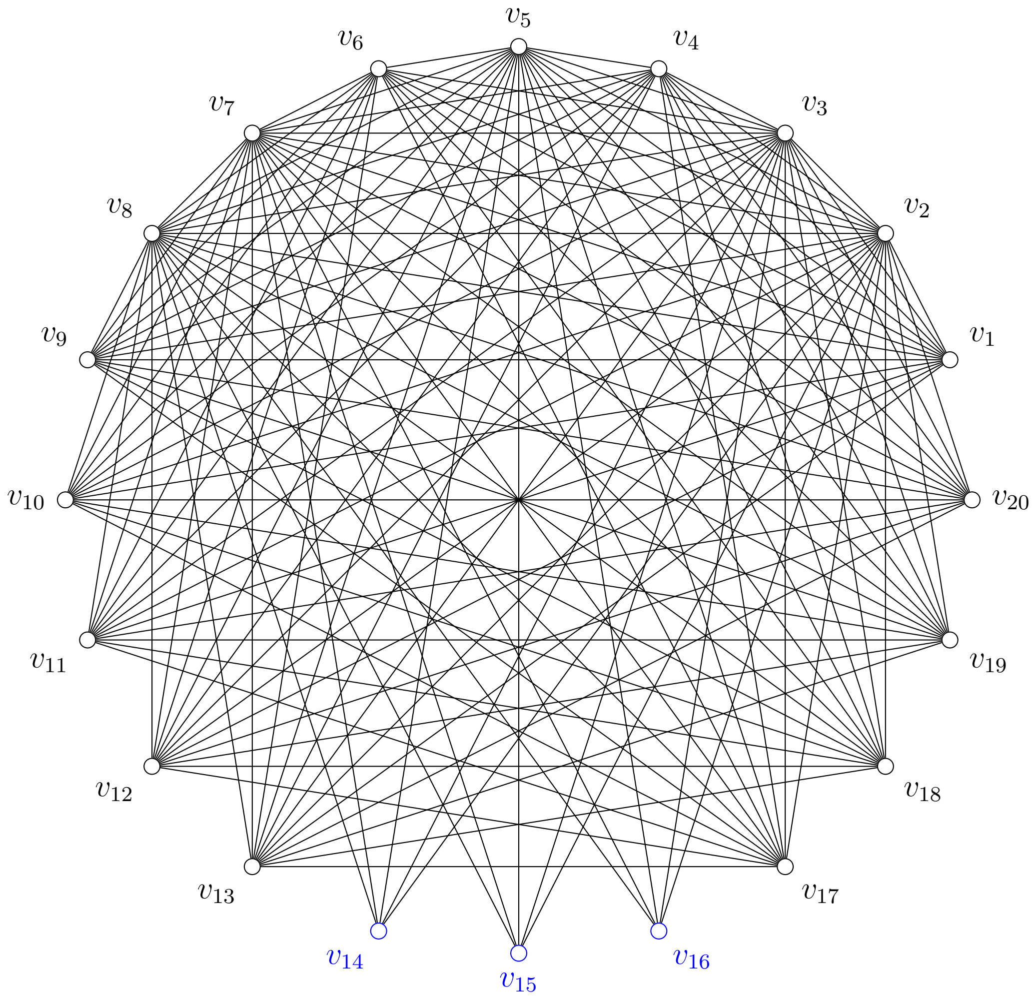


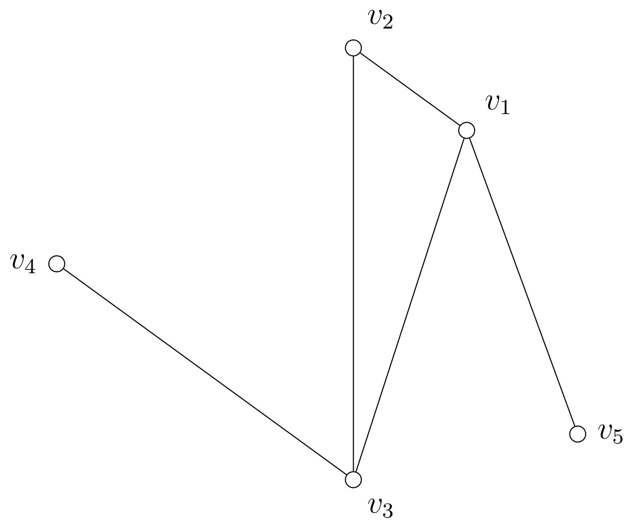
Figure 4.10: Graph $G = C_{20} \setminus S$.

4.C Monge solutions for graphs with inner hubs and gluing of them

The main result of this section (Theorem 4.12) ensures that under regularity conditions on two of the marginals, the surplus associated to a graph with inner hub provides a unique solution for the Monge-Kantorovich problem.

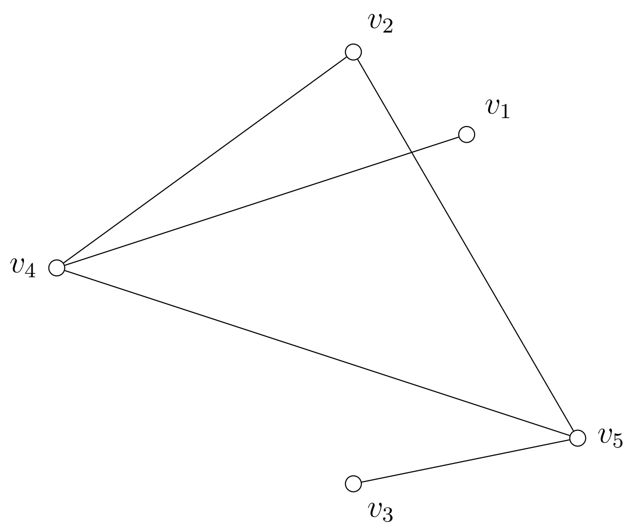
Before stating the main result of this section, we present the following simple example, which illustrates part of the motivation for Theorem 4.12 and Propositions 4.13 and 4.15.

Example 4.11. *Let b be the surplus associated to the graph G below.*



The second assertion of Proposition 4.6 implies that b is not twisted on splitting sets, and there are in fact choices μ_1, μ_2, μ_3 and μ_4 of marginals such that μ_1 is absolutely continuous with respect to $\mathcal{L}^n$ and the solution to (KPG) is of non-Monge form and non-unique (explicitly, take μ_3 to be a Dirac mass and the other marginals to be uniform on bounded domains). However, it is clear that the problem does admit a unique, Monge type solution as soon as both μ_1 and μ_3 are absolutely continuous. The reason for this is one may solve the three marginal problem with μ_1, μ_2, and μ_3 and the surplus $x_1 \cdot x_2 + x_1 \cdot x_3 + x_2 \cdot x_3$ via the Gangbo-Święch theorem [27], obtaining unique optimal maps T_2, T_3, and then solve independently the two marginal problems between μ_3 and μ_4 with surplus $x_3 \cdot x_4$, yielding a unique optimal map $\bar{T}_4$, and between μ_1 and μ_5, with surplus $x_1 \cdot x_5$, yielding a unique optimal map T_5. Since x_4 only interacts with x_3, and x_5 only interacts with x_1, $(T_2, T_3, T_4, T_5) := (T_2, T_3, \bar{T}_4 \circ T_3, T_5)$ is then the unique Monge solution for the overall problem.

This sort of result is not captured by Theorem 4.9, as the graph extracted from the complete graph C_5 to yield G, depicted below:



does not have an inner hub; we develop in this section a framework that encapsulates simple examples like this one, as well as more complicated ones which cannot be treated with adhoc arguments like the one sketched above.

4.C.1 Monge solutions for graphs with inner hubs

We now proceed to state and prove our second main result.

Theorem 4.12. *Let G be a graph with inner hub A and maximal cliques $S_1, \ldots, S_l$, with $m = |V(G)|$, and b its associated surplus. Let μ_i be probability measures over X_i, $i = 1, \ldots, m$, with μ_1 absolutely continuous with respect to $\mathcal{L}^n$. If there exists $p \in I(A)$ such that μ_p is absolutely continuous with respect to $\mathcal{L}^n$, then every solution to the Kantorovich problem* (KPG) *with surplus b is induced by a map.*

Proof. Let μ be a solution to the Kantorovich problem with surplus b and $(u_1, \ldots, u_m)$

a b-conjugate solution to its dual. Set:

$$\widetilde{W_p} = \Big\{(x_1, \ldots, x_m) : Du_1(x_1) \quad \text{and} \quad Du_p(x_p) \quad \text{exist,} \quad \text{and} \quad \sum_{i=1}^{m} u_i(x_i) = b(x_1, \ldots, x_m)\Big\}.$$

As in Theorem 4.9, we obtain $\mu(\widetilde{W_p}) = 1$. Moreover, by fixing x_1^0 where $u_1(x_1)$ is differentiable, we get for any $(m-1)$-tuple $(x_2^0, \ldots, x_m^0)$ satisfying $(x_1^0, \ldots, x_m^0) \in \widetilde{W_p}$,

$$Du_1(x_1^0) = D_{x_1} b(x_1^0, \ldots, x_m^0).$$

Let us show that the map

$$(x_2, \ldots, x_m) \mapsto D_{x_1} b(x_1^0, x_2, \ldots, x_m)$$

is injective on $\widetilde{W}_{x_{1p}^0} := \Big\{(x_2, \ldots, x_m) : (x_1^0, x_2, \ldots, x_m) \in \widetilde{W_p}\Big\}$. Indeed, assume

$$D_{x_1} b(x_1^0, x_2^1, \ldots, x_m^1) = \sum_{s \in I(N(v_1))} x_s^1 = \sum_{s \in I(N(v_1))} x_s^2 = D_{x_1} b(x_1^0, x_2^2, \ldots, x_m^2), \tag{4.26}$$

where $(x_1^0, x_2^1, \ldots, x_m^1), (x_1^0, x_2^2, \ldots, x_m^2) \in \widetilde{W_p}$, and $x_1^0 := x_1^1 := x_1^2$. Recall that $G = \bigcup_{j=1}^{l} S_j$, and without lost of generality assume $v_1 \in V(S_1)$. Then $v_1 \in B_1$ or $v_1 \in A$, where $B_j = V(S_j) \setminus A$, $j \in \{1, \ldots, l\}$. For the case $v_1 \in B_1$, we split the proof into several steps.

Step 1. Since S_1 is complete, for every $s \in I(B_1)$, $N(v_s) = V(S_1) \setminus \{v_s\}$, which implies $\overline{N}(v_1) = \overline{N}(v_s)$. Then, by part 4a of Lemma 4.8 we get

$$x_s^1 = x_s^2 \quad \text{for all} \quad s \in I(B_1). \tag{4.27}$$

Hence, the equalities $N(v_1) = V(S_1) \setminus \{v_1\} = (B_1 \cup A) \setminus \{v_1\}$, and (4.26), show that

$$\sum_{s \in I(A)} x_s^1 = \sum_{s \in I(A)} x_s^2. \tag{4.28}$$

Step 2. From part 2 of Lemma 4.8, for every $t \in I(A)$ we get

$$\left(x_t^2 - x_t^1\right) \cdot \sum_{s \in I(N(v_t))} \left(x_s^1 - x_s^2\right) \leq 0, \tag{4.29}$$

and by the definition of A,

$$\begin{aligned} N(v_t) &= V(G) \setminus \{v_t\} \\ &= \left(\bigcup_{j=1}^{l} V(S_j)\right) \setminus \{v_t\} \\ &= (A \setminus \{v_t\}) \bigcup \left(\bigcup_{j=1}^{l} B_j\right). \end{aligned} \tag{4.30}$$

Thus, we can write (4.29) as

$$\left(x_t^2 - x_t^1\right) \cdot \sum_{s \in I(A) \setminus \{t\}} \left(x_s^1 - x_s^2\right) + \left(x_t^2 - x_t^1\right) \cdot \sum_{s \in \bigcup_{j=1}^{l} I(B_j)} \left(x_s^1 - x_s^2\right) \leq 0.$$

It follows from (4.28) that

$$\|x_t^2 - x_t^1\|^2 + \left(x_t^2 - x_t^1\right) \cdot \sum_{s \in \bigcup_{j=1}^{l} I(B_j)} \left(x_s^1 - x_s^2\right) \leq 0, \tag{4.31}$$

hence, one easily deduces

$$\left(x_t^2 - x_t^1\right) \cdot \sum_{s \in \bigcup_{j=1}^{l} I(B_j)} \left(x_s^1 - x_s^2\right) \leq 0. \tag{4.32}$$

Summing over $t \in I(A)$ we get

$$\sum_{t \in I(A)} \left(x_t^2 - x_t^1\right) \cdot \sum_{s \in \bigcup_{j=1}^{l} I(B_j)} \left(x_s^1 - x_s^2\right) \leq 0,$$

and by (4.28), we must have equality in (4.32) for every $t \in I(A)$. Therefore, from (4.31) we get

$$x_t^1 = x_t^2 \quad \text{for all} \quad t \in I(A). \tag{4.33}$$

In particular, $x_p^1 = x_p^2$ and so, x_p^2 belongs to

$$\text{Argmax}\Big\{x_p \mapsto \Big(\sum_{s\in I(N(v_p))} x_s^1\Big)\cdot x_p - u_p(x_p)\Big\} \bigcap \text{Argmax}\Big\{x_p \mapsto \Big(\sum_{s\in I(N(v_p))} x_s^2\Big)\cdot x_p - u_p(x_p)\Big\}.$$

It follows that

$$\sum_{s\in I(N(v_p))} x_s^1 = Du_p(x_p^2) = \sum_{s\in I(N(v_p))} x_s^2,$$

or equivalently, invoking (4.30),

$$\sum_{s\in I(A)\setminus\{p\}} x_s^1 + \sum_{s\in \bigcup_{j=1}^l I(B_j)} x_s^1 = \sum_{s\in I(A)\setminus\{p\}} x_s^2 + \sum_{s\in \bigcup_{j=1}^l I(B_j)} x_s^2.$$

It immediately implies by (4.33) that

$$\sum_{s\in \bigcup_{j=1}^l I(B_j)} x_s^1 = \sum_{s\in \bigcup_{j=1}^l I(B_j)} x_s^2. \tag{4.34}$$

Step 3. Fix $k \in \{2, \dots, l\}$. From definition 4.2, $\{B_j\}_{j=1}^l$ is a disjoint collection of sets and every $j \in \{1, \dots, l\}$ satisfies $N(v_s) = V(S_j) \setminus \{v_s\} = (B_j \cup A) \setminus \{v_s\}$, for every $s \in I(B_j)$. Since $(x_1^0, x_2^1, \dots, x_m^1) \in \widetilde{W}_p$, we get

$$\{x_s^1\}_{s\in \bigcup_{\substack{j=1\\ j\neq k}}^l I(B_j)} \in \text{Argmax}\Bigg\{ \{x_s\}_{s\in \bigcup_{\substack{j=1\\ j\neq k}}^l I(B_j)} \mapsto \Big(\sum_{s\in I(A)} x_s^1\Big)\cdot \sum_{s\in \bigcup_{\substack{j=1\\ j\neq k}}^l I(B_j)} x_s$$
$$+ \sum_{\substack{j=1\\ j\neq k}}^l \sum_{\substack{s,t\in I(B_j)\\ s<t}} x_s\cdot x_t - \sum_{s\in \bigcup_{\substack{j=1\\ j\neq k}}^l I(B_j)} u_s(x_s)\Bigg\},$$

and by (4.33),

$$\{x_s^1\}_{s\in\bigcup_{\substack{j=1\\j\neq k}}^l I(B_j)} \in \operatorname{Argmax}\Bigg\{ \{x_s\}_{s\in\bigcup_{\substack{j=1\\j\neq k}}^l I(B_j)} \mapsto \Big(\sum_{s\in I(A)} x_s^2\Big)\cdot \sum_{s\in\bigcup_{\substack{j=1\\j\neq k}}^l I(B_j)} x_s + \sum_{\substack{j=1\\j\neq k}}^{l} \sum_{\substack{s,t\in I(B_j)\\s<t}} x_s\cdot x_t - \sum_{s\in\bigcup_{\substack{j=1\\j\neq k}}^l I(B_j)} u_s(x_s)\Bigg\}.$$

Hence, setting $y := (y_1, y_2, \ldots, y_m)$ with

$$y_s = \begin{cases} x_s^2 & \text{if} \quad s \in \{1,2,\ldots,m\}\setminus \bigcup_{\substack{j=1\\j\neq k}}^l I(B_j) = I(V(S_k)) \\ x_s^1 & \text{if} \quad s\in \bigcup_{\substack{j=1\\j\neq k}}^l I(B_j), \end{cases}$$

we get $y \in \widetilde{W}_p$, as $(x_1^0, x_2^2, \ldots, x_m^2) \in \widetilde{W}_p$. Since (4.34) holds true for every $(x_1^0, x_2^1, \ldots, x_m^1)$, $(x_1^0, x_2^2, \ldots, x_m^2) \in \widetilde{W}_p$; in particular, it is true for $(x_1^0, x_2^1, \ldots, x_m^1)$ and y, which implies that

$$\sum_{s\in I(B_k)} x_s^1 = \sum_{s\in I(B_k)} x_s^2.$$

Using (4.33) we can write the above equality as

$$\sum_{s\in I(V(S_k))} x_s^1 = \sum_{s\in I(V(S_k))} x_s^2.$$

Hence, all the elements of $I(B_k)$ satisfy (4.5), as each $s \in I(B_k)$ satisfies $N(v_s) = V(S_k)\setminus\{v_s\}$. Then, by part 3 of Lemma 4.8, $x_s^1 = x_s^2$ for all $s \in I(B_k)$. We thus conclude by (4.27) and (4.33) that $x_s^1 = x_s^2$ for all $s \in \bigcup_{j=1}^l I(B_j) \cup I(A) = I(G) = \{1, 2, \ldots, m\}$. This completes the proof for the case $v_1 \in B_1$.

Finally, for the case $v_1 \in A$, note that every $s \in I(A)$ satisfies $N(v_s) = V(G)\setminus\{v_s\}$, hence for any $s \in I(A)$ we get

$$\overline{N}(v_s) = \{v_s\} \cup N(v_s) = V(G) = (V(G) \setminus \{v_1\}) \cup \{v_1\} = N(v_1) \cup \{v_1\} = \overline{N}(v_1).$$

Therefore, by part 4a of Lemma 4.8 we get (4.33), and then, (4.26) reduces to (4.34). The rest of the proof runs exactly as the proof in Step 3, but instead of fixing k in $\{2, \ldots, l\}$, we fix it in $\{1, \ldots, l\}$, completing the proof of the theorem. □

4.C.2 Monge solutions for graphs glued on cliques

We now turn to a natural extension of Theorem 4.12. The next proposition states, roughly speaking, that gluing together several graphs with inner hubs via the procedure formulated in Definition 4.4, leads to a solution of Monge type.

Proposition 4.13. *Let S_1 be a graph with inner hub A_1 and $\{S_{1j}\}_{j=1}^{l}$ its collection of maximal cliques. Let $E \subset \{2, \ldots, l\}$ such that for every $\alpha \in E$, S_α is a graph with inner hub $A_\alpha \neq \emptyset$, and with collection of maximal cliques $\{S_{\alpha j}\}_{j=1}^{k_\alpha}$. Assume $A_\alpha \cap A_1 = \emptyset$ for every $\alpha \in E$, and set $G = \bigcup_{\alpha \in E \cup \{1\}} S_\alpha$ and $m = |V(G)|$. Let μ_i be probability measures over X_i, $i = 1, \ldots, m$ and assume:*

1. *S_α and S_1 are glued on a clique for all $\alpha \in E$.*
2. *$V(S_\alpha) \bigcap V(S_\beta) = A_1$ for all $\alpha \neq \beta$, $\alpha, \beta \in E$.*
3. *For each $\alpha \in E \cup \{1\}$, there exists $p_\alpha \in I(A_\alpha)$ such that μ_{p_α} is absolutely continuous with respect to $\mathcal{L}^n$.*
4. *μ_1 is absolutely continuous with respect to $\mathcal{L}^n$ and $v_1 \in V(S_{11})$.*

Then every solution to the Kantorovich problem (KPG) with surplus associated to G is concentrated on a graph of a measurable map.

Proof. The strategy of the proof is similar to the strategy used in Theorem 4.12. Let μ be a solution to the Kantorovich problem with surplus $b(x_1, \ldots, x_m)$, where b is the surplus associated to G. Let $(u_1,\ldots,u_m)$ be a b-conjugate solution to its dual and set

$$\widetilde{W}=\Big\{(x_1,\ldots,x_m):Du_1(x_1)\quad\text{and}\quad Du_{p_\alpha}(x_{p_\alpha})\quad\text{exist}\quad\text{for all }\alpha\in E\cup\{1\},\\ \text{and}\quad\sum_{i=1}^m u_i(x_i)=b(x_1,\ldots,x_m)\Big\}.$$

Fix $x_1^0\in spt(\mu_1)$, where $u_1(x_1)$ is differentiable. Then $Du_1(x_1^0)=D_{x_1}b(x_1^0,\ldots,x_m^0)$, for every $(x_1^0,\ldots,x_m^0)\in\widetilde{W}$. We want to prove that the map $(x_2,\ldots,x_m)\mapsto D_{x_1}b(x_1^0,x_2,\ldots,x_m)$ is injective on

$$\widetilde{W}_{x_1^0}:=\Big\{(x_2,\ldots,x_m):(x_1^0,x_2,\ldots,x_m)\in\widetilde{W}\Big\}.$$

Assume

$$D_{x_1}b(x_1^0,x_2^1,\ldots,x_m^1)=\sum_{s\in I(N(v_1))}x_s^1=\sum_{s\in I(N(v_1))}x_s^2=D_{x_1}b(x_1^0,x_2^2,\ldots,x_m^2),\tag{4.35}$$

with $(x_1^0,x_2^1,\ldots,x_m^1),(x_1^0,x_2^2,\ldots,x_m^2)\in\widetilde{W}$ and $x_1^0:=x_1^1=x_1^2$. Note that if $E=\emptyset$, we get $G=S_1$, and then, by Theorem 4.12 we get a solution of Monge type. Assume $E\neq\emptyset$ and set $B_j=V(S_{1j})\setminus A_1$, where $j\in\{1,\ldots,l\}$. Since $A_j\cap A_1=\emptyset$ for every $j\in E$, $B_j\neq\emptyset$ for every $j\in E$ and

$$N(v_s)=V(S_1)\setminus\{v_s\}=\bigcup_{j=1}^{l}B_j\cup(A_1\setminus\{v_s\}),\quad\text{for every}\quad s\in I(A_1).\tag{4.36}$$

Furthermore, by assumption 1 we can assume without lost of generality that $S_{1\alpha}=S_{\alpha 1}$ for every $\alpha\in E$. As in Theorem 4.12, we consider two cases, $v_1\in B_1$ or $v_1\in A_1$. Let us divide the proof of case $v_1\in B_1$ into several steps:

Step 1. We proceed to make a straightforward adaptation of the arguments used in Step 3 of the proof of Theorem 4.12. First, note that $N(v_s)=V(S_{11})\setminus\{v_s\}$, for every $s \in I(B_1)$, then, using the differentiability of $u_{p_1}(x_{p_1})$ at $x^1_{p_1}$ and $x^2_{p_1}$, and the equalities (4.35) and (4.36), we can mirror steps 1 and 2 in the proof of Theorem 4.12 to get:

$$x^1_s = x^2_s \quad \text{for all} \quad s \in I(B_1), \tag{4.37}$$

$$x^1_s = x^2_s \quad \text{for all} \quad s \in I(A_1), \tag{4.38}$$

and

$$\sum_{s \in \bigcup_{j=1}^{l} I(B_j)} x^1_s = \sum_{s \in \bigcup_{j=1}^{l} I(B_j)} x^2_s. \tag{4.39}$$

Step 2. Fix $\alpha \in \{2, \ldots, l\}$ and set $S_\beta = S_{1\beta}$ for any $\beta \in \{2, \ldots, l\} \setminus E$. Define $\mathcal{I}_1 = \bigcup_{\substack{\beta=2 \\ \beta \neq \alpha}}^{l} I(V(S_\beta) \setminus A_1) \cup I(B_1) = \{t_1, \ldots, t_d\}$ and $\mathcal{I}_2 = I(V(S_\alpha) \setminus A_1) = \{r_1, \ldots, r_e\}$. We claim that $x^1_s = x^2_s$ for all $s \in \mathcal{I}_2$, this will complete the proof. Indeed, note that

$$\begin{aligned}
\{1, \ldots, m\} &= \bigcup_{\beta=2}^{l} I(V(S_\beta)) \cup I(B_1) \\
&= \left(\bigcup_{\substack{\beta=2 \\ \beta \neq \alpha}}^{l} I(V(S_\beta)) \cup I(B_1) \right) \cup I(V(S_\alpha)) \\
&= \left(\bigcup_{\substack{\beta=2 \\ \beta \neq \alpha}}^{l} I(V(S_\beta) \setminus A_1) \cup I(B_1) \right) \cup I(V(S_\alpha) \setminus A_1) \cup A_1 \\
&= \mathcal{I}_1 \cup \mathcal{I}_2 \cup A_1
\end{aligned} \tag{4.40}$$

Furthermore, the last union is disjoint by assumptions 1 and 2. Now, let $g_1(x_{t_1}, \ldots, x_{t_d})$ and $g_2(x_{r_1}, \ldots, x_{r_e})$ be the functions formed by all the terms of b that depend only on the variables with index in $\mathcal{I}_1$ and $\mathcal{I}_2$ respectively.

From Definition 2.3 and assumptions 1 and 2, it is not hard to deduces that

$$\bigcup_{s\in\mathcal{I}_k} N(v_s) = \{v_s\}_{s\in\mathcal{I}_k} \cup A_1, k = 1, 2.$$

Combining the above equalities, (4.40) and (4.36) we get

$$\begin{aligned} b(x_1,\ldots,x_m) &= g_1(x_{t_1},\ldots,x_{t_d}) + g_2(x_{r_1},\ldots,x_{r_e}) \\ &\quad + \Big(\sum_{s\in I(A_1)} x_s\Big)\cdot \sum_{s\in\bigcup_{j=1}^{l} I(B_j)} x_s + \sum_{\substack{s,t\in I(A_1)\\ s<t}} x_s\cdot x_t \\ &= g_1(x_{t_1},\ldots,x_{t_d}) + g_2(x_{r_1},\ldots,x_{r_e}) \end{aligned}$$

$$+ \Big(\sum_{s\in I(A_1)} x_s\Big)\cdot \sum_{s\in\bigcup_{\substack{j=1\\ j\neq\alpha}}^{l} I(B_j)} x_s + \Big(\sum_{s\in I(A_1)} x_s\Big)\cdot \sum_{s\in I(B_\alpha)} x_s + \sum_{\substack{s,t\in I(A_1)\\ s<t}} x_s\cdot x_t.$$

Note that

$$\bigcup_{\substack{j=1\\ j\neq\alpha}}^{l} I(B_j) \subset \mathcal{I}_1, \tag{4.41}$$

and the only terms of b that depend on the variables with index in $\mathcal{I}_1$ are $g_1(x_{t_1},\ldots,x_{t_d})$ and $\Big(\sum_{s\in I(A_1)} x_s\Big)\cdot\sum_{s\in\bigcup_{\substack{j=1\\ j\neq\alpha}}^{l} I(B_j)} x_s$. Hence,

$$\begin{aligned} \{x_s^1\}_{s\in\mathcal{I}_1} \in \text{Argmax}\Bigg\{ & \{x_s\}_{s\in\mathcal{I}_1} \mapsto \Big(\sum_{s\in I(A_1)} x_s^1\Big)\cdot \sum_{s\in\bigcup_{\substack{j=1\\ j\neq\alpha}}^{l} I(B_j)} x_s + g_1(x_{t_1},\ldots,x_{t_d}) \\ & - \sum_{s\in\mathcal{I}_1} u_s(x_s) + g_2(x_{r_1}^1,\ldots,x_{r_e}^1) + \Big(\sum_{s\in I(A_1)} x_s^1\Big)\cdot \sum_{s\in I(B_\alpha)} x_s^1 \\ & + \sum_{\substack{s,t\in I(A_1)\\ s<t}} x_s^1\cdot x_t^1 - \sum_{s\in\mathcal{I}_2\cup I(A_1)} u_s(x_s^1)\Bigg\} \end{aligned}$$

$$= \operatorname{Argmax}\left\{ \{x_s\}_{s\in\mathcal{I}_1} \mapsto \Big(\sum_{s\in I(A_1)} x_s^1 \Big) \cdot \sum_{s\in\bigcup_{\substack{j=1\\ j\neq\alpha}}^{l} I(B_j)} x_s + g_1(x_{t_1},\dots,x_{t_d}) - \sum_{s\in\mathcal{I}_1} u_s(x_s) \right\},$$

and by (4.38),

$$\{x_s^1\}_{s\in\mathcal{I}_1} \in \operatorname{Argmax}\left\{ \{x_s\}_{s\in\mathcal{I}_1} \mapsto \Big(\sum_{s\in I(A_1)} x_s^2 \Big) \cdot \sum_{s\in\bigcup_{\substack{j=1\\ j\neq\alpha}}^{l} I(B_j)} x_s + g_1(x_{t_1},\dots,x_{t_d}) - \sum_{s\in\mathcal{I}_1} u_s(x_s) \right\}.$$

Since $(x_1^0, x_2^2, \dots, x_m^2) \in \widetilde{W}$, we obtain $y := (y_1, y_2, \dots, y_m) \in \widetilde{W}$, where

$$y_s = \begin{cases} x_s^2 & \text{if} \quad s \in \{1,2,\dots,m\} \setminus \mathcal{I}_1 \\ x_s^1 & \text{if} \quad s \in \mathcal{I}_1 \end{cases}$$

Therefore, (4.39) holds true for y and $(x_1^0, x_2^1, \dots, x_m^1) \in \widetilde{W}$; that is,

$$\sum_{s\in\bigcup_{j=1}^{l} I(B_j)} y_s = \sum_{s\in\bigcup_{j=1}^{l} I(B_j)} x_s^1,$$

or equivalently,

$$\sum_{s\in I(B_\alpha)} y_s + \sum_{s\in\bigcup_{\substack{j=1\\ j\neq\alpha}}^{l} I(B_j)} y_s = \sum_{s\in I(B_\alpha)} x_s^1 + \sum_{s\in\bigcup_{\substack{j=1\\ j\neq\alpha}}^{l} I(B_j)} x_s^1.$$

By the above equality, (4.41) and construction of y we get

$$\sum_{s\in I(B_\alpha)} x_s^2 = \sum_{s\in I(B_\alpha)} x_s^1. \tag{4.42}$$

Step 3. Since $B_\alpha = V(S_{1\alpha}) \setminus A_1$, by (4.38) and the above equality we can write,

$$\sum_{s \in I(V(S_{1\alpha}))} x_s^1 = \sum_{s \in I(V(S_{1\alpha}))} x_s^2. \tag{4.43}$$

Now, if $\alpha \in \{2, \ldots, m\} \setminus E$, then $S_\alpha = S_{1\alpha}$ and $N(v_s) = V(S_{1\alpha}) \setminus \{v_s\}$ for any $s \in I(B_\alpha)$. Hence, from (4.43) we get (4.5) on $I(B_\alpha)$, implying $x_s^1 = x_s^2$ on $I(V(S_\alpha))$, by part 3 of Lemma 4.8 and (4.38).

On the other hand, if $\alpha \in E$, the equality $A_\alpha \cap A_1 = \emptyset$ implies $A_\alpha \subseteq V(S_{\alpha 1}) \setminus A_1 = V(S_{1\alpha}) \setminus A_1 = B_\alpha$. It follows by (4.43) that equality (4.5) holds for the elements of $I(B_\alpha \setminus A_\alpha)$, as every $s \in I(B_\alpha \setminus A_\alpha)$ satisfies $N(v_s) = V(S_{1\alpha}) \setminus \{v_s\}$. Then, by part 3 of Lemma 4.8,

$$x_s^1 = x_s^2 \quad \text{for all} \quad s \in I(B_\alpha \setminus A_\alpha), \tag{4.44}$$

and by (4.42),

$$\sum_{s \in I(A_\alpha)} x_s^1 = \sum_{s \in I(A_\alpha)} x_s^2.$$

Note that by the differentiability of $u_{p_\alpha}(x_{p_\alpha})$ at $x_{p_\alpha}^1$ and $x_{p_\alpha}^2$, we can apply to the graph $S_\alpha = \bigcup_{j=1}^{k_\alpha} S_{\alpha j}$, the same arguments discussed in Step 2 of the proof of Theorem 4.12, getting

$$x_s^1 = x_s^2 \quad \text{for all} \quad s \in I(A_\alpha), \quad \text{and} \quad \sum_{s \in \bigcup_{j=1}^{k_\alpha} I(B_{\alpha j})} x_s^1 = \sum_{s \in \bigcup_{j=1}^{k_\alpha} I(B_{\alpha j})} x_s^2, \tag{4.45}$$

where $B_{\alpha j} = V(S_{\alpha j}) \setminus A_\alpha$ for all $j \in \{1, \ldots, k_\alpha\}$. By the left-hand equality of the above results, (4.38) and (4.44), we get

$$x_s^1 = x_s^2 \ \text{ on } \ I(V(S_{1\alpha})) = I(V(S_{\alpha 1})). \tag{4.46}$$

Next, we fix $r \in \{2, \ldots, k_\alpha\}$ and proceed to apply the same strategy used in step 2: we set $\mathcal{I}_1' = \{1,\ldots,m\} \setminus I(V(S_{\alpha r})) = \{e_1,\ldots,e_q\}$ and $\mathcal{I}_2' = I(B_{\alpha r}) = \{d_1,\ldots,d_f\}$, and consider $g_1'(x_{e_1},\ldots,x_{e_q})$ and $g_2'(x_{d_1},\ldots,x_{d_f})$, the functions formed by all the terms of b that depend only on the vertices with index in $\mathcal{I}_1'$ and $\mathcal{I}_2'$ respectively. Noting that $\bigcup_{s\in\mathcal{I}_j'} N(v_s) = \{v_s\}_{s\in\mathcal{I}_j'} \cup A_\alpha$, $j = 1,2$, and using the left-hand equality in (4.45), we follow the arguments of Step 2 to get

$$\begin{aligned}
\{x_s^1\}_{s\in\mathcal{I}_1'} \in \operatorname{Argmax}\Bigg\{ &\{x_s\}_{s\in\mathcal{I}_1'} \mapsto \Big(\sum_{s\in I(A_\alpha)} x_s^1\Big)\cdot \sum_{s\in\bigcup_{\substack{j=1\\ j\neq r}}^{k_\alpha} I(B_{\alpha j})} x_s + g_1'(x_{e_1},\ldots,x_{e_q}) \\
&- \sum_{s\in\mathcal{I}_1'} u_s(x_s) + g_2'(x_{d_1}^1,\ldots,x_{d_f}^1) + \Big(\sum_{s\in I(A_\alpha)} x_s^1\Big)\cdot \sum_{s\in I(B_{\alpha r})} x_s^1 \\
&+ \sum_{\substack{s,t\in I(A_\alpha)\\ s<t}} x_s^1\cdot x_t^1 - \sum_{s\in\mathcal{I}_2'\cup I(A_\alpha)} u_s(x_s^1)\Bigg\} \\
= \operatorname{Argmax}\Bigg\{ &\{x_s\}_{s\in\mathcal{I}_1'} \mapsto \Big(\sum_{s\in I(A_\alpha)} x_s^1\Big)\cdot \sum_{s\in\bigcup_{\substack{j=1\\ j\neq r}}^{k_\alpha} I(B_{\alpha j})} x_s + g_1'(x_{e_1},\ldots,x_{e_q}) \\
&- \sum_{s\in\mathcal{I}_1'} u_s(x_s)\Bigg\} \\
= \operatorname{Argmax}\Bigg\{ &\{x_s\}_{s\in\mathcal{I}_1'} \mapsto \Big(\sum_{s\in I(A_\alpha)} x_s^2\Big)\cdot \sum_{s\in\bigcup_{\substack{j=1\\ j\neq r}}^{k_\alpha} I(B_{\alpha j})} x_s + g_1'(x_{e_1},\ldots,x_{e_q}) \\
&- \sum_{s\in\mathcal{I}_1'} u_s(x_s)\Bigg\},
\end{aligned}$$

and then, using the right-hand equality in (4.45), we get the equality (4.42) on $I(B_{\alpha r})$; that is,

$$\sum_{s\in I(B_{\alpha r})} x_s^2 = \sum_{s\in I(B_{\alpha r})} x_s^1. \tag{4.47}$$

Finally, we combine the above equality with the left-hand equality in (4.45) to get (4.5) for all $s \in I(B_{\alpha r})$, since $N(v_s) = V(S_{\alpha r}) \setminus \{v_s\}$ for every $s \in I(B_{\alpha r})$ and $B_{\alpha j} = V(S_{\alpha j}) \setminus A_\alpha$. Hence, by part 3 of Lemma 4.8, $x_s^1 = x_s^2$ for all $s \in I(B_{\alpha r})$. Thus, $x_s^1 = x_s^2$ for all $s \in \bigcup_{j=2}^{k_\alpha} I(B_{\alpha j}) = I(V(S_\alpha) \setminus V(S_{\alpha 1}))$. Hence, from (4.46) we conclude $x_s^1 = x_s^2$ on $I(V(S_\alpha))$, and so, $x_s^1 = x_s^2$ for all $s \in \bigcup_{\alpha=1}^{l} I(V(S_\alpha)) = \{1, \ldots, m\}$, completing the proof of the case $v_1 \in B_1$.

For the case $v_1 \in A_1$, every $s \in I(A_1)$ satisfies $N(v_s) = V(S_1) \setminus \{v_s\}$, then any $s \in I(A_1)$ satisfies $\overline{N}(v_s) = \overline{N}(v_1)$. Therefore, by part 4a of Lemma 4.8 we get (4.38), and (4.35) reduces to (4.39). For the rest of the proof we fix $\alpha \in \{1, \ldots, l\}$ and mimic the proof of the case $v_1 \in B_1$, completing the proof of the theorem. □

Remark 4.14. *The results developed in this section, for graphs with inner hubs glued on their cliques, are neither more or less general than Theorem 4.9, which applies to graphs obtained by extracting subgraphs with inner hubs from complete graphs. To see this, note that in Theorem 4.9, if $m = 4$ and $l = 2$, with $S_1 = \{x_1, x_3\}$ and $S_2 = \{x_2, x_4\}$, we get the surplus associated to the graph in Figure 4.3, which clearly cannot be obtained from the results of Section 4.C. On the other hand, we can find examples of surplus functions covered by the framework presented in Section 4.C, but not covered by Theorem 4.9. For instance, Figure 4.13a and 4.13b are graphs whose respective surplus are not covered by Theorem 4.9, as we need more than two absolutely continuous measures and clearly, these conditions are necessary.*

We next turn to a slight generalization of Proposition 4.13, where, roughly speaking, any two graphs (with inner hubs) can be glued together (unlike in the preceding proposition, where each S_α, $\alpha \in E$ was glued to S_1). The proof is a straightforward modification of the proof of Proposition 4.13 and is therefore omitted.

In order to facilitate the description of the next Proposition we will introduce a natural higher level notion of graph. For this, we interpret any collection of graphs with inner hubs $\{G_\alpha\}_{\alpha=1}^{l}$, as the vertices of a graph $\mathcal{G}$, whose edges are glueings on cliques between the G_α and G_β; that is,

$$V(\mathcal{G}) = \{G_\alpha\}_{\alpha=1}^{l}$$

and

$$E(\mathcal{G}) = \{\{G_\alpha, G_\beta\} : G_\alpha \text{ is glued on a clique to } G_\beta\}.$$

Proposition 4.15. *Let $\{G_\alpha\}_{\alpha=1}^{l}$ be a collection of graphs with inner hubs A_α, and $\mathcal{G}$ its associated higher order graph (described above). Let $m = |\bigcup_{\alpha=1}^{l} V(G_\alpha)|$ and μ_i be probability measures over X_i, $i = 1, \ldots, m$, where without loss of generality $v_1 \in V(G_1)$. Assume:*

1. *For each distinct $\alpha \neq \beta$, $A_\alpha \cap A_\beta = \emptyset$ and $V(G_\alpha) \cap V(G_\beta)$ is either:*
   - *empty,*
   - *the vertex set $V(S)$, where S is a maximal clique S of both G_α and G_β (in this case G_α and G_β are glued on a clique S), or*
   - *A_λ for some other G_λ (as when G_α and G_β are both glued to G_λ).*
2. *μ_1 is absolutely continuous with respect to $\mathcal{L}^n$, and, for each $\alpha \in \{1, \ldots, l\}$, there exists $p_\alpha \in I(A_\alpha)$ such that μ_{p_α} is absolutely continuous with respect to $\mathcal{L}^n$.*
3. *For at least one maximal clique S having v_1 as one of its vertices, G_1 is not glued to any other G_α on S.*
4. *$\mathcal{G}$ is a tree.*

Then every solution to the Kantorovich problem with surplus $\bigcup_{\alpha=1}^{l} G_\alpha$ is induced by a map.

Remark 4.16. *Using the terminology developed above, the assumptions in Proposition 4.13 are equivalent to the assumptions in Proposition 4.15, except that the* *hypothesis that $\mathcal{G}$ is a tree is replaced with the hypothesis that $\mathcal{G}$ is a star with internal node S_1. Therefore, Proposition 4.15 is a direct generalization of Proposition 4.13.*

4.C.3 Examples

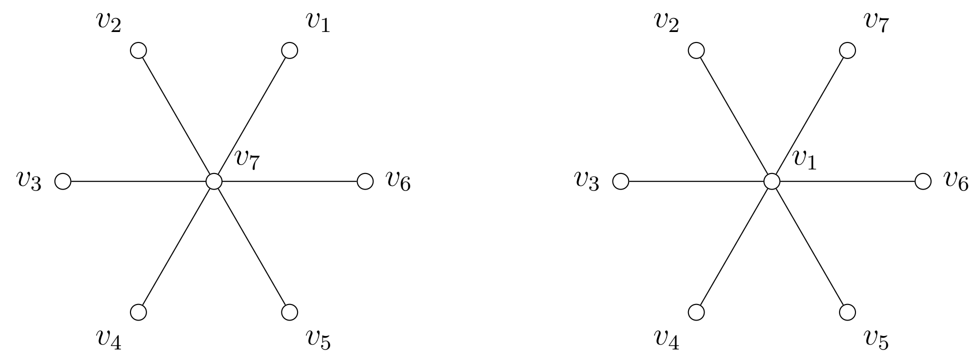
Let us illustrate the results obtained in this section throughout some examples. In what follows, μ_1 is absolutely continuous.

Examples 4.17. *(i) In Theorem 4.12, if $S_j = S_k$ for every $j, k \in \{1, \dots, l\}$, then $\bigcup_{j=1}^{l} S_j$ reduces to the Gangbo and Święch surplus.*

(ii) By Theorem 4.12, the graph $S_1 \cup S_2 \cup S_3$ in Example 4.3 provides a Monge solution, with μ_p absolutely continuous for some $p \in \{6, 7, 8\}$.

(iii) In Example 4.5, if there are $p_1 \in I(A)$ and $p_2 \in I(A')$ such that μ_{p_1} and μ_{p_2} are absolutely continuous, then by Proposition 4.13 the graph $\left(\bigcup_{j=1}^{3} S_j\right) \bigcup \left(\bigcup_{j=1}^{4} S'_j\right)$ provides a solution of Monge type .

(iv) By Theorem 4.12, any graph of the form $K_{1,k}$ (known as a star graph) provides a solution of Monge type, under at most two regularity conditions (see pictures below). Note that $|V(K_{1,k})| = k+1$ and there exists $v \in V(K_{1,k})$ such that $N(v) = \{v_1, \dots, v_k\}$. Additionally, $N(v_s) = \{v\}$ for all $s \in \{1, \dots, k\}$. This is one of the most simple graphs providing Monge solutions that we could obtain, since a graph with inner hub have in fact a "star shape". Note that, in the general setting, the single set $\{v\}$ is replaced by the inner hub A and $\{v_j\}$ is replaced by $B_j := V(S_j) \setminus A$, $j = 1, \dots, k$, where $\{S_j\}_{j=1}^{l}$ is the collection of maximal cliques. See for instance Figure 4.12.



(a) $K_{1,6}$, with $V_1 = \{v_7\}$ and $V_2 = \{v_i\}_{i=1}^{6}$. Here, we need regularity conditions on μ_1 and μ_7.

(b) $K_{1,6}$, with $V_1 = \{v_1\}$ and $V_2 = \{v_i\}_{i=2}^{7}$. Here, we only need a regularity condition on μ_1.

Figure 4.11: Star Graphs.

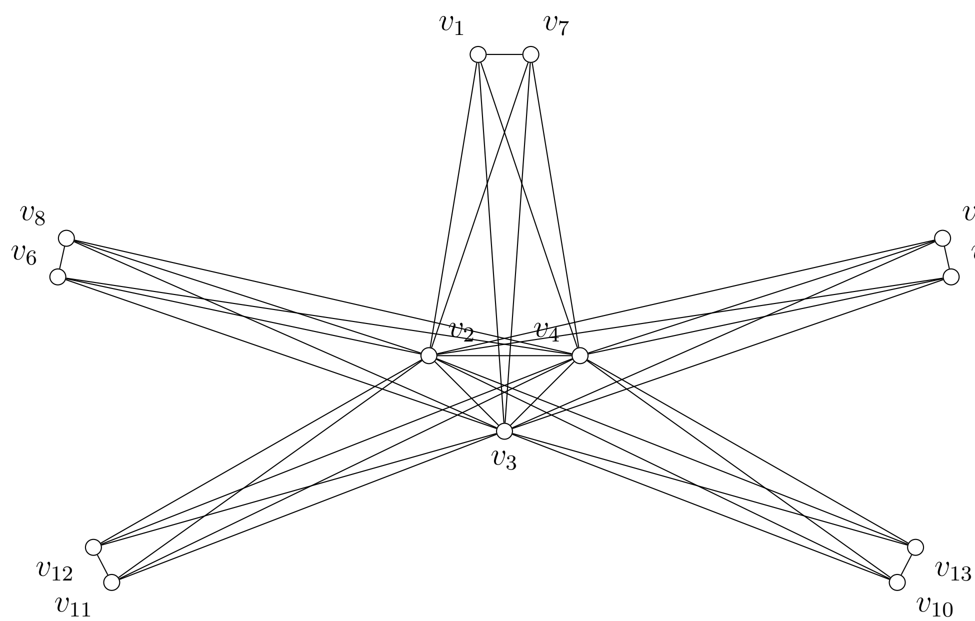
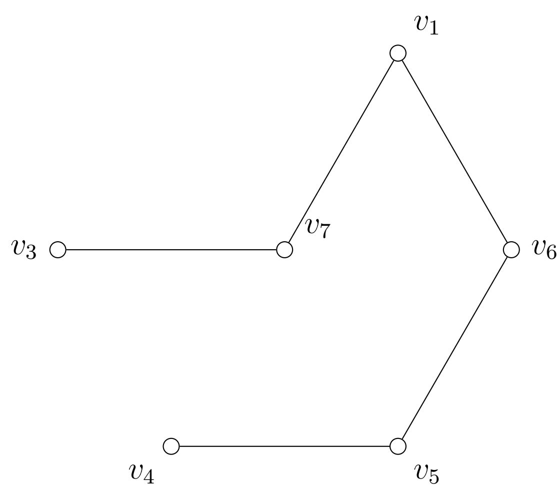


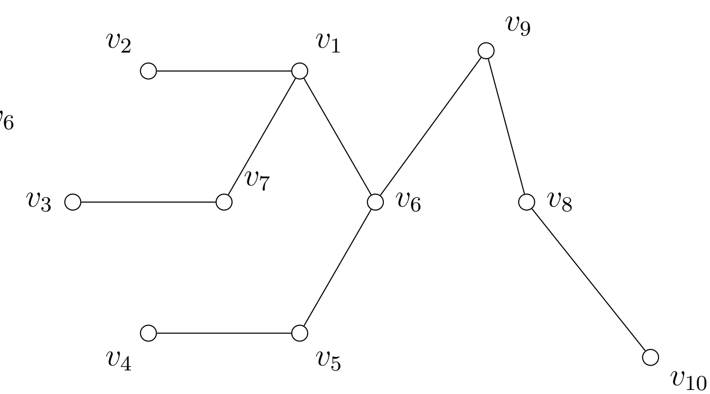
Figure 4.12: Graph $G = \bigcup_{j=1}^{5} S_j$ generated by the collection of its maximal cliques $\{S_j\}_{j=1}^{5}$, where $V(S_1) = \{v_2, v_3, v_4, v_1, v_7\}$, $V(S_2) = \{v_2, v_3, v_4, v_6, v_8\}$, $V(S_3) = \{v_2, v_3, v_4, v_{11}, v_{12}\}$, $V(S_4) = \{v_2, v_3, v_4, v_{10}, v_{13}\}$ and $V(S_5) = \{v_2, v_3, v_4, v_5, v_9\}$. Clearly, $A = \{v_2, v_3, v_4\}$ is the inner hub of G.

(v) *Let G be a graph tree with $V(G) = \{v_1, \ldots, v_m\}$ and $\mathcal{D} = \{s \in \{1, \ldots, m\} : |N(v_s)| = 1\}$. Assume μ_s is absolutely continuous for every $s \in \{2, \ldots, m\} \setminus$*

$\mathcal{D}$. Monge solutions for these graphs could be easily deduced by adapting the reasoning presented in Example 4.11; the solution will be the composition of optimal maps for two marginal problems along any path. Alternatively, these can be seen as special cases of Proposition 4.15.



(a) Path with vertex sequence $(x_3, x_7, x_1, x_6, x_5, x_4)$. Here, we need regularity conditions on $\mu_1, \mu_5, \mu_6, \mu_7$.



(b) Here, we need regularity conditions on μ_k, for every $k \in \{5, 6, 7, 8, 9\}$.

Figure 4.13: Trees.

(vi) Consider the graphs G_1 and G_2 below (Figures 4.14 and 4.15)

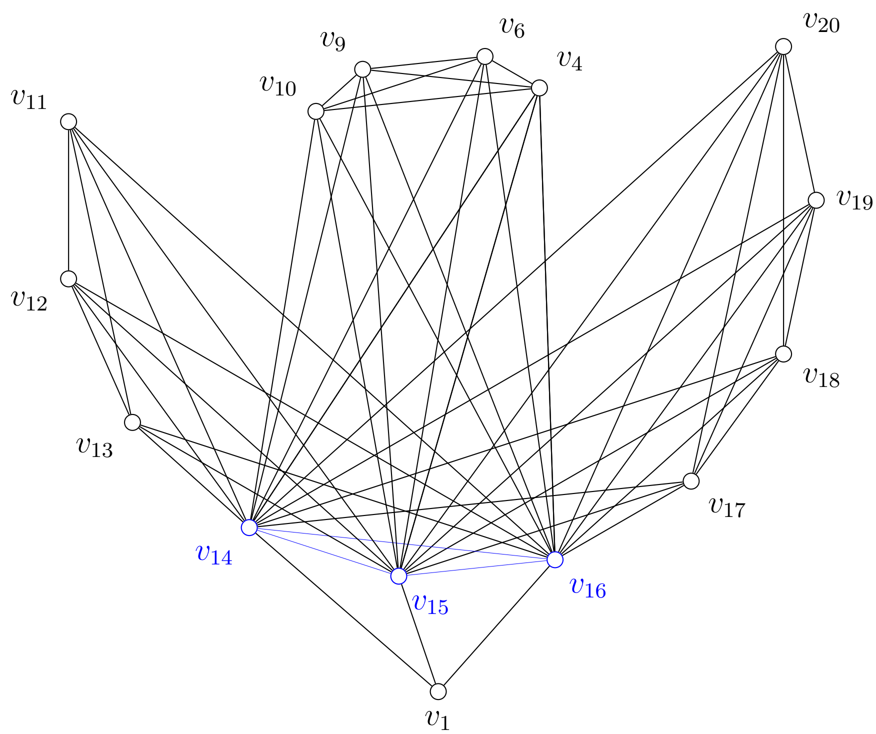


Figure 4.14: Graph G_1 with inner hub $\{v_{14}, v_{15}, v_{16}\}$.

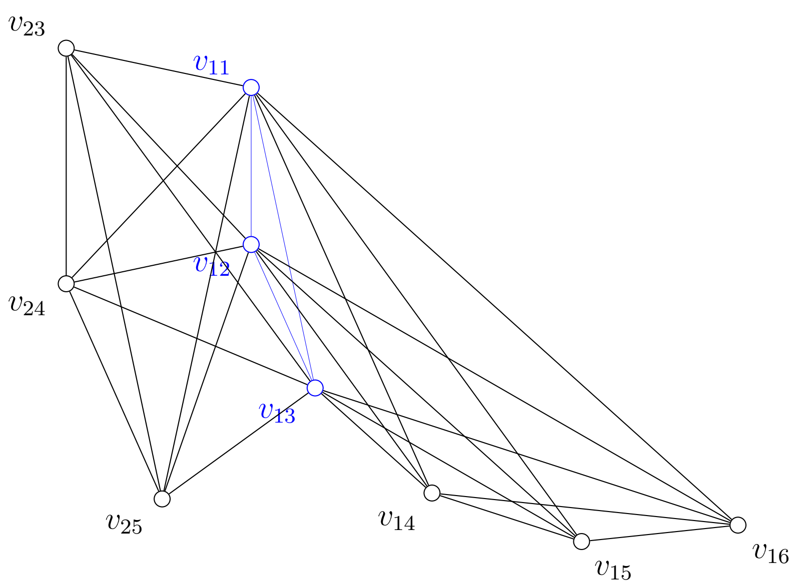


Figure 4.15: Graph G_2 with inner hub $\{v_{11}, v_{12}, v_{13}\}$.

Note that the graphs G_1 and G_2 have a common clique with vertices $\{v_{11}, v_{12}, v_{13}, v_{14}, v_{15}, v_{16}\}$ and they do not have any other common vertex; that is, G_1 and G_2 are glued on a clique, as shows the graph below.

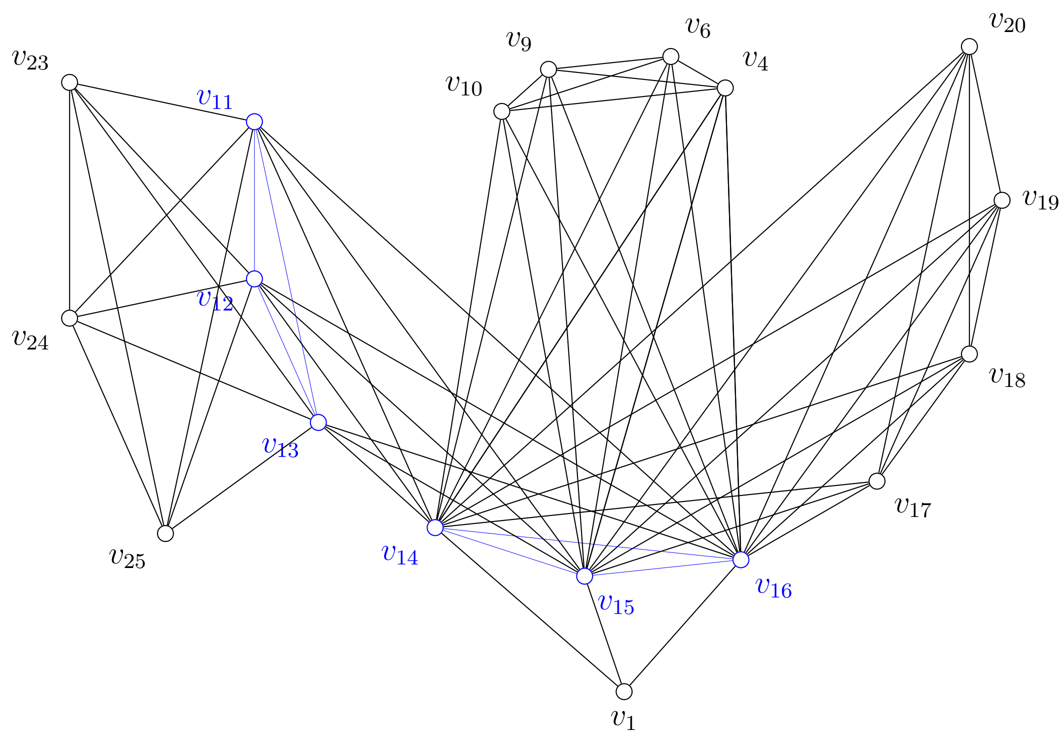


Figure 4.16: Graph $G_1 \cup G_2$.

Similarly, the graphs G_3 and G_4 in Figures 4.17 and 4.18 have inner hubs $\{v_4, v_6, v_9, v_{10}\}$ and $\{v_{21}, v_{22}\}$ respectively. Also, they have a common clique with vertices $\{v_4, v_6, v_9, v_{10}, v_{21}, v_{22}\}$ with no other common vertex. Then, they are glued on a clique. See Figure 4.19.

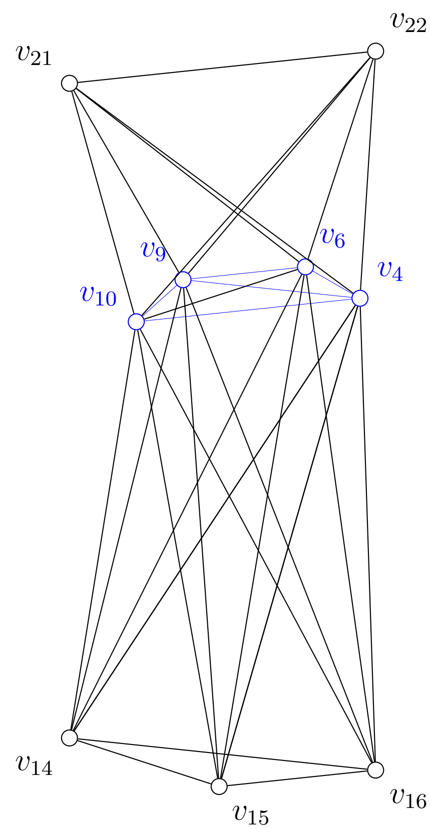


Figure 4.17: Graph G_3 with inner hub $\{v_4, v_6, v_9, v_{10}\}$.

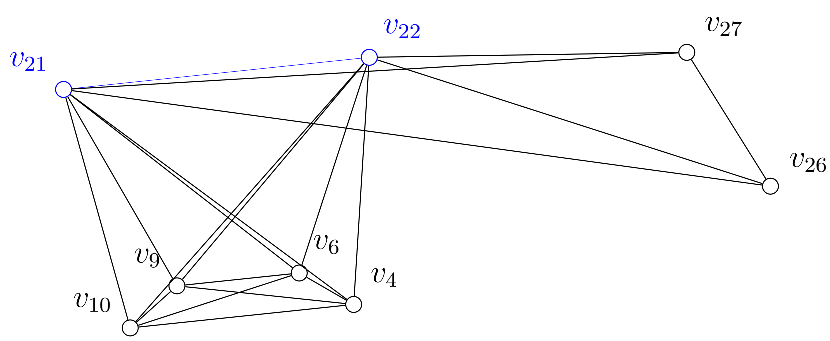


Figure 4.18: Graph G_4 with inner hub $\{v_{21}, v_{22}\}$.

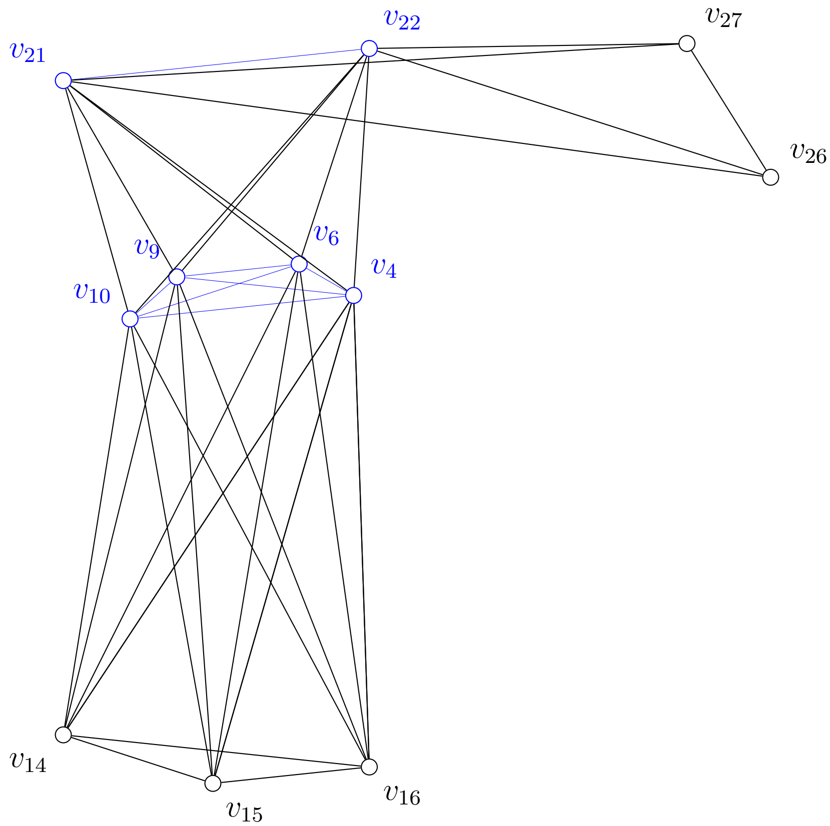


Figure 4.19: Graph $G_3 \cup G_4$.

It is clear that under some conditions the graphs in Figures 4.14, 4.15, 4.16, 4.17, 4.18 and 4.19 provide uniqueness in the Monge-Kantorovich problem, by Proposition 4.13. Note also that by Proposition 4.15, the graph $G_1 \cup G_2 \cup G_3 \cup G_4$ also provides uniqueness in the Monge-Kantorovich problem. See graph below.

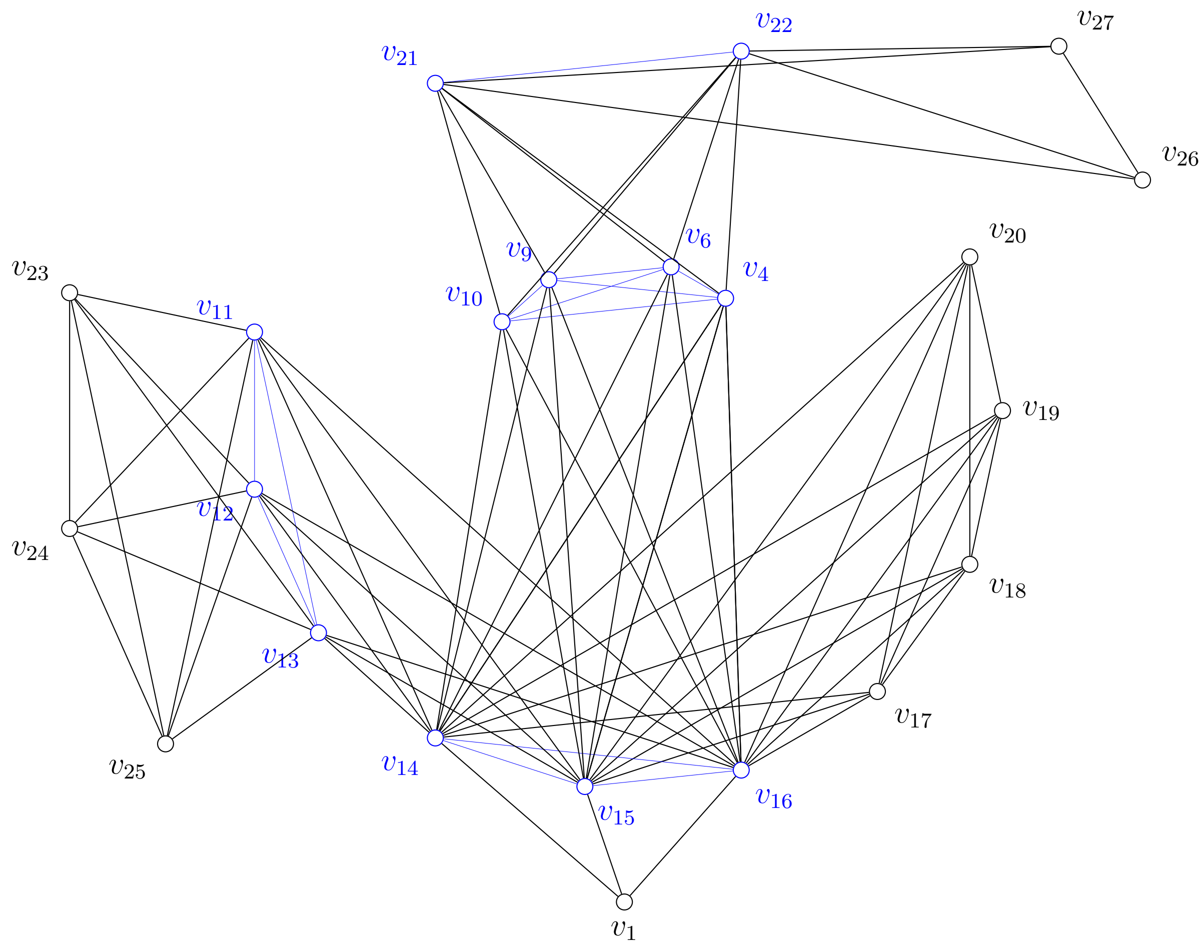


Figure 4.20: Graph $G_1 \cup G_2 \cup G_3 \cup G_4$.

4.D Uniqueness

Here, we include a standard argument, showing that in situations where all solutions are of Monge type, the solution to (KPG) must be unique.

Corollary 4.18. *Under the hypotheses in any of Theorem 4.9, Theorem 4.12, Proposition 4.13 or Proposition 4.15, the solution to the Kantorovich problem* (KPG) *is unique.*

Proof. If there are two such solutions, μ_0 and μ_1, linearity of the Kantorovich functional implies that their interpolant $\mu_{1/2} = \frac{1}{2}\mu_0 + \frac{1}{2}\mu_1$ is also a solution; under any of the collections of hypotheses listed in the statement of the corollary, the corresponding result then asserts that each of μ_0, μ_1 and $\mu_{1/2}$ must concentrate on the graph of a function. This is clearly not possible, as if μ_0, μ_1 concentrate on the graphs of T_0 and T_1, respectively, $\mu_{1/2}$ concentrates on the union of these two graphs, which is itself a single graph only if $T_0 = T_1$ μ_1 almost everywhere, in which case $\mu_0 = \mu_1$. □

4.E Discussion and negative examples

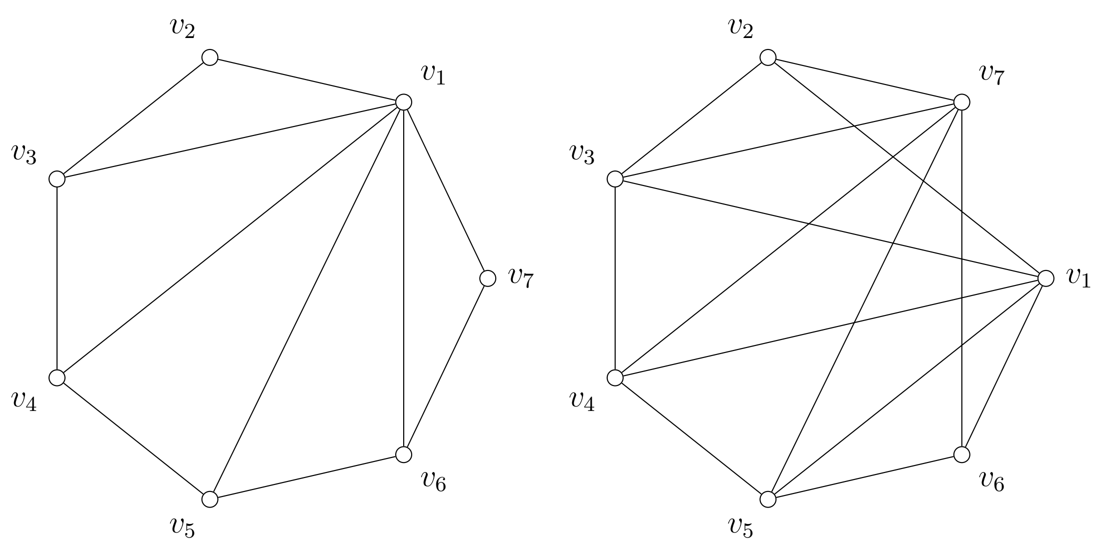
This section has identified a wide class of graphs leading to Monge solution and uniqueness results in the multi-marginal optimal transport problem (MP) with corresponding surplus (4.2), under appropriate conditions on the marginals; see Theorems 4.9 and 4.12 as well as Propositions 4.13 and 4.15. To the best of our knowledge, such results are not known for any graph which is not covered here. Furthermore, Part 2 of Proposition 4.6 verifies that the extra regularity conditions on the marginals imposed here are necessary in order to obtain Monge solution and uniqueness results.

There are many graphs to which none of Theorem 4.9, Theorem 4.12, Proposition 4.13 or 4.15 apply, and for most of these we do not know whether or not Monge solution and uniqueness results might hold, assuming for simplicity that all the marginals are absolutely continuous. A notable exception to this is the cycle graph for $m \geq 5$ (see Figure 1.2 for the case $m = 7$); in a recent work [48], we showed the existence of absolutely continuous marginals generating non-Monge solutions for the corresponding surplus (1.2). For illustrative purposes, we close by mentioning a class of graphs falling outside the scope of this paper, for which Monge solution and uniqueness remain completely open. For this, recall that for graphs G_1 and G_2 with disjoint vertex sets V_1 and V_2, the *graph join* $G_1 + G_2$ is defined as the graph union $G_1 \cup G_2$ together with all edges joining vertices in V_1 with vertices in V_2. Also, for any graph G, the *graph complement* (denoted $\overline{G}$) is the graph with vertices $V(G)$ and set of edges $E(\overline{G}) = \{\{v, w\} : v, w \in V(G) \text{ and } \{v, w\} \notin E(G)\}$.

Definition 4.19. *Let P_n be a path with n vertices and $\overline{K_r}$ the complement of the complete graph with r vertices K_r (so $N_{\overline{K_r}}(v) = \emptyset$ for every $v \in V(\overline{K_r})$). The fan graph $F_{r,n}$ is defined as the graph $\overline{K_r} + P_n$.*

Example 4.20. *Let us illustrate the above definition with some basic examples.*

- *The graphs $F_{1,1}$ and $F_{1,2}$ reduce to complete graphs with two and three vertices respectively.*
- *The graph $F_{1,3}$ reduces to the extraction of the graph consisting of only one edge from the complete graph K_4.*
- *The graphs $F_{1,6}$ and $F_{2,5}$, where for $F_{1,6}$ we denote the only vertex of K_1 as v_1, and for $F_{2,5}$ we denote the vertices of K_2 as v_1 and v_7. See figures below*



(a) $F_{1,6}$ (b) $F_{2,5}$

Figure 4.21: Fan Graphs.

Proposition 4.21. *Let $F_{k,n}$ be a fan graph.*

1. *If $n \geq 4$, then $F_{r,n}$ does not belong to the class of graphs in Theorem 4.9, Theorem 4.12, Proposition 4.13 or Proposition 4.15.*

2. *If $n < 4$, then $F_{r,n}$ belongs to the class of graphs considered in Theorem 4.9.*

The proof of Part 1 of the above proposition will be divided into two cases. In both cases the next lemma will be used during the proofs.

Lemma 4.22. *Assume $n \geq 4$. Then $F_{1,n}$ does not have an inner hub.*

Proof. Assume $F_{1,n}$ has an inner hub. Since $F_{1,n}$ is connected, every vertex in the nonempty hub is adjacent to all the other vertices. Now, the only vertex of $F_{1,n}$ satisfying this property is the vertex of $\overline{K_1} = K_1$ (so $V(K_1)$ is the hub of $F_{1,n}$). This implies by definition of inner hub that P_n is complete or it is the disjoint union of complete graphs. This is a contradiction as $n > 2$ (so P_n can not be complete) and it is connected, completing the proof of the lemma. □

Proof of Proposition 4.21. Since Proposition 4.15 generalizes Theorem 4.12 and Proposition 4.13, it suffices to prove Part 1 for Theorem 4.9 and Proposition 4.15. For this, we set $m = r + n$ and consider two cases.

Case 1. Assume $r = 1$. If $F_{1,n} = K_m \setminus S$ for some subgraph S of K_m, then $S = \overline{P_n}$ or $S = K_1 \cup \overline{P_n}$. Since $n \geq 4$, $\overline{P_n}$ is connected and there is not a vertex in $V(\overline{P_n})$ adjacent to all the other vertices; that is, $\overline{P_n}$ can not have an inner hub. Also, the only way that the disconnected graph $K_1 \cup \overline{P_n}$ has an inner hub is when $\overline{P_n}$ is complete (as it is connected), which is clearly not the case. Hence, the structure of $F_{1,n}$ does not correspond to the graphs considered in Theorem 4.9. On the other hand, note that the vertex of $\overline{K_1} = K_1$ is connected to all the other vertices of $F_{1,n}$, and the only case in Proposition 4.15 where a vertex of a graph $\bigcup_{\alpha=1}^{l} G_\alpha$ (where $\{G_\alpha\}_{\alpha=1}^{l}$ is a collection of graphs with inner hubs A_α satisfying the conditions in Proposition 4.15) satisfies this condition is when $l = 1$; that is, if there exists one of these collections satisfying $\bigcup_{\alpha=1}^{l} G_\alpha = F_{1,n}$, then $F_{1,n}$ would be a graph with an inner hub, contradicting Lemma 4.22. This proves that $F_{1,n}$ does not belong to the class of graphs in Proposition 4.15, completing the proof of Case 1.

Case 2. Assume $r \geqslant 2$. If $F_{r,n} = K_m \setminus S$ for some subgraph S of K_m, then $S = K_r \cup \overline{P_n}$. Note that S is disconnected with connected components K_r and $\overline{P_n}$ (as $n \geqslant 4$), so if S has inner hub then it must be empty, which implies $\overline{P_n}$ is complete. This clearly is not possible as $n > 1$. Hence, $F_{r,n}$ does not belong to the class of graphs in Theorem 4.9. For the other part of the assertion, consider $\{G_\alpha\}_{\alpha=1}^{l}$ a collection of graphs with inner hubs A_α satisfying the conditions imposed in Proposition 4.15 and assume $F_{r,n} = \bigcup_{\alpha=1}^{l} G_\alpha$. Fix any vertex v in $V(\overline{K_r}) \subseteq V(F_{r,n})$, then there exists β such that $v \in A_\beta$ or $v \in V(S_\beta) \setminus A_\beta$ for some maximal clique S_β of G_β. If $v \in A_\beta$, then

$$\begin{aligned} V(G_\beta) &= (V(G_\beta) \setminus \{v\}) \cup \{v\} \\ &= N_{\bigcup_{\alpha=1}^{l} G_\alpha}(v) \cup \{v\} \\ &= N_{F_{r,n}}(v) \cup \{v\} \\ &= V(P_n) \cup \{v\} \qquad\qquad \text{as } v \in V(\overline{K_r}) \end{aligned}$$

This implies that $G_\beta = P_n \cup K_{v,V(P_n)}$ where $K_{v,V(P_n)}$ is a bi-partite graph with set partition $\{\{v\}, V(P_n)\}$ (alternatively we can interpret it as a star graph with ”center” v); that is, G_β is a graph of the form $F_{1,n}$ having an inner hub. This is a contradiction by Lemma 4.22. This proves that $F_{r,n}$ does not satisfy the graph structure condition in Proposition 4.15. Now, assume $v \in V(S_\beta) \setminus A_\beta$ and without lost of generality assume $v \notin A_\alpha$ for any $\alpha \neq \beta$ (otherwise we apply the same arguments as in the case $v \in A_\beta$ above), then $V(S_\beta) = N_{F_{r,n}}(v) \cup \{v\} = V(P_n) \cup \{v\}$; that is, $P_n \cup K_{v,V(P_n)}$ is a complete graph (so it has inner hub $V(P_n \cup K_{v,V(P_n)})$), contradicting Lemma 4.22. Hence, $F_{r,n}$ does not belong to the class of graphs in Proposition 4.15, completing the proof of Part 1.

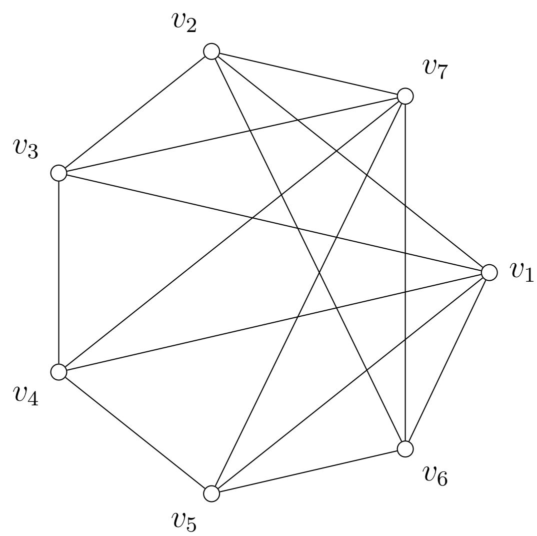
To prove Part 2, note that if $n \in \{1, 2, 3\}$ the graph $\overline{P_n}$ can be trivially expressed as a union of disjoint complete graphs, so $K_r \cup \overline{P_n}$ is a disjoint union of complete graphs and can be interpreted as a graph with empty inner hub. Since $F_{r,n} = K_m \backslash (K_r \cup \overline{P_n})$, we immediately conclude that $F_{r,n}$ belong to the class of graphs in Theorem 4.9. This completes the Proof of part 2. □

We note that the essential ideas in the proposition above can in fact be adapted to a more abstract class of graphs. The next lemma describes such a class, which therefore also falls outside the scope of the results in this paper and for which the Monge solution and uniqueness questions remain open.

Lemma 4.23. *Let G be a connected graph satisfying $N_G(v) \cup \{v\} \neq V(G)$, for all $v \in V(G)$ and consider the graph $\mathcal{F}_{r,G} := \overline{K_r} + G$. Then $\mathcal{F}_{r,G}$ does not belong to the classes of graph considered in Theorem 4.9 and Proposition 4.15.*

Proof. Note that the condition $N_G(v) \cup \{v\} \neq V(G)$, for all $v \in V(G)$ implies that G does not have an inner hub (there is not a vertex in $V(G)$ adjacent to all the other vertices). Also, since G is connected, $N_{\overline{G}}(v) \cup \{v\} \neq V(\overline{G})$ for all $v \in V(\overline{G}) = V(G)$, so $\overline{G}$ also has no inner hub. In particular, G and $\overline{G}$ are not complete and can not be expressed as a disjoint union of complete graphs. Knowing this, it is not hard to follow the arguments of Lemma 4.22 to prove that $\mathcal{F}_{1,G}$ does not have inner hub, and then, by mimicking the proof of the above proposition the proof is completed. □

Lemma 4.23 allows one to construct many graphs for which Monge solution and uniqueness results are not known, with more adhoc structure than the fan graphs considered above. One such possibility is illustrated in the figure below.


v_2
v_7
v_3
v_1
v_4
v_6
v_5

Chapter 5

A general condition for uniqueness in the Monge-Kantorovich problem

Our goal in this chapter is to generalize and extend the twist on c-splitting sets condition.

In the next section we formulate the conditions we will need and show a preliminary result. In Section 5.B we formulate and state our main result and in Section 5.C we illustrate our condition through several examples.

5.A Essential definitions and preliminary results

Here we establish the main background concepts used in this Chapter. For this, we first introduce some convenient notation. Assume $\{k_i\}_{i=1}^r \subseteq \{2, \dots, m\}$, with $k_1 < k_2 < \dots < k_r$.

- Let $S \subseteq \prod_{i=1}^m X_i$ be a c-splitting set and $(u_1, \dots, u_m)$ an $m-$tuple of c-splitting functions for S. Given $x_1^0 \in \pi_1(S)$, where π_1 is the canonical projection from $\prod_{i=1}^m X_i$ to X_1, we define

$$W_{x_1^0 k_1 \ldots k_r}(u_1, \ldots, u_m, S) := \Big\{(x_2, \ldots, x_m) \in \prod_{i=2}^m X_i : (x_1^0, x_2, \ldots, x_m) \in S \text{ and } Du_{k_i}(x_{k_i}) \text{ exists for each } i = 1, \ldots r\Big\}.$$

- For a given m-tuple of Borel functions $(u_1', \ldots, u_m')$ satisfying inequality (2.2), and $x_1^0 \in X_1$, we define

$$M_{x_1^0 k_1 \ldots k_r}(u_1', \ldots, u_m') := \Big\{(x_2, \ldots, x_m) \in \prod_{i=2}^m X_i : Du_{k_i}'(x_{k_i}) \text{ exists for each } i = 1, \ldots, r \text{ and } u_1'(x_1^0) + \sum_{i=2}^m u_i'(x_i) = c(x_1^0, x_2, \ldots, x_m)\Big\}.$$

From now on, if there is not danger of confusion, we will write $W_{x_1^0 k_1 \ldots k_r}$ and $M_{x_1^0 k_1 \ldots k_r}$ for $W_{x_1^0 k_1 \ldots k_r}(u_1, \ldots, u_m, S)$ and $M_{x_1^0 k_1 \ldots k_r}(u_1', \ldots, u_m')$ respectively.

Remark 5.1. *Note that $W_{x_1^0 k_1 \ldots k_r}(u_1, \ldots, u_m, S) \subseteq M_{x_1^0 k_1 \ldots k_r}(u_1, \ldots, u_m)$, for any c-splitting set S, m-tuple of c-splitting functions $(u_1, \ldots, u_m)$ for S and $x_1^0 \in \pi_1(S)$. Hence, for any fixed $(u_1, \ldots, u_m)$ satisfying inequality (2.2) and $x_1^0 \in X_1$, we get*

$$\bigcup_{S \in \mathcal{F}} W_{x_1^0 k_1 \ldots k_r}(u_1, \ldots, u_m, S) \subseteq M_{x_1^0 k_1 \ldots k_r}(u_1, \ldots, u_m),$$

where

$$\mathcal{F} := \Big\{S \subseteq \prod_{i=1}^m X_i : x_1^0 \in \pi_1(S) \text{ and } S \text{ is a splitting set having } (u_1, \ldots, u_m) \text{ as } c\text{-splitting functions}\Big\}.$$

On the other hand, for any $(x_2, \ldots, x_m) \in M_{x_1^0 k_1 \ldots k_r}(u_1, \ldots, u_m)$, the singleton $\bar{S} = \{x_1^0, x_2, \ldots, x_m\}$ is trivially a c-splitting set satisfying $(x_2, \ldots, x_m) \in$

$W_{x_1^0 k_1 \ldots k_r}(u_1, \ldots, u_m, \bar{S})$ *with* $\bar{S} \in \mathcal{F}$*. This immediately implies*

$$\bigcup_{S \in \mathcal{F}} W_{x_1^0 k_1 \ldots k_r}(u_1, \ldots, u_m, S) = M_{x_1^0 k_1 \ldots k_r}(u_1, \ldots, u_m).$$

Definition 5.2. *Let* c *be a continuous semi-concave cost function, and let* $\{k_i\}_{i=1}^r \subseteq \{2, \ldots, m\}$*, with* $k_1 < k_2 < \ldots < k_r$*. We say* c *is twisted on* c*-splitting sets with respect to the variables* $x_1, x_{k_1}, \ldots, x_{k_r}$*, if for each* c*-splitting set* $S \subseteq \prod_{i=1}^m X_i$ *and* m*-tuple* $(u_1, \ldots, u_m)$ *of c-splitting functions for* S*, the map*

$$(x_2, \ldots, x_m) \mapsto D_{x_1} c(x_1^0, x_2, \ldots, x_m)$$

is injective on the subset of $W_{x_1^0 k_1 \ldots k_r}$ *where* $D_{x_1} c(x_1^0, x_2, \ldots, x_m)$ *exists, for each fixed* $x_1^0 \in \pi_1(S)$ *satisfying* $W_{x_1^0 k_1 \ldots k_r} \neq \emptyset$*.*

Remark 5.3. *Note that Definition 2.5 is equivalent to* c *being twisted on* c*-splitting sets with respect to the variable* x_1*. Hence, our main result (Theorem 5.6), generalizes the main result in [35] (see Remark 2.6).*

We now proceed to prove a lemma, which provides an alternative way to check the condition above.

Lemma 5.4. *Let* c *be a continuous, semi-concave cost function. Let* $\{k_i\}_{i=1}^r \subseteq \{2, \ldots, m\}$*, with* $k_1 < k_2 < \ldots < k_r$*. The cost* c *is twisted on* c*-splitting sets with respect to the variables* $x_1, x_{k_1}, \ldots, x_{k_r}$ *if and only if for every* m*-tuple of Borel functions* $(u_1, \ldots, u_m)$ *satisfying inequality (2.2) and for every* $x_1^0 \in X_1$ *with* $M_{x_1^0 k_1 \ldots k_r} \neq \emptyset$*, we get that the map*

$$(x_2, \ldots, x_m) \mapsto D_{x_1} c(x_1^0, x_2, \ldots, x_m)$$

is injective on the subset of $M_{x_1^0 k_1 \ldots k_r}$ *where* $D_{x_1} c(x_1^0, x_2, \ldots, x_m)$ *exists.*

Proof. The converse is straightforward, as for every c-splitting set $S \subseteq \prod_{i=1}^m X_i$ and

m-tuple $(u_1, \ldots, u_m)$ of c-splitting functions for S, we have $W_{x_1^0 k_1 \ldots k_r} \subseteq M_{x_1^0 k_1 \ldots k_r}$ for each fixed $x_1^0 \in \pi_1(S)$. Hence, if $W_{x_1^0 k_1 \ldots k_r} \neq \emptyset$ we get $M_{x_1^0 k_1 \ldots k_r} \neq \emptyset$, which implies that the map $(x_2, \ldots, x_m) \mapsto D_{x_1} c(x_1^0, x_2, \ldots, x_m)$ is injective on the subset of $M_{x_1^0 k_1 \ldots k_r}$ where $D_{x_1} c(x_1^0, x_2, \ldots, x_m)$ exists, in particular, it is injective on the subset of $W_{x_1^0 k_1 \ldots k_r}$ where $D_{x_1} c(x_1^0, x_2, \ldots, x_m)$ exists; that is, c is twisted on c-splitting sets with respect to the variables $x_1, x_{k_1}, \ldots, x_{k_r}$. Assume now that the cost c is twisted on c-splitting sets with respect to the variables $x_1, x_{k_1}, \ldots, x_{k_r}$. Let $(u_1, \ldots, u_m)$ be an m-tuple of Borel functions satisfying inequality (2.2), and fix $x_1^0 \in X_1$. Assume $M_{x_1^0 k_1 \ldots k_r} \neq \emptyset$, and set

$$
\begin{aligned}
S &:= \Big\{ (x_1^0, x_2, \ldots, x_m) \in \prod_{i=1}^m X_i : (x_2, \ldots, x_m) \in M_{x_1^0 k_1 \ldots k_r} \Big\} \\
&= \Big\{ (x_1^0, x_2, \ldots, x_m) \in \prod_{i=1}^m X_i : u_1(x_1^0) + \sum_{i=2}^m u_i(x_i) = c(x_1^0, x_2, \ldots, x_m), \text{ and} \\
&\qquad\qquad Du_{k_i}(x_{k_i}) \text{ exists for each } i = 1, \ldots r \Big\}.
\end{aligned}
$$

Clearly, S is a c-splitting set, $\pi_1(S) = \{x_1^0\}$ and $W_{x_1^0 k_1 \ldots k_r} = M_{x_1^0 k_1 \ldots k_r} \neq \emptyset$. This immediately implies, by assumption that the map $(x_2, \ldots, x_m) \mapsto D_{x_1} c(x_1^0, x_2 \ldots, x_m)$ is injective on the subset of $M_{x_1^0 k_1 \ldots k_r}$ where $D_{x_1} c(x_1^0, x_2, \ldots, x_m)$ exists, completing the proof of the lemma. □

Remark 5.5. *Note that by Lemma 2.8, if $(u_1, \ldots, u_m)$ is an m-tuple of Borel functions satisfying inequality (2.2) and $Du_1(x_1^0)$ exists for some $x_1^0 \in X_1$ satisfying $M_{x_1^0 k_1 \ldots k_r} \neq \emptyset$, then the map*

$$(x_2, \ldots, x_m) \mapsto D_{x_1} c(x_1^0, x_2, \ldots, x_m)$$

is injective on the subset of $M_{x_1^0 k_1 \ldots k_r}$ where $D_{x_1} c(x_1^0, x_2, \ldots, x_m)$ exists if and only if $M_{x_1^0 k_1 \ldots k_r}$ is a singleton. As we shall see in the next two sections, this fact will be convenient for the proof of our main result (Theorem 5.6) and the propositions in

Section 5.C.

5.B Existence and Uniqueness to Monge Problem

We now state and prove our main result.

Theorem 5.6. *Assume the measures $\mu_1, \mu_{k_1}, \ldots, \mu_{k_r}$ are absolutely continuous with respect to $\mathcal{L}^n$, with $\{k_i\}_{i=1}^r \subseteq \{2, \ldots, m\}$, $k_1 < k_2 < \ldots < k_r$. Assume c is twisted on c-splitting sets with respect to the variables $x_1, x_{k_1}, \ldots, x_{k_r}$. Then the solution μ in (KP) is concentrated on a graph of a measurable map and it is unique.*

Proof. Let us first prove that μ is induced by a map. The uniqueness assertion will follows immediately by a standard argument. By Theorem 2.2 there exists an m-tuple $(u_1, \ldots, u_m)$ of c-splitting functions for $spt(\mu)$ satisfying (2.1). Fix $i \in \{0, 1, \ldots, r\}$ and set $k_0 = 1$. Note that the function $u_{k_i}(x_{k_i})$ is semi-concave for each k_i, as it is the infimum of semi-concave functions. Hence, $u_{k_i}(x_{k_i})$ is differentiable almost everywhere with respect to $\mathcal{L}^n$. It follows that $u_{k_i}(x_{k_i})$ is differentiable μ_{k_i} almost everywhere, as the measure μ_{k_i} is absolutely continuous with respect to $\mathcal{L}^n$. It implies that $\mu(S) = 1$, where

$$S := \Big\{(x_1, x_2, \ldots, x_m) \in \prod_{i=1}^m X_i : Du_1(x_1) \text{ and } Du_{k_i}(x_{k_i}) \text{ exist for each } i = 1, \ldots r, \text{ and } \sum_{i=1}^m u_i(x_i) = c(x_1, x_2, \ldots, x_m)\Big\}.$$

Fix $x_1^0 \in \pi_1(S)$. Clearly, $M_{x_1^0 k_1 \ldots k_r} \neq \emptyset$, and so by Lemma 5.4 the map $(x_2, \ldots, x_m) \mapsto D_{x_1} c(x_1^0, x_2, \ldots, x_m)$ is injective on the subset of $M_{x_1^0 k_1 \ldots k_r}$ where $D_{x_1} c(x_1^0, x_2, \ldots, x_m)$ exists, this happens if and only if the set $M_{x_1^0 k_1 \ldots k_r}$ is a singleton (see Remark 5.5), which implies $W_{x_1^0 k_1 \ldots k_r}$ is also a singleton. This completes the proof that μ is induced by a map. To prove that μ is unique note that for any pair of solutions μ_1 and μ_2 (which are induced by maps T_1 and T_2), we have $\frac{1}{2}(\mu_1 + \mu_2)$ is also a solution (by the convexity of the set $\Pi(\mu_1, \ldots, \mu_m)$), which implies that it is also concentrated on the graph of some map. However, $\frac{1}{2}(\mu_1 + \mu_2)$ must be concentrated on the union of the graphs of T_1 and T_2. We conclude $T_1 = T_2$ μ_1-a.e., completing the proof of the theorem. □

Remark 5.7. *Note from the above proof that the regularity condition on the first marginal (which is a standard assumption in the classical and multi-marginal optimal transport for uniqueness results), allow us to focus on the set* $\{x_1 \in X_1 : Du_1(x_1) \;\; \text{exists}\}$, *for every* m*-tuple* $(u_1, \ldots, u_m)$ *of Borel functions satisfying inequality* (2.2). *In what follows such regularity condition holds, so to get uniqueness of solutions in the Monge-Kantorovich problem it suffices to prove that the set* $M_{x_1^0 k_1 \ldots k_r}(u_1, \ldots, u_m)$ *is a singleton for every* $x_1^0 \in \{x_1 \in X_1 : Du_1(x_1) \;\; \text{exists}\}$ *fixed, for every* m*-tuple* $(u_1, \ldots, u_m)$ *of Borel functions satisfying inequality* (2.2) *(see also Remark 5.5).*

5.C Examples.

Here, we illustrate the result obtained in Theorem 5.6 throughout several examples.

Proposition 5.8 (One dimensional sub-modular type costs)**.** *Assume* $c(x_1, \ldots, x_m)$ *is semi-concave and* C^2, *where* $X_i = \mathbb{R}$ *for all* $i = 1, \ldots, m$. *Let* G *be an undirected simple graph on* $\{1, \ldots, m\}$ *and assume*

1. $\dfrac{\partial^2 c}{\partial x_i \partial x_j} \leq 0$ *for all* $i \neq j$ *and* $\dfrac{\partial^2 c}{\partial x_i \partial x_j} < 0$ *for all* $\{i, j\} \in E(G)$.

2. *There exists a set* $P := \{k_1, \ldots, k_r\} \subseteq \{1, \ldots, m\}$ *such that for every* $i \in \{1, \ldots, m\}$ *not adjacent to* 1, *there is a path* $\{\{1, i_1\}, \{i_1, i_2\}, \ldots, \{i_{l-1}, i_l\}, \{i_l, i\}\}$ *in* G *with* $\{i_1, \ldots, i_l\} \subseteq P$.

Then c *is twisted on* c*-splitting sets with respect to the variables* $x_1, x_{k_1}, \ldots, x_{k_r}$.

Proof. Let $(u_1, \ldots, u_m)$ be an m-tuple of Borel functions satisfying inequality (2.2) and fix $x_1^0 \in X_1$ such that $Du_1(x_1^0)$ exists and $M_{x_1^0 k_1 \ldots k_r} \neq \emptyset$. We want to prove that $M_{x_1^0 k_1 \ldots k_r}$ is a singleton. This will complete the proof.

Let $(x_2, \ldots, x_m), (\overline{x}_2, \ldots, \overline{x}_m) \in M_{x_1^0 k_1 \ldots k_r}$ and set $x = (x_1^0, x_2, \ldots, x_m)$ and $\overline{x} = (x_1^0, \overline{x}_2, \ldots, \overline{x}_m)$. Consider

$$x^+ = (x_1^0, x_2^+, \ldots, x_m^+) \quad \text{where} \quad x_k^+ = \max\{x_k, \overline{x}_k\},$$

$$x^- = (x_1^0, x_2^-, \ldots, x_m^-) \quad \text{where} \quad x_k^- = \min\{x_k, \overline{x}_k\}.$$

From definition of $M_{x_1^0 k_1 \ldots k_r}$ the set $\{x, \overline{x}\}$ is a c-splitting set, so it is cyclically monotone. Then

$$c(x) + c(\overline{x}) \leq c(x^+) + c(x^-). \tag{5.1}$$

We claim that the reverse inequality also holds. To get this consider $x(t) = tx^+ + (1-t)x$ and $y(t) = t\overline{x} + (1-t)x^-$ for $s \in [0,1]$. Next, write

$$\begin{aligned} c(x^+) - c(x) &= \int_0^1 \frac{d}{dt} c(x(t))dt \\ &= \int_0^1 \sum_{i=2}^m \frac{\partial c(x(t))}{\partial x_i}(x_i^+ - x_i)dt, \end{aligned} \tag{5.2}$$

and

$$\begin{aligned} c(x^-) - c(\overline{x}) &= -\int_0^1 \frac{d}{dt} c(y(t))dt \\ &= -\int_0^1 \sum_{i=2}^m \frac{\partial c(y(t))}{\partial x_i}(\overline{x}_i - x_i^-)dt. \end{aligned} \tag{5.3}$$

Since for each $i \in \{2, \ldots, m\}$ we have

$$x_i^+ - x_i = \overline{x}_i - x_i^- = \begin{cases} \overline{x}_i - x_i & \text{if} \quad \overline{x}_i > x_i \\ 0 & \quad\;\; \overline{x}_i \leq x_i, \end{cases} \tag{5.4}$$

the addition of (5.2) and (5.3) gives

$$c(x^+) - c(x) + c(x^-) - c(\overline{x}) = \int_0^1 \sum_{i=2}^{m} \left[\frac{\partial c(x(t))}{\partial x_i} - \frac{\partial c(y(t))}{\partial x_i} \right] (x_i^+ - x_i) dt \quad (5.5)$$

Now, set $x(t,s) = sx(t) + (1-s)y(t)$, with $t \in [0,1]$ fixed. Then for each $i \in \{2, \dots, m\}$ we have

$$\begin{aligned}
\frac{\partial c(x(t))}{\partial x_i} - \frac{\partial c(y(t))}{\partial x_i} &= \int_0^1 \sum_{j=2}^{m} \frac{\partial^2 c(x(t,s))}{\partial x_i \partial x_j} \left[t x_j^+ + (1-t)x_j - \left(t\overline{x}_j + (1-t)x_j^- \right) \right] ds \\
&= \int_0^1 \sum_{j=2}^{m} \frac{\partial^2 c(x(t,s))}{\partial x_i \partial x_j} \left[t(x_j^+ - x_j) - t(\overline{x}_j - x_j^-) + x_j - x_j^- \right] ds \\
&= \int_0^1 \sum_{j=2}^{m} \frac{\partial^2 c(x(t,s))}{\partial x_i \partial x_j} \left(x_j - x_j^- \right) ds. \qquad \text{by (5.4)}
\end{aligned}$$

Substituting it into (5.5) we get

$$c(x^+) - c(x) + c(x^-) - c(\overline{x}) = \int_0^1 \int_0^1 \sum_{i,j=2}^{m} \frac{\partial^2 c(x(t,s))}{\partial x_i \partial x_j} \left(x_i^+ - x_i \right) \left(x_j - x_j^- \right) ds dt.$$

Now note that $x_i^+ - x_i, x_j - x_j^- \geq 0$, then by Assumption 1, $c(x^+) - c(x) + c(x^-) - c(\overline{x}) \leq 0$. This implies that equality holds in (5.1), completing the proof of the claim. Also, note that if one of the inequalities

$$u_1(x_1^0) + \sum_{i=2}^{m} u_i(x_i^+) \leq c(x^+) \quad (5.6)$$

$$u_1(x_1^0) + \sum_{i=2}^{m} u_i(x_i^-) \leq c(x^-) \quad (5.7)$$

is strict, we would have

$$2u_1(x_1^0) + \sum_{i=2}^{m} u_i(x_i^+) + \sum_{i=2}^{m} u_i(x_i^-) < c(x^+) + c(x^-)$$

$$
\begin{aligned}
&= c(x) + c(\overline{x}) \\
&= 2u_1(x_1^0) + \sum_{i=2}^{m} u_i(x_i) + \sum_{i=2}^{m} u_i(\overline{x_i}) \\
&= 2u_1(x_1^0) + \sum_{i=2}^{m} u_i(x_i^+) + \sum_{i=2}^{m} u_i(x_i^-),
\end{aligned}
$$

which is clearly not posible; that is, equality holds in (5.6) and (5.7). Hence

$$
x^+, x^- \in M_{x_1^0 k_1 \dots k_r}. \tag{5.8}
$$

Furthermore, from Lemma 2.8 we get

$$
\frac{\partial c(x^+)}{\partial x_1} = Du_1(x_1^0) = \frac{\partial c(x^-)}{\partial x_1},
$$

or equivalently,

$$
\int_0^1 \sum_{i=2}^{m} \frac{\partial^2 c(r(t))}{\partial x_1 \partial x_i}(x_i^+ - x_i^-)dt = 0,
$$

where $r(t) = tx^+ + (1-t)x^-$, $t \in [0,1]$. We then must have

$$
\frac{\partial^2 c(r(t))}{\partial x_1 \partial x_i}(x_i^+ - x_i^-) = 0
$$

for every $i \in \{2, \dots, m\}$, as $\frac{\partial^2 c(r(t))}{\partial x_1 \partial x_i}(x_i^+ - x_i^-) \leq 0$ on $\{2, \dots, m\}$. We next use Assumption 1 to deduce $x_i^+ = x_i^-$ for all i adjacent to 1; that is,

$$
x_i = \overline{x}_i \ \text{ for all } i \text{ adjacent to } 1. \tag{5.9}
$$

Now, if 1 is adjacent to all the other vertices, the proof is completed. If there is a vertex not adjacent to 1, then 1 must be adjacent to some $i \in P$ (by Assumption 2), which implies $x_i = \overline{x}_i$ by (5.9). Combining this with (5.8) and Lemma 2.8 we get

$$\frac{\partial c(x^+)}{\partial x_i} = Du_i(x_i) = \frac{\partial c(x^-)}{\partial x_i},$$

so we can mimic the arguments presented in the proof of (5.9) (beginning from (5.8)) to get $x_j = \overline{x}_j$ for every j adjacent to i. Following this iterative process we can prove that $x_j = \overline{x}_j$ for every $j \in V(G)$, as Assumption 2 implies that every vertex of $V(G)$ is adjacent to at least one vertex in P, completing the proof of the proposition. □

Remark 5.9. *Note that if the graph G is complete, we can take $P = \emptyset$ and Condition 1 basically means that c is strictly sub-modular. Unique Monge type solutions for strictly sub-modular costs was established by Carlier [16]. It was observed in [41] that this condition is equivalent (up to a change of variables) to the compatibility condition, which states that*

$$\left(\frac{\partial^2 c}{\partial x_i \partial x_j}\right)\left(\frac{\partial^2 c}{\partial x_k \partial x_j}\right)^{-1}\left(\frac{\partial^2 c}{\partial x_k \partial x_i}\right) < 0$$

everywhere, for all distinct i, j, k, and so compatible costs yield unique Monge solutions as well.

We can easily see that the next result is a generalization of a special case of Theorem 4.9. Note that here we do not require f being symmetric.

Proposition 5.10. *Let $\{I_1, I_2, I_3\}$ be a partition of $\{1, \ldots, m\}$. Let $f : \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}$ be a function satisfying:*

1. *f is bi-linear,*

2. *$f(x, x) \leq 0$ for every $x \in \mathbb{R}^n$,*

3. *f is bi-twisted; that is, for each $x_0, y_0 \in \mathbb{R}^n$ fixed, the maps $y \mapsto D_x f(x_0, y)$ and $x \mapsto D_y f(x, y_0)$ are injective on $\{x_0\} \times \mathbb{R}^n$ and $\mathbb{R}^n \times \{y_0\}$ respectively.*

Assume $1 \in I_1$ *and fix* $p \in I_2 \cup I_3$*, then the cost function*

$$c(x_1, \ldots, x_m) = \sum_{s \in I_1} \sum_{t \in I_2 \cup I_3} f(x_s, x_t) + \sum_{s \in I_3} \sum_{t \in I_2} f(x_s, x_t) + \sum_{\substack{s,t \in I_3 \\ s<t}} f(x_s, x_t) \quad (5.10)$$

is twisted on c*-splitting sets with respect to* x_1 *and* x_p*.*

Proof. Firstly, by hypothesis 1 we can write

$$c(x_1, \ldots, x_m) = f(\sum_{s \in I_1} x_s, \sum_{t \in I_2 \cup I_3} x_t) + f(\sum_{s \in I_3} x_s, \sum_{t \in I_2} x_t) + \sum_{\substack{s,t \in I_3 \\ s<t}} f(x_s, x_t). \quad (5.11)$$

Let $(u_1, \ldots, u_m)$ be an m-tuple of Borel functions satisfying inequality (2.2) and fix $x_1^0 \in \{x_1 \in X_1 : Du_1(x_1) \text{ exists}\}$, with $M_{x_1^0 p}(u_1, \ldots, u_m) \neq \emptyset$. We want to prove that $M_{x_1^0 p}$ is a singleton, this will complete the proof.

Let $(x_2^1, \ldots, x_m^1), (x_2^2, \ldots, x_m^2) \in M_{x_1^0 p}$. Since $\{I_1, I_2, I_3\}$ is a partition and $1 \in I_1$, we get from Lemma 2.8 and (5.11),

$$\begin{aligned} D_{x_1} f(x_1^0, \sum_{t \in I_2 \cup I_3} x_t^1) &= D_{x_1} c(x_1^0, x_2^1, \ldots, x_m^1) \\ &= Du_1(x_1^0) \\ &= D_{x_1} c(x_1^0, x_2^2, \ldots, x_m^2) \\ &= D_{x_1} f(x_1^0, \sum_{t \in I_2 \cup I_3} x_t^2). \end{aligned}$$

It follows that

$$\sum_{t \in I_2 \cup I_3} x_t^1 = \sum_{t \in I_2 \cup I_3} x_t^2, \quad (5.12)$$

by Assumption 3.

Claim 5.11. *For every $N \subseteq I_1$ we get $y_N := (y_2, \dots, y_m) \in M_{x_1^0 p}$, where*

$$y_s = \begin{cases} x_s^2 & \textit{if} \quad s \in \{2, \dots, m\} \setminus N \\ x_s^1 & \textit{if} \quad s \in N. \end{cases} \tag{5.13}$$

Proof of Claim 1. Note that from (5.11), we can write

$$c(x_1, \dots, x_m) = f(\sum_{s \in N} x_s, \sum_{t \in I_2 \cup I_3} x_t) + f(\sum_{s \in I_1 \setminus N} x_s, \sum_{t \in I_2 \cup I_3} x_t) + f(\sum_{s \in I_3} x_s, \sum_{t \in I_2} x_t) + \sum_{\substack{s,t \in I_3 \\ s<t}} f(x_s, x_t). \tag{5.14}$$

Since $(x_2^1, \dots, x_m^1) \in M_{x_1^0 p}$, we get

$$\begin{aligned} \{x_s^1\}_{s \in N} \in \text{Argmin}\Big\{ & \{x_s\}_{s \in N} \mapsto f(\sum_{s \in N} x_s, \sum_{t \in I_2 \cup I_3} x_t^1) - \sum_{s \in N} u_s(x_s) + f(\sum_{s \in I_1 \setminus N} x_s^1, \sum_{t \in I_2 \cup I_3} x_t^1) \\ & + f(\sum_{s \in I_3} x_s^1, \sum_{t \in I_2} x_t^1) + \sum_{\substack{s,t \in I_3 \\ s<t}} f(x_s^1, x_t^1) - \sum_{s \in \{1, \dots, m\} \setminus N} u_s(x_s^1) \Big\} \\ = \text{Argmin}\Big\{ & \{x_s\}_{s \in N} \mapsto f(\sum_{s \in N} x_s, \sum_{t \in I_2 \cup I_3} x_t^1) - \sum_{s \in N} u_s(x_s) \Big\} \\ = \text{Argmin}\Big\{ & \{x_s\}_{s \in N} \mapsto f(\sum_{s \in N} x_s, \sum_{t \in I_2 \cup I_3} x_t^2) - \sum_{s \in N} u_s(x_s) \Big\}, \end{aligned}$$

by (5.12). We deduce $y_N \in M_{x_1^0 p}$, as $(x_2^2, \dots, x_m^2) \in M_{x_1^0 p}$. This complete the proof of Claim 1.

Claim 5.12. $x_s^1 = x_s^2$ *for every* $s \in I_2$.

Proof of Claim 2. From Claim 1, $(y_2, \dots, y_m) \in M_{x_1^0 p}$ where

$$y_s = \begin{cases} x_s^2 & \text{if} \quad s \in \{2, \dots, m\} \setminus I_1 \\ x_s^1 & \text{if} \quad s \in I_1. \end{cases} \tag{5.15}$$

Then, by fixing $r \in I_2$ we get

$$x_r^1 \in \text{Argmin}\left\{ x_r \mapsto c(x_1^0, x_2^1, \ldots, x_{r-1}^1, x_r, x_{r+1}^1, \ldots, x_m^1) - u_r(x_r) \right\},$$

$$y_r = x_r^2 \in \text{Argmin}\left\{ x_r \mapsto c(x_1^0, y_2, \ldots, y_{r-1}, x_r, y_{r+1}, \ldots, y_m) - u_r(x_r) \right\}.$$

Then

$$c(x_1^0, x_2^1, \ldots, x_m^1) - u_r(x_r^1) \leq c(x_1^0, x_2^1, \ldots, x_{r-1}^1, x_r^2, x_{r+1}^1, \ldots, x_m^1) - u_r(x_r^2), \tag{5.16}$$

$$c(x_1^0, y_2, \ldots, y_m) - u_r(x_r^2) \leq c(x_1^0, y_2, \ldots, y_{r-1}, x_r^1, y_{r+1}, \ldots, y_m) - u_r(x_r^1), \tag{5.17}$$

which implies

$$\begin{aligned} c(x_1^0, x_2^1, \ldots, x_m^1) + c(x_1^0, y_2, \ldots, y_m) \leq\ & c(x_1^0, x_2^1, \ldots, x_{r-1}^1, x_r^2, x_{r+1}^1, \ldots, x_m^1) \\ & + c(x_1^0, y_2, \ldots, y_{r-1}, x_r^1, y_{r+1}, \ldots, y_m). \end{aligned} \tag{5.18}$$

Now, from bi-linearity of f we can write

$$\begin{aligned} c(x_1, \ldots, x_m) &= f(\sum_{s \in I_1} x_s, \sum_{t \in I_2 \cup I_3} x_t) + f(\sum_{s \in I_3} x_s, \sum_{t \in I_2} x_t) + \sum_{\substack{s,t \in I_3 \\ s<t}} f(x_s, x_t) \\ &= f(\sum_{s \in I_1 \cup I_3} x_s, \sum_{t \in I_2} x_t) + f(\sum_{s \in I_1} x_s, \sum_{t \in I_3} x_t) + \sum_{\substack{s,t \in I_3 \\ s<t}} f(x_s, x_t) \\ &= f(\sum_{s \in I_1 \cup I_3} x_s, x_r) + f(\sum_{s \in I_1 \cup I_3} x_s, \sum_{t \in I_2 \setminus \{r\}} x_t) + f(\sum_{s \in I_1} x_s, \sum_{t \in I_3} x_t) + \sum_{\substack{s,t \in I_3 \\ s<t}} f(x_s, x_t). \end{aligned}$$

Combining this with (5.18) we get

$$\begin{aligned}
&f(\sum_{s\in I_1\cup I_3} x_s^1, x_r^1) + f(\sum_{s\in I_1\cup I_3} x_s^1, \sum_{t\in I_2\setminus\{r\}} x_t^1) + f(\sum_{s\in I_1} x_s^1, \sum_{t\in I_3} x_t^1) + \sum_{\substack{s,t\in I_3\\ s<t}} f(x_s^1, x_t^1)\\
&+f(\sum_{s\in I_1\cup I_3} y_s, y_r) + f(\sum_{s\in I_1\cup I_3} y_s, \sum_{t\in I_2\setminus\{r\}} y_t) + f(\sum_{s\in I_1} y_s, \sum_{t\in I_3} y_t) + \sum_{\substack{s,t\in I_3\\ s<t}} f(y_s, y_t)\\
\le\ & f(\sum_{s\in I_1\cup I_3} x_s^1, x_r^2) + f(\sum_{s\in I_1\cup I_3} x_s^1, \sum_{t\in I_2\setminus\{r\}} x_t^1) + f(\sum_{s\in I_1} x_s^1, \sum_{t\in I_3} x_t^1) + \sum_{\substack{s,t\in I_3\\ s<t}} f(x_s^1, x_t^1)\\
&+f(\sum_{s\in I_1\cup I_3} y_s, x_r^1) + f(\sum_{s\in I_1\cup I_3} y_s, \sum_{t\in I_2\setminus\{r\}} y_t) + f(\sum_{s\in I_1} y_s, \sum_{t\in I_3} y_t) + \sum_{\substack{s,t\in I_3\\ s<t}} f(y_s, y_t).
\end{aligned}$$

Then, the above inequality reduces to

$$f(\sum_{s\in I_1\cup I_3} x_s^1, x_r^1) + f(\sum_{s\in I_1\cup I_3} y_s, y_r) \le f(\sum_{s\in I_1\cup I_3} x_s^1, x_r^2) + f(\sum_{s\in I_1\cup I_3} y_s, x_r^1). \quad (5.19)$$

By construction of y and linearity we have

$$f(\sum_{s\in I_1\cup I_3} y_s, y_r) = f(\sum_{s\in I_1} y_s, y_r) + f(\sum_{s\in I_3} y_s, y_r) = f(\sum_{s\in I_1} x_s^1, x_r^2) + f(\sum_{s\in I_3} x_s^2, x_r^2),$$

$$f(\sum_{s\in I_1\cup I_3} y_s, x_r^1) = f(\sum_{s\in I_1} y_s, x_r^1) + f(\sum_{s\in I_3} y_s, x_r^1) = f(\sum_{s\in I_1} x_s^1, x_r^1) + f(\sum_{s\in I_3} x_s^2, x_r^1).$$

Substituting it into (5.19) and eliminating similar terms we get

$$f(\sum_{s\in I_3} x_s^1, x_r^1) + f(\sum_{s\in I_3} x_s^2, x_r^2) \le f(\sum_{s\in I_3} x_s^1, x_r^2) + f(\sum_{s\in I_3} x_s^2, x_r^1),$$

and by (5.12), we get

$$f(\sum_{s\in I_2} (x_s^2 - x_s^1), x_r^1 - x_r^2) = f(\sum_{s\in I_3} (x_s^1 - x_s^2), x_r^1 - x_r^2) \le 0;$$

that is,

$$f(\sum_{s\in I_2}(x_s^2-x_s^1),x_r^2-x_r^1)\geq 0.$$

Summing over $r\in I_2$ we get

$$f(\sum_{s\in I_2}(x_s^2-x_s^1),\sum_{r\in I_2}(x_r^2-x_r^1))\geq 0.$$

Combining this with hypothesis 2 we get $f(\sum_{s\in I_2}(x_s^2-x_s^1),\sum_{r\in I_2}(x_r^2-x_r^1))=0.$ Then, we must have, in particular, equality in (5.16). It follows that $(x_2^1,\dots,x_{r-1}^1,x_r^2,x_{r+1}^1,\dots,x_m^1)\in M_{x_1^0p}$, and so

$$\begin{aligned}D_{x_1}f(x_1^0,\sum_{t\in I_2\cup I_3}x_t^1)&=D_{x_1}c(x_1^0,x_2^1,\dots,x_m^1)\\&=Du_1(x_1^0)\\&=D_{x_1}c(x_1^0,x_2^1,\dots,x_{r-1}^1,x_r^2,x_{r+1}^1,\dots,x_m^1)\\&=D_{x_1}f(x_1^0,x_r^2+\sum_{t\in I_2\cup I_3\setminus\{r\}}x_t^1).\end{aligned}$$

Thus, $x_r^1=x_r^2$, as f is twisted. This completes the proof of Claim 2.

Claim 5.13. *For every n, equation (5.12) implies $x_{t_j}^1=x_{t_j}^2$ for $1\leq j\leq n$, where $I_3:=\{t_1,\dots,t_n\}$.*

Proof of Claim 3. From Claim 2 and (5.12) we get

$$\sum_{t\in I_3}x_t^1=\sum_{t\in I_3}x_t^2. \tag{5.20}$$

We proceed to apply induction on n. Indeed, when $n=1$ it is clearly true. Assume the statement is true when $n=k-1$. Note that

$$f(\sum_{s\in I_1}x_s,\sum_{t\in I_3}x_t)=f(\sum_{s\in I_1}x_s,x_{t_k})+f(\sum_{s\in I_1}x_s,\sum_{t\in I_3\setminus\{t_k\}}x_t),$$

$$f(\sum_{s\in I_3} x_s, \sum_{t\in I_2} x_t) = f(x_{t_k}, \sum_{t\in I_2} x_t) + f(\sum_{s\in I_3\setminus\{t_k\}} x_s, \sum_{t\in I_2} x_t),$$

$$\sum_{\substack{s,t\in I_3\\ s<t}} f(x_s, x_t) = \sum_{s\in I_3\setminus\{t_k\}} f(x_s, x_{t_k}) + \sum_{\substack{s,t\in I_3\setminus\{t_k\}\\ s<t}} f(x_s, x_t).$$

Hence,

$$\begin{aligned}
c(x_1,\ldots,x_m) &= f(\sum_{s\in I_1} x_s, \sum_{t\in I_2\cup I_3} x_t) + f(\sum_{s\in I_3} x_s, \sum_{t\in I_2} x_t) + \sum_{\substack{s,t\in I_3\\ s<t}} f(x_s, x_t) \\
&= f(\sum_{s\in I_1} x_s, \sum_{t\in I_2} x_t) + f(\sum_{s\in I_1} x_s, \sum_{t\in I_3} x_t) + f(\sum_{s\in I_3} x_s, \sum_{t\in I_2} x_t) + \sum_{\substack{s,t\in I_3\\ s<t}} f(x_s, x_t) \\
&= f(\sum_{s\in I_1} x_s, \sum_{t\in I_2} x_t) + f(\sum_{s\in I_1} x_s, x_{t_k}) + f(\sum_{s\in I_1} x_s, \sum_{t\in I_3\setminus\{t_k\}} x_t) + f(x_{t_k}, \sum_{t\in I_2} x_t) \\
&\quad + f(\sum_{s\in I_3\setminus\{t_k\}} x_s, \sum_{t\in I_2} x_t) + \sum_{s\in I_3\setminus\{t_k\}} f(x_s, x_{t_k}) + \sum_{\substack{s,t\in I_3\setminus\{t_k\}\\ s<t}} f(x_s, x_t).
\end{aligned}$$

Since the only terms of c depending on x_{t_k} are $f(\sum_{s\in I_1} x_s, x_{t_k}), f(x_{t_k}, \sum_{t\in I_2} x_t)$ and $\sum_{s\in I_3\setminus\{t_k\}} f(x_s, x_{t_k})$, we get

$$x^1_{t_k} \in \text{Argmin}\left\{x_{t_k} \mapsto f(\sum_{s\in I_1} x^1_s, x_{t_k}) + f(x_{t_k}, \sum_{t\in I_2} x^1_t) + \sum_{s\in I_3\setminus\{t_k\}} f(x^1_s, x_{t_k}) - u_{t_k}(x_{t_k})\right\}.$$

Furthermore, defining y as in (5.15) we get

$$y_{t_k} = x^2_{t_k} \in \text{Argmin}\left\{x_{t_k} \mapsto f(\sum_{s\in I_1} y_s, x_{t_k}) + f(x_{t_k}, \sum_{t\in I_2} y_t) + \sum_{s\in I_3\setminus\{t_k\}} f(y_s, x_{t_k}) - u_{t_k}(x_{t_k})\right\}.$$

We deduce

$$\begin{aligned}
&f(\sum_{s\in I_1} x^1_s, x^1_{t_k}) + f(x^1_{t_k}, \sum_{t\in I_2} x^1_t) + \sum_{s\in I_3\setminus\{t_k\}} f(x^1_s, x^1_{t_k}) - u_{t_k}(x^1_{t_k}) \\
\leq\ & f(\sum_{s\in I_1} x^1_s, x^2_{t_k}) + f(x^2_{t_k}, \sum_{t\in I_2} x^1_t) + \sum_{s\in I_3\setminus\{t_k\}} f(x^1_s, x^2_{t_k}) - u_{t_k}(x^2_{t_k}),
\end{aligned} \tag{5.21}$$

$$f(\sum_{s\in I_1} y_s, x^2_{t_k}) + f(x^2_{t_k}, \sum_{t\in I_2} y_t) + \sum_{s\in I_3\setminus\{t_k\}} f(y_s, x^2_{t_k}) - u_{t_k}(x^2_{t_k})$$
$$\leq f(\sum_{s\in I_1} y_s, x^1_{t_k}) + f(x^1_{t_k}, \sum_{t\in I_2} y_t) + \sum_{s\in I_3\setminus\{t_k\}} f(y_s, x^1_{t_k}) - u_{t_k}(x^1_{t_k}).$$

Adding the above inequalities, using Claim 2 and construction of y we get

$$\sum_{s\in I_3\setminus\{t_k\}} f(x^1_s, x^1_{t_k}) + \sum_{s\in I_3\setminus\{t_k\}} f(x^2_s, x^2_{t_k}) \leq \sum_{s\in I_3\setminus\{t_k\}} f(x^1_s, x^2_{t_k}) + \sum_{s\in I_3\setminus\{t_k\}} f(x^2_s, x^1_{t_k}).$$

Combining this with (5.20) we get

$$f(x^2_{t_k} - x^1_{t_k}, x^2_{t_k} - x^1_{t_k}) = f(\sum_{s\in I_3\setminus\{t_k\}} (x^1_s - x^2_s), x^2_{t_k} - x^1_{t_k}) \geq 0.$$

From hypothesis 2 we then get $f(x^2_{t_k} - x^1_{t_k}, x^2_{t_k} - x^1_{t_k}) = 0$. Hence, equality holds in (5.21) and $(x^1_2, \dots, x^1_{t_k-1}, x^2_{t_k}, x^1_{t_k+1}, \dots, x^1_m) \in M_{x^0_1 p}$. This implies

$$\begin{aligned} D_{x_1} f(x^0_1, \sum_{t\in I_2\cup I_3} x^1_t) &= D_{x_1} c(x^0_1, x^1_2, \dots, x^1_m) \\ &= Du_1(x^0_1) \\ &= D_{x_1} c(x^0_1, x^1_2, \dots, x^1_{t_k-1}, x^2_{t_k}, x^1_{t_k+1}, \dots, x^1_m) = D_{x_1} f(x^0_1, x^2_{t_k} + \sum_{t\in I_2\cup I_3\setminus\{t_k\}} x^1_t). \end{aligned}$$

Thus, $x^1_{t_k} = x^2_{t_k}$, as f is twisted. Hence, from (5.20) and Claim 2 we can write $\sum_{t\in I_2\cup I_3\setminus\{t_k\}} x^1_t = \sum_{t\in I_2\cup I_3\setminus\{t_k\}} x^2_t$. It follows that $x^1_{t_2} = x^2_{t_2}, \dots, x^1_{t_{k-1}} = x^2_{t_{k-1}}$, by induction hypothesis. This completes the proof of Claim 3.

Claim 5.14. $x^1_s = x^2_s$ *for every* $s \in I_1$.

Proof of Claim 3. Since $p \in I_2 \cup I_3$, $x^1_p = x^2_p$ by Claim 2 and 3. Hence,

$$\begin{aligned} D_{x_p} c(x^0_1, x^1_2, \dots, x^1_m) &= Du_p(x^1_p) \\ &= Du_p(x^2_p) \end{aligned}$$

$$= D_{x_p} c(x_1^0, x_2^2, \ldots, x_m^2).$$

Combining the above equality, Claim 2 and 3, and (5.11) we get $D_{x_p} f(\sum_{t \in I_1} x_t^1, x_p^1) = D_{x_p} f(\sum_{t \in I_1} x_t^2, x_p^2)$. It follows that

$$\sum_{t \in I_1} x_t^1 = \sum_{t \in I_1} x_t^2. \tag{5.22}$$

Now, fix $t \in I_1 \backslash \{1\}$. Setting $N = \{t\}$, we use Claim 1 to get $y_N = (x_2^2, \ldots, x_{t-1}^2, x_t^1, x_{t+1}^2, \ldots, x_m^2) \in M_{x_1^0 p}$. Since (5.22) holds true for every $(x_2^1, \ldots, x_m^1), (x_2^2, \ldots, x_m^2) \in M_{1p}$, in particular, it is true for y_N and $(x_2^2, \ldots, x_m^2)$. It immediately implies $x_t^1 = x_t^2$, completing the proof of Claim 4.

This completes the proof of the Proposition. □

The next result focuses on a cost with a cycle structure that generalizes Theorem 3.1.

Proposition 5.15. *Consider*

$$c(x_1, x_2, x_3, x_4) = c_1(x_1, x_2) + c_2(x_2, x_3) + c_3(x_3, x_4) + c_4(x_4, x_1), \tag{5.23}$$

with c_i semi-concave for each $i = 1, 2, 3, 4$. Assume

1. *For every 4-tuple of Borel functions (u_1, u_2, u_3, u_4) satisfying inequality (2.2) and $x_1^0 \in \{x_1 \in X_1 : Du_1(x_1)$ exists$\}$, we get*

$$c_2(x_2^1, x_3^1) + c_3(x_3^1, x_4^1) + c_2(x_2^2, x_3^2) + c_3(x_3^2, x_4^2) \geq c_2(x_2^1, x_3^2) + c_3(x_3^2, x_4^1) + c_2(x_2^2, x_3^1) + c_3(x_3^1, x_4^2), \tag{5.24}$$

for every $(x_2^1, x_3^1, x_4^1), (x_2^2, x_3^2, x_4^2) \in M_{x_1^0 4}$.

2. *c_3 is twisted with respect to x_4; that is, for every x_4 fixed the map $x_3 \mapsto D_{x_4}c_3(x_3, x_4)$ is injective on the subset of $X_3 \times \{x_4\}$ where c_3 is differentiable with respect to x_4.*

3. *c_1 and c_4 are twisted with respect to x_1 respectively; that is, for every x_1 fixed the maps $x_2 \mapsto D_{x_1}c_1(x_1, x_2)$ and $x_4 \mapsto D_{x_1}c_4(x_4, x_1)$ are injective on the subsets of $\{x_1\} \times X_2$ and $X_4 \times \{x_1\}$ where c_1 and c_4 are differentiable with respect to x_1 respectively.*

Then, c is twisted on c-splitting sets with respect to x_1 and x_4.

Proof. Let (u_1, u_2, u_3, u_4) be a 4-tuple of Borel functions satisfying inequality (2.2). Fix $x_1^0 \in \{x_1 \in X_1 : Du_1(x_1) \text{ exists}\}$ and let $(x_2^1, x_3^1, x_4^1), (x_2^2, x_3^2, x_4^2) \in M_{x_1^0 4}$. We want to show $x_i^1 = x_i^2$, $i = 2, 3, 4$. For this, observe that

$$(x_3^k, x_4^k) \in \text{Argmin}\Big\{(x_3, x_4) \mapsto c(x_1^0, x_2^k, x_3, x_4) - u_3(x_3) - u_4(x_4)\Big\},\ k = 1, 2.$$

Then

$$c(x_1^0, x_2^1, x_3^1, x_4^1) - u_3(x_3^1) - u_4(x_4^1) \le c(x_1^0, x_2^1, x_3^2, x_4^1) - u_3(x_3^2) - u_4(x_4^1), \quad (5.25)$$

$$c(x_1^0, x_2^2, x_3^2, x_4^2) - u_3(x_3^2) - u_4(x_4^2) \le c(x_1^0, x_2^2, x_3^1, x_4^2) - u_3(x_3^1) - u_4(x_4^2). \quad (5.26)$$

Adding the above inequalities and eliminating similar terms we get

$$c_2(x_2^1, x_3^1)+c_3(x_3^1, x_4^1)+c_2(x_2^2, x_3^2)+c_3(x_3^2, x_4^2) \le c_2(x_2^1, x_3^2)+c_3(x_3^2, x_4^1)+c_2(x_2^2, x_3^1)+c_3(x_3^1, x_4^2). \quad (5.27)$$

By Assumption 1, the above inequality is in fact equality, which implies that we must have equality in (5.25) and (5.26). In particular, $(x_2^1, x_3^2, x_4^1) \in M_{x_1^0 4}$, so by Lemma 2.8 we get

$$D_{x_4}c(x_1^0, x_2^1, x_3^2, x_4^1) = Du_4(x_4^1) = D_{x_4}c(x_1^0, x_2^1, x_3^1, x_4^1),$$

or equivalently,

$$D_{x_4}c_3(x_3^2,x_4^1)+D_{x_4}c_4(x_4^1,x_1^0)=Du_4(x_4^1)=D_{x_4}c_3(x_3^1,x_4^1)+D_{x_4}c_4(x_4^1,x_1^0).$$

The above equalities gives $D_{x_4}c_3(x_3^2,x_4^1)=D_{x_4}c_3(x_3^1,x_4^1)$, and by Assumption 2, $x_3^1=x_3^2$. Now, note that

$$\begin{aligned}x_4^1 &\in \text{Argmin}\left\{x_4\mapsto c(x_1^0,x_2^1,x_3^1,x_4)-u_4(x_4)\right\}\\ &=\text{Argmin}\left\{x_4\mapsto c_3(x_3^1,x_4)+c_4(x_4,x_1^0)-u_4(x_4)\right\}\\ &=\text{Argmin}\left\{x_4\mapsto c_3(x_3^2,x_4)+c_4(x_4,x_1^0)-u_4(x_4)\right\}\\ &=\text{Argmin}\left\{x_4\mapsto c(x_1^0,x_2^2,x_3^2,x_4)-u_4(x_4)\right\},\end{aligned}$$

as $(x_2^2,x_3^2,x_4^2)\in M_{x_1^0 4}$. Hence, $(x_1^0,x_2^2,x_3^1,x_4^1)=(x_1^0,x_2^2,x_3^2,x_4^1)\in M_{x_1^0 4}$, and by Lemma 2.8 we get $D_{x_1}c(x_1^0,x_2^2,x_3^1,x_4^1)=Du_1(x_1^0)=D_{x_1}c(x_1^0,x_2^1,x_3^1,x_4^1)$; that is, $D_{x_1}c_1(x_1^0,x_2^2)+D_{x_1}c_4(x_4^1,x_1^0)=Du_1(x_1^0)=D_{x_1}c_1(x_1^0,x_2^1)+D_{x_1}c_4(x_4^1,x_1^0)$. Thus, $D_{x_1}c_1(x_1^0,x_2^2)=D_{x_1}c_1(x_1^0,x_2^1)$ and by Assumption 3, $x_2^1=x_2^2$. Finally, we clearly have $(x_2^2,x_3^2,x_4^1)=(x_2^1,x_3^1,x_4^1)\in M_{x_1^0 4}$, hence applying one more time Lemma 2.8 we get $D_{x_1}c(x_1^0,x_2^2,x_3^2,x_4^1)=Du_1(x_1^0)=D_{x_1}c(x_1^0,x_2^2,x_3^2,x_4^2)$. It follows that $D_{x_1}c_4(x_4^1,x_1^0)=D_{x_1}c_4(x_4^2,x_1^0)$, and by Assumption 3, $x_4^1=x_4^2$. This completes the proof of the proposition. □

Note that it is not hard to find costs of the form (5.23) satisfying Assumptions 2 and 3. Assumption 1, on the other hand, is less common. We proceed now to illustrate the previous proposition with an example, which can also be seen as a slight generalization of Proposition 5.10 when $m=4$, I_3 is empty, $I_1=\{1,3\}$ and $I_2=\{2,4\}$. Note that the bi-linearity assumption from Proposition 5.10 is relaxed here.

Example 5.16. *For the cost* (5.23)*, take* $c_1(x_1,x_2)=f(x_1,x_2)$*,* $c_2(x_2,x_3)=f(x_3,x_2)$*,* $c_3(x_3,x_4)=f(x_3,x_4)$ *and* $c_4(x_4,x_1)=f(x_1,x_4)$*, where* $f:\mathbb{R}^n\times\mathbb{R}^n\mapsto$

$\mathbb{R}$ *is a map satisfying:*

(i) f *is additive with respect to the second coordinate; that is,* $f(x, y+z) = f(x,y) + f(x,z)$ *for every* x *fixed.*

(ii) f *is bi-twisted; that is, the maps* $y \mapsto D_x f(x,y)$ *and* $x \mapsto D_y f(x,y)$ *are injective.*

Substituting into (5.23) *and using* (i) *we get*

$$\begin{aligned} c(x_1,x_2,x_3,x_4) &= f(x_1,x_2) + f(x_3,x_2) + f(x_3,x_4) + f(x_1,x_4) \\ &= f(x_1, x_2 + x_4) + f(x_3, x_2 + x_4) \end{aligned}$$

Now, let (u_1,u_2,u_3,u_4) *be a 4-tuple of Borel functions satisfying inequality* (2.2). *Fix* $x_1^0 \in \{x_1 \in X_1 : Du_1(x_1)$ *exists*$\}$ *and let* $(x_2^1, x_3^1, x_4^1), (x_2^2, x_3^2, x_4^2) \in M_{x_1^0 4}$. *From Lemma* 2.8*,*

$$\begin{aligned} D_{x_1} f(x_1^0, x_2^1 + x_4^1) &= D_{x_1} c(x_1^0, x_2^1, x_3^1, x_4^1) \\ &= Du_1(x_1^0) \\ &= D_{x_1} c(x_1^0, x_2^2, x_3^2, x_4^2) \\ &= D_{x_1} f(x_1^0, x_2^2 + x_4^2). \end{aligned}$$

From Assumption (ii), we deduce

$$x_2^1 + x_4^1 = x_2^2 + x_4^2. \tag{5.28}$$

It follows that

$$\begin{aligned} c_2(x_2^1,x_3^1) + c_3(x_3^1,x_4^1) + c_2(x_2^2,x_3^2) + c_3(x_3^2,x_4^2) &= f(x_3^1,x_2^1) + f(x_3^1,x_4^1) + f(x_3^2,x_2^2) + f(x_3^2,x_4^2) \\ &= f(x_3^1, x_2^1 + x_4^1) + f(x_3^2, x_2^2 + x_4^2) \\ &= f(x_3^1, x_2^2 + x_4^2) + f(x_3^2, x_2^1 + x_4^1) \end{aligned}$$

$$= f(x_3^2, x_2^1) + f(x_3^2, x_4^1) + f(x_3^1, x_2^2) + f(x_3^1, x_4^2)$$
$$= c_2(x_2^1, x_3^2) + c_3(x_3^2, x_4^1) + c_2(x_2^2, x_3^1) + c_3(x_3^1, x_4^2).$$

Thus, Condition 1 in Proposition 5.15 is trivially satisfied. Since Conditions 2 and 3 are also satisfied (by (ii)), we obtain that c is twisted on c-splitting sets with respect to x_1 and x_4.

The next Proposition was obtained from some of the essential ideas of Theorem 5.1 in [35], which provides Monge structure and uniqueness of the optimal measures in infimal convolution type examples.

Proposition 5.17. *Let $m_0 := 1 < m_1 < \ldots < m_n := m$. Set $Y_j := (x_{m_{j-1}+1}, \ldots, x_{m_j})$ and $(x_{m_{j-1}}, Y_j) := (x_{m_{j-1}}, x_{m_{j-1}+1}, \ldots, x_{m_j})$, where $j = 1, \ldots, n$, and $(x_{m_0}, Y_1, Y_2, \ldots, Y_n) = (x_1, \ldots, x_m)$. Consider the cost*

$$c(x_1, \ldots, x_m) = \sum_{j=1}^{n} c_j(x_{m_{j-1}}, Y_j), \tag{5.29}$$

and assume

1. *c_j semi-concave for each j.*

2. *c_j is twisted on c_j-splitting sets with respect to $x_{m_{j-1}}$; that is, for each c_j-splitting set $S^j \subseteq X_{m_{j-1}} \times \ldots \times X_{m_j}$ and $x_{m_{j-1}} \in \pi_{m_{j-1}}(S^j)$, where $\pi_{m_{j-1}} : X_{m_{j-1}} \times \ldots \times X_{m_j} \mapsto X_{m_{j-1}}$ is the canonical projection, the map $Y_j \mapsto D_{x_{m_{j-1}}} c_j(x_{m_{j-1}}, Y_j)$ is injective on the subset of S^j where $D_{x_{m_{j-1}}} c_j(x_{m_{j-1}}, Y_j)$ exists.*

Then, the cost $c(x_1, \ldots, x_m)$ is twisted on c-splitting sets with respect to $x_1, x_{m_1}, \ldots, x_{m_{n-1}}$.

Proof. Fix $j \in \{1, \ldots, n\}$. Let us first prove that for every c-splitting set $S \subseteq \prod_{i=1}^{m} X_i$, the set $S^j := \pi_{x_{m_{j-1}} \ldots x_{m_j}}(S)$ is a c_j-splitting set on $\prod_{i=m_{j-1}}^{m_j} X_i$, or equivalently, a c_j-cyclical monotone set on $\prod_{i=m_{j-1}}^{m_j} X_i$, where $\pi_{x_{m_{j-1}}\cdots x_{m_j}} : X \mapsto \prod_{i=m_{j-1}}^{m_j} X_i$ is the canonical projection. Indeed, fix S a c-splitting set on X, and let $\left\{(x^k_{m_{j-1}}, \ldots, x^k_{m_j})\right\}_{k=1}^{p} \subseteq S^j$ and $\sigma_{m_{j-1}}, \ldots, \sigma_{m_j} \in S_P$, where S_P denotes the set of permutations of $P = \{1, \ldots, p\}$. We want to show

$$\sum_{k=1}^{p} c_j(x^k_{m_{j-1}}, Y^k_j) = \sum_{k=1}^{p} c_j(x^k_{m_{j-1}}, \ldots, x^k_{m_j}) \leq \sum_{k=1}^{p} c_j(x^{\sigma_{m_{j-1}}(k)}_{m_{j-1}}, \ldots, x^{\sigma_{m_j}(k)}_{m_j}). \tag{5.30}$$

Note that for each $k \in P$, there are $Y^k_s = (x^k_{m_{s-1}+1}, \ldots, x^k_{m_s})$, $s \neq j$, such that $(x^k_1, Y^k_1, Y^k_2, \ldots, Y^k_n) \in S$. Set

$$\sigma_i = \begin{cases} \sigma_{m_{j-1}} & \text{if} \quad 1 \leq i \leq m_{j-1} \\ \sigma_{m_j} & \text{if} \quad m_j \leq i \leq m. \end{cases} \tag{5.31}$$

Since S is c-cyclically monotone we get

$$\begin{aligned}
&\sum_{k=1}^{p} c_j(x^k_{m_{j-1}}, Y^k_j) + \sum_{s=1}^{j-1}\sum_{k=1}^{p} c_s(x^k_{m_{s-1}}, Y^k_s) + \sum_{s=j+1}^{n}\sum_{k=1}^{p} c_s(x^k_{m_{s-1}}, Y^k_s) \\
&= \sum_{s=1}^{n}\sum_{k=1}^{p} c_s(x^k_{m_{s-1}}, Y^k_s) \\
&= \sum_{k=1}^{p} c(x^k_1, \ldots, x^k_m) \\
&\leq \sum_{k=1}^{p} c(x^{\sigma_1(k)}_1, \ldots, x^{\sigma_m(k)}_m) \\
&= \sum_{k=1}^{p} c_j(x^{\sigma_{m_{j-1}}(k)}_{m_{j-1}}, x^{\sigma_{m_{j-1}+1}(k)}_{m_{j-1}+1}, \ldots, x^{\sigma_{m_j}(k)}_{m_j}) + \sum_{s=1}^{j-1}\sum_{k=1}^{p} c_s(x^{\sigma_{m_{s-1}}(k)}_{m_{s-1}}, x^{\sigma_{m_{s-1}+1}(k)}_{m_{s-1}+1}, \ldots, x^{\sigma_{m_s}(k)}_{m_s}) \\
&+ \sum_{s=j+1}^{n}\sum_{k=1}^{p} c_s(x^{\sigma_{m_{s-1}}(k)}_{m_{s-1}}, x^{\sigma_{m_{s-1}+1}(k)}_{m_{s-1}+1}, \ldots, x^{\sigma_{m_s}(k)}_{m_s})
\end{aligned} \tag{5.32}$$

From (5.31) we have

$$\sum_{s=1}^{j-1}\sum_{k=1}^{p} c_s(x_{m_{s-1}}^{\sigma_{m_{s-1}}(k)}, x_{m_{s-1}+1}^{\sigma_{m_{s-1}+1}(k)}, \dots, x_{m_s}^{\sigma_{m_s}(k)}) = \sum_{s=1}^{j-1}\sum_{k=1}^{p} c_s(x_{m_{s-1}}^{\sigma_{m_{j-1}}(k)}, x_{m_{s-1}+1}^{\sigma_{m_{j-1}}(k)}, \dots, x_{m_s}^{\sigma_{m_{j-1}}(k)})$$
$$= \sum_{s=1}^{j-1}\sum_{k=1}^{p} c_s(x_{m_{s-1}}^{k}, x_{m_{s-1}+1}^{k}, \dots, x_{m_s}^{k})$$
$$= \sum_{s=1}^{j-1}\sum_{k=1}^{p} c_s(x_{m_{s-1}}^{k}, Y_s^k), \qquad (5.33)$$

$$\sum_{s=j+1}^{n}\sum_{k=1}^{p} c_s(x_{m_{s-1}}^{\sigma_{m_{s-1}}(k)}, x_{m_{s-1}+1}^{\sigma_{m_{s-1}+1}(k)}, \dots, x_{m_s}^{\sigma_{m_s}(k)}) = \sum_{s=j+1}^{n}\sum_{k=1}^{p} c_s(x_{m_{s-1}}^{\sigma_{m_j}(k)}, x_{m_{s-1}+1}^{\sigma_{m_j}(k)}, \dots, x_{m_s}^{\sigma_{m_j}(k)})$$
$$= \sum_{s=j+1}^{n}\sum_{k=1}^{p} c_s(x_{m_{s-1}}^{k}, x_{m_{s-1}+1}^{k}, \dots, x_{m_s}^{k})$$
$$= \sum_{s=j+1}^{n}\sum_{k=1}^{p} c_s(x_{m_{s-1}}^{k}, Y_s^k). \qquad (5.34)$$

Substituting the above equalities into inequality (5.32) we get (5.30); that is, S^j is a c_j-splitting set on $\prod_{i=m_{j-1}}^{m_j} X_i$.

Now, let $(u_1, \dots, u_m)$ be an m-tuple of c-splitting functions for S and fix $x_1^0 \in \pi_1(S)$. Assume $D_{x_1}c(x_1^0, x_2^1, \dots, x_m^1)$ and $D_{x_1}c(x_1^0, x_2^2, \dots, x_m^2)$ exist, and

$$D_{x_1}c(x_1^0, x_2^1, \dots, x_m^1) = D_{x_1}c(x_1^0, x_2^2, \dots, x_m^2),$$

where $(x_2^1, \dots, x_m^1), (x_2^2, \dots, x_m^2) \in W_{x_1^0 m_1 \dots m_{n-1}}$. Since c_j does not depend on x_1 for every $j \in \{2, \dots, n\}$, we immediately get

$$D_{x_1}c_1(x_1^0, x_2^1, \dots, x_{m_1}^1) = D_{x_1}c_1(x_1^0, x_2^2, \dots, x_{m_1}^2),$$

then

$$x_j^1 = x_j^2 \text{ for every } j \in \{2, \dots, m_1\}, \tag{5.35}$$

as clearly $(x_1^0, x_2^1, \dots, x_{m_1}^1), (x_1^0, x_2^2, \dots, x_{m_1}^2) \in S^1$ and c_1 is twisted on the c_1-splitting set S^1. In particular, $x_{m_1}^1 = x_{m_1}^2$ and by Lemma 2.8,

$$D_{x_{m_1}} c(x_1^0, x_2^1, \dots, x_m^1) = Du_{m_1}(x_{m_1}^1) = Du_{m_1}(x_{m_1}^2) = D_{x_{m_1}} c(x_1^0, x_2^2, \dots, x_m^2)$$

(here the differentiability of u_{m_1} at $x_{m_1}^1$ follows from the fact that $(x_2^1, \dots, x_m^1), (x_2^2, \dots, x_m^2) \in W_{x_1^0 m_1 \dots m_{n-1}}$). Hence,

$$D_{x_{m_1}} c_1(x_1^0, x_2^1, \dots, x_{m_1}^1) + D_{x_{m_1}} c_2(x_{m_1}^1, \dots, x_{m_2}^1) = D_{x_{m_1}} c_1(x_1^0, x_2^2, \dots, x_{m_1}^2) + D_{x_{m_1}} c_2(x_{m_1}^2, \dots, x_{m_2}^2).$$

Combining this with (5.35) we get

$$D_{x_{m_1}} c_2(x_{m_1}^1, x_{m_1+1}^1 \dots, x_{m_2}^1) = D_{x_{m_1}} c_2(x_{m_1}^1, x_{m_1+1}^2 \dots, x_{m_2}^2).$$

Since c_2 is twisted on the c_2-splitting set S^2 and $(x_{m_1}^1, x_{m_1+1}^1 \dots, x_{m_2}^1), (x_{m_1}^1, x_{m_1+1}^2 \dots, x_{m_2}^2) \in S^2$, we deduce $x_j^1 = x_j^2$ for every $j \in \{m_1 + 1, \dots, m_2\}$. Note that this is an iterative process, so continuing with this inductive reasoning we get $x_j^1 = x_j^2$ for every $j \in \{2, \dots, m\}$. This completes the proof of the proposition. □

In the following proposition, for a given subset $Y := \{x_{t_1}, \dots, x_{t_s}\} \subseteq V = \{x_1, \dots, x_m\}$ with $t_1 < \dots < t_s$ and $x \in V \setminus Y$, we will write $(Y, x) := (x_{t_1}, \dots, x_{t_s}, x)$ and $(X^k, x^k) := (x_{t_1}^k, \dots, x_{t_s}^k, x^k)$, $k = 1, 2$.

Proposition 5.18. *Fix $s \in \{2, \dots, m-1\}$. Consider a sequence $\{t_\alpha\}_{\alpha=1}^{m-(s+1)}$ and sets $Y_2, \dots, Y_{m-s+1}$ such that $x_{t_\alpha} \in Y_{\alpha+1}$, $\alpha = 1, \dots, m-(s+1)$ and $Y_j \subseteq \{x_2, \dots, x_{s+j-2}\} \setminus \{x_{t_\alpha}\}_{\alpha=1}^{j-2}$ for every $j = 2, \dots, m-s+1$. Consider the cost*

$$c(x_1, \dots, x_m) = c_1(x_1, \dots, x_s) + \sum_{j=2}^{m-s+1} c_j(Y_j, x_{s+j-1}) \tag{5.36}$$

where c_j is semi-concave for each j, and suppose

1. *c_1 is twisted on $\pi_{1,\ldots,s}(S)$ for every c-splitting set S, where $\pi_{1,\ldots,s} : \prod_{i=1}^{m} X_i \mapsto \prod_{i=1}^{s} X_i$ is the canonical projection; that is, for every c-splitting set S and $x_1^0 \in \pi_1(S)$, the map*

$$(x_2, \ldots, x_s) \mapsto D_{x_1} c_1(x_1^0, x_2, \ldots, x_s)$$

   is injective on $\{(x_2, \ldots, x_s) : (x_1^0, x_2, \ldots, x_s) \in \pi_{1,\ldots,s}(S)\}$.

2. *c_j is $(x_{t_{j-1}}, x_{s+j-1})$ twisted for all $j = 2, \ldots, m-s+1$; that is, the map $x_{s+j-1} \mapsto D_{x_{t_{j-1}}} c_j(Y_j, x_{s+j-1})$ is injective on the subset of X_{s+j-1} where $D_{x_{t_{j-1}}} c_j(Y_j, x_{s+j-1})$ exists, for every $j = 2, \ldots, m-s+1$ and Y_j fixed.*

Then, c is twisted on c-splitting sets with respect to the variables $x_1, x_{t_1}, \ldots, x_{t_{m-s}}$.

Proof. Let $S \subseteq X_1 \times \ldots \times X_m$ be a c-splitting set and $(u_1, \ldots, u_m)$ an m-tuple of c-splitting functions for S. Fix $x_1^0 \in \pi_1(S)$ and assume $D_{x_1} c(x_1^0, x_2^1, \ldots, x_m^1) = D_{x_1} c(x_1^0, x_2^2, \ldots, x_m^2)$, where $(x_2^1, \ldots, x_m^1)$, $(x_2^2, \ldots, x_m^2) \in W_{x_1^0, t_1, \ldots, t_{m-s}}$. We want to show that $x_j^1 = x_j^2$ for every $j = 2, \ldots, m$. Indeed, since the costs $c_2, \ldots, c_{m-s+1}$ do not depend on x_1, we immediately get

$$D_{x_1} c_1(x_1^0, x_2^1, \ldots, x_s^1) = D_{x_1} c_1(x_1^0, x_2^2, \ldots, x_s^2).$$

Hence, by Assumption 1 we get

$$x_j^1 = x_j^2 \text{ for } 2 \leq j \leq s. \tag{5.37}$$

To prove that $x_{s+j}^1 = x_{s+j}^2$ for $1 \leq j \leq m-s$ we use induction on j. For $j = 1$, note that $x_{t_1} \in Y_2 \subseteq \{x_2, \ldots, x_s\}$, so $x_{t_1}^1 = x_{t_1}^2$ by (5.37), and by Lemma 2.8

$$D_{x_{t_1}} c(x_1^0, x_2^1, \ldots, x_m^1) = Du_{t_1}(x_{t_1}^1) = Du_{t_1}(x_{t_1}^2) = D_{x_{t_1}} c(x_1^0, x_2^2, \ldots, x_m^2).$$

Since $x_{t_1} \notin Y_j$ for $3 \leq j \leq m-s+1$, we deduce

$$D_{x_{t_1}} c_1(x_1^0, x_2^1, \ldots, x_s^1) + D_{x_{t_1}} c_2(Y_2^1, x_{s+1}^1) = D_{x_{t_1}} c_1(x_1^0, x_2^2, \ldots, x_s^2) + D_{x_{t_1}} c_2(Y_2^2, x_{s+1}^2),$$

then by (5.37),

$$D_{x_{t_1}} c_2(Y_2^1, x_{s+1}^1) = D_{x_{t_1}} c_2(Y_2^2, x_{s+1}^2)$$

and $Y_2^1 = Y_2^2$. Consequently, we must have $x_{s+1}^1 = x_{s+1}^2$, as c_2 is (x_{t_1}, x_{s+1}) twisted on c_2-splitting sets, by Assumption 2.

Assume $x_{s+1}^1 = x_{s+1}^2, \ldots, x_{s+k-1}^1 = x_{s+k-1}^2$, where $1 < k = j \leq m-s$. Combining this and (5.37) we get $x_{t_k}^1 = x_{t_k}^2$, as $x_{t_k} \in Y_{k+1} \subseteq \{x_2, \ldots, x_{s+k-1}\} \setminus \{x_{t_1}, \ldots, x_{t_{k-1}}\}$. Then

$$D_{x_{t_k}} c(x_1^0, x_2^1, \ldots, x_m^1) = Du_{t_k}(x_{t_k}^1) = Du_{t_k}(x_{t_k}^2) = D_{x_{t_k}} c(x_1^0, x_2^2, \ldots, x_m^2).$$

Since $x_{t_k} \notin Y_j$ for $k+2 \leq j \leq m-s+1$, we get

$$D_{x_{t_k}} c_1(x_1^0, x_2^1, \ldots, x_s^1) + \sum_{j=2}^{k+1} D_{x_{t_k}} c_j(Y_j^1, x_{s+j-1}^1) = D_{x_{t_k}} c_1(x_1^0, x_2^2, \ldots, x_s^2) + \sum_{j=2}^{k+1} D_{x_{t_k}} c_j(Y_j^2, x_{s+j-1}^2). \tag{5.38}$$

Now, by induction hypothesis and (5.37), $D_{x_{t_k}} c_1(x_1^0, x_2^1, \ldots, x_s^1) = D_{x_{t_k}} c_1(x_1^0, x_2^2, \ldots, x_s^2)$, $D_{x_{t_k}} c_j(Y_j^1, x_{s+j-1}^1) = D_{x_{t_k}} c_j(Y_j^2, x_{s+j-1}^2)$ for every $j = 2, \ldots, k$, and $Y_{k+1}^1 = Y_{k+1}^2$. Hence, (5.38) reduces to

$$D_{x_{t_k}} c_{k+1}(Y_{k+1}^1, x_{s+k}^1) = D_{x_{t_k}} c_{k+1}(Y_{k+1}^1, x_{s+k}^2).$$

We then conclude $x_{s+k}^1 = x_{s+k}^2$, as c_{k+1} is (x_{t_k}, x_{s+k}) twisted by Assumption 2. This completes the proof of the proposition. □

Bibliography

[1] Agueh, M., Carlier, G. Barycenters in the Wasserstein space. *SIAM J. Math. Anal.*, 43(2):904–924, 2011.

[2] Ambrosio, L., Gigli, N. *A user's guide to optimal transport*, volume In B. Piccolli and M. Rascle, editors, Modelling and Optimisation of Flows on Networks, volume 2062 of Lecture Notes in Mathematics. Springer, 2013.

[3] Arnold, V. Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. *Ann. Inst. Fourier (Grenoble)*, 16:319–361, 1966.

[4] Ballantine, C.S. Products of positive definite matrices i. *Pacific J. Math.*, 23(3):427–433, 1967.

[5] Ballantine, C.S. Products of positive definite matrices ii. *Pacific J. Math.*, 24(1):7–17, 1968.

[6] Benamou, J-D., Carlier, G., Nenna, L. Generalized incompressible flows, multi-marginal transport and Sinkhorn algorithm. *Numerische Mathematik*, 142(1):33–54, 2019.

[7] Bento, J., Mi, L. Multi-marginal optimal transport defines a generalized metric. *Preprint arXiv:2001.11114v3*.

[8] Bernot, M., Figalli, A., Santambrogio, F. Generalized solutions for the euler equations in one and two dimensions. *J. Math. Pures Appl.*, 91(9):137–155, 2009.

[9] Bouchitté, G., Buttazzo, G., Champion, T., De Pascale, L. Dissociating limit in density functional theory with coulomb optimal transport cost. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.*, 2021.

[10] Bouchitté, G., Buttazzo, G., Champion, T., De Pascale, L. Relaxed multi-marginal costs and quantization effects. *Ann. Inst. H. Poincaré Anal. Non Linéaire.*, 38:61–90, 2021.

[11] Brenier, Y. Decomposition polaire et rearrangement monotone des champs de vecteurs. *C.R. Acad. Sci. Pair. Ser. I Math.*, 305:805–808, 1987.

[12] Brenier, Y. The least action principle and the related concept of generalized flows for incompressible perfect fluids. *J. Am. Math. Soc.*, 2:225–255, 1989.

[13] Brenier, Y. Generalized solutions and hydrostatic approximation of the euler equations. *Physica D*, 237(14-17):1982–1988, 2008.

[14] Buttazzo, G., Champion, T., De Pascale, L. Continuity and estimates for multimarginal optimal transportation problems with singular costs. *Appl. Math. Optim.*, 78(1):185–200, 2018.

[15] Buttazzo, G., De Pascale, L., Gori-Giorgi, P. Optimal-transport formulation of electronic density-functional theory. *Physical Review A*, 85(6):062502, 2012.

[16] Carlier, G. On a class of multidimensional optimal transportation problems. *J. Convex Anal.*, 10(2):517–529, 2010.

[17] Carlier, G., and Nazaret, B. Optimal transportation for the determinant. *ESAIM Control Optim. Calc. Var.*, 14(4):678–698, 2008.

[18] Carlier, G., Ekeland, I. Matching for teams. *Econom. Theory.*, 42(2):397–418, 2010.

[19] Chiapporri, P-A., McCann, R., Nesheim, L. Hedonic price equilibria, stable matching and optimal transport; equivalence, topology and uniqueness. *Econom. Theory.*, 42(2):317–354, 2010.

[20] Colombo, M., De Pascale, L., Di Marino, S. Multimarginal optimal transport maps for one-dimensional repulsive costs. *Canad. J. Math.*, 67(2):350–368, 2015.

[21] Cotar, C., Friesecke, G., Klüppelberg, C. Density functional theory and optimal transportation with coulomb cost. *Communications on Pure and Applied Mathematics*, 66(4):548–599, 2013.

[22] Ebin, D.G., Marsden, J. Groups of diffeomorphisms and the motion of an ideal incompressible fluid. *Ann. of Math.*, 92(2):102–163, 1970.

[23] Eckstein, S., Guo, G., Lim, T., Obloj, J. Robust pricing and hedging of options on multiple assets and its numerics. *SIAM J. Finan. Math.*, 12(1):158–188, 2021.

[24] Gallouet, T., Mérigot, Q. A lagrangian scheme for the incompressible Euler equation using optimal transport. *Preprint arXiv:1605.00568*, 2016.

[25] Gangbo, W. Habilitation Thesis. *Universite de Metz*, 1995.

[26] Gangbo, W., McCann, R.J. The geometry of optimal transportation. *Acta Math.*, 177:113–161, 1996.

[27] Gangbo, W., Święch, A. Optimal maps for the multidimensional Monge-Kantorovich problem. *Comm. Pure Appl. Math.*, 51(1):23–45, 1998.

[28] Gerolin, A., Kausamo, A., Rajala, T. Non-existence of optimal transport maps for the multimarginal repulsive harmonic cost. *SIAM J. Math. Anal.*, 51(3):2359–2371, 2019.

[29] Griessler, C. c-cyclical monotonicity as a sufficient criterion for optimality in the multimarginal monge-kantorovich problem. *Proc. Amer. Math. Soc.*, 146:4735–4740, 2018.

[30] Haasler, I., Ringh, A., Chen, Y., Karlsson, J. Multi-marginal optimalt transport with a tree-structured cost and the Schrödinger Bridge Problem. *SIAM J. Control Optim.*, 59(4), 2428–2453.

[31] Haasler, I., Singh, R., Zhang Q., Karlsson, J., Chen, Y. Multi-marginal optimal transport and probabilistic graphical models. *arXiv preprint arXiv:2006.14113*, 2020.

[32] Heinich, H. Probleme de Monge pour n probabilities. *C.R. Math. Acad. Sci. Paris*, 334(9):793–795, 2002.

[33] Kantorovich, L. On the transfer of masses. *Dokl. Acad. Nauk. USSR*, 37:7–8, 1942.

[34] Kellerer, H.G. Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete*, 67:399–432, 1984.

[35] Kim, Y.-H., Pass, B. A general condition for Monge solutions in the multi-marginal optimal transport problem. *SIAM J. Math. Anal*, 46:1538–1550, 2014.

[36] Kim, Y.-H., Pass, B. Multi-marginal optimal transport on Riemannian manifolds. *Amer. J. Math.*, 137:1045–1060, 2015.

[37] McCann, R. Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.*, 11:589–608, 2001.

[38] Mirebeau, J-M. Numerical resolution of Euler equations, through semidiscrete optimal transport. *Journées Équations Aux Dérivées Partielles*, VII:16, 2015.

[39] Moameni, A., Pass, B. Solutions to multi-marginal optimal transport problems concentrated on several graphs. *ESAIM Control Optim. Calc. Var.*, 23:551–567, 2017.

[40] Monge, G. Mémoire sur la théorie des déblais et de remblais. *In Histoire de l'Académie Royale des Sciences de Paris, avec les Mémoires de Mathématique et de Physique, pour la meme annee*, page 666–704, 1781.

[41] Pass, B. PhD thesis. *University of Toronto*, 2011.

[42] Pass, B. On the local structure of optimal measures in the multi-marginal optimal transportation problem. *Calc. Var. and PDE*, 43:529–536, 2012.

[43] Pass, B. Remarks on the semi-classical hohenberg-kohn functional. *Nonlinearity*, 26(9):2731–2744, 2013.

[44] Pass, B. Multi-marginal optimal transport and multi-agent matching problems: Uniqueness and structure of solutions. *Discrete & Continuous Dynamical Systems-A*, 34(4):1623–1639, 2014.

[45] Pass, B. Multi-marginal optimal transport: theory and applications. *ESAIM: Math. Model. Numer. Anal.*, 49:1771–1790, 2015.

[46] Pass, B. Uniqueness and Monge solutions in the multi-marginal optimal transportation problem. *SIAM J. Math. Anal.*, 43:2758–2775, 2015.

[47] Pass, B., Pinamonti, A., Vedovato, M. Multi-marginal optimal transport on the Heisenberg group. *Methods Appl. Anal.*, 28:61–76, 2021.

[48] Pass, B., Vargas-Jiménez, A. Multi-marginal optimal transportation problem for cyclic costs. *SIAM J. Math. Anal.*, 53:4386–4400, 2021.

[49] Pass, B.,Vargas-Jiménez, A. Monge solutions and uniqueness in multi-marginal optimal transport via graph theory. *ArXiv:2104.09488*, 2021.

[50] Rüschendorf, L. On c-optimal random variables. *Statist. Probab. Lett.*, 27:267–270, 1996.

[51] F. Santambrogio. *Optimal transport for Applied Mathematicians: Calculus of Variations, PDEs and Modeling*, volume 87 of Progress in Nonlinear Differentials Equations and Their Applications. Birkhäuser/Springer, Heidelberg, 2015.

[52] C. Villani. *Topics in optimal transportation*, volume 58 of Graduate Studies in Mathematics. American Mathematical Society, Providence, 2003.

[53] C. Villani. *Optimal Transport: old and new*, volume 338 of Grundlehren de mathematischen Wissenschaften. Springer, New York, 2009.

[54] Zemel, Y., Panaretos, V. Frechet means and Procrustes analysis in Wasserstein space. *Bernoulli*, 25:932–976, 2019.