Isometric structure of transportation cost spaces on finite metric spaces

Sofiya Ostrovska and Mikhail I. Ostrovskii*

July 18, 2022

Abstract: The paper is devoted to isometric Banach-space-theoretical structure of transportation cost (TC) spaces on finite metric spaces. The TC spaces are also known as Arens-Eells, Lipschitz-free, or Wasserstein spaces. A new notion of a roadmap pertinent to a transportation problem on a finite metric space has been introduced and used to simplify proofs for the results on representation of TC spaces as quotients of ℓ_1 spaces on the edge set over the cycle space. A Tolstoi-type theorem for roadmaps is proved, and directed subgraphs of the canonical graphs, which are supports of maximal optimal roadmaps, are characterized. Possible obstacles for a TC space on a finite metric space X preventing them from containing subspaces isometric to ℓ_∞^n have been found in terms of the canonical graph of X. The fact that TC spaces on diamond graphs do not contain ℓ_∞^4 isometrically has been derived. In addition, a short overview of known results on the isometric structure of TC spaces on finite metric spaces is presented.

MSC2020: 46B04 (Primary) 46B85 (Secondary)

1 Basic definitions and results

The theory of transportation cost spaces launched by Kantorovich and Gavurin [Kan42, KG49] initially had been developed as a study of special norms pertaining to function spaces on finite metric spaces. However, its further progress turned into the direction of infinite and continuous setting rather than the discrete one in papers of Kantorovich and Rubinstein [Kan42, Kan48, KR57, KR58], and this stream has become dominant. See [ABS21, AGS08, FG21, Gar18, KA82, Vil03, Vil09, Wea18].

Nevertheless, researches within Theoretical Computer Science on the transportation cost, which computer scientists renamed to *earth mover distance* [RTG98], along with some of the recent works in metric geometry and Banach space theory focus on the case of finite metric spaces bringing this area back into spotlight. See, for example, [AFGZ21, BMSZ20+, Cha02, DKO20, DKO21, IT03, KKMR09, KMO20, KN06, Nao21, NR17, NS07, RTG98]. More details on the history of the subject can be found in [Vil09, Chapter 3] and [OO19, Section 1.6].

---

*Corresponding author, `ostrovsm@stjohns.edu`

This work deals with the isometric theory of transportation cost spaces on finite metric spaces. It has to be pointed out that the transportation cost spaces have been studied from different perspectives and found numerous applications in various disciplines. In this connection, there exists a relatively broad assortment of terms for the same notion. We follow the terminology going back to Kantorovich up to adding new concepts and names. In our opinion, this terminology provides the most intuitive description of the area and, as such, makes it more accessible and attractive to both mathematicians and researchers working with practical applications. We presented more arguments in favor of this selection in [OO19, Section 1.6].

Apart from presenting new results (mentioned in the abstract, see Sections 2 and 3 for more details), our goal in this paper is to demonstrate how the convenient terminology and notation allow to simplify the proofs of some already available results. For the convenience of the reader, we present all of the necessary definitions related to the transportation cost spaces in this article, although most of them can be found in [OO19, Section 1.6].

Let (X, d) be a metric space. Consider a real-valued finitely supported function f on X with a zero sum, that is, $\sum_{v \in \mathrm{supp} f} f(v) = 0$. A natural and important interpretation of such a function is considering it as a *transportation problem*: one needs to transport certain product from locations where $f(v) > 0$ to locations where $f(v) < 0$.

One can easily see that a transportation problem f can be represented as

$$f = a_1(\mathbf{1}_{x_1} - \mathbf{1}_{y_1}) + a_2(\mathbf{1}_{x_2} - \mathbf{1}_{y_2}) + \cdots + a_n(\mathbf{1}_{x_n} - \mathbf{1}_{y_n}), \tag{1}$$

where $a_i \geq 0$, $x_i, y_i \in X$, and $\mathbf{1}_u(x)$ for $u \in X$ is the *indicator function* of u. We call each such representation a *transportation plan* for f, it can be interpreted as a plan of moving a_i units of the product from x_i to y_i. A pair x_i, y_i for which $a_i > 0$ will be called a *transportation pair*. In some contexts we shorten "transportation plan" to *plan* if it does not cause any confusion.

The *cost* of the transportation plan (1) is defined as $\sum_{i=1}^{n} a_i d(x_i, y_i)$.

We denote the real vector space of all transportation problems by $\mathrm{TP}(X)$. We introduce the *transportation cost norm* (or just *transportation cost*) $\|f\|_{\mathrm{TC}}$ of a transportation problem f as the infimum of costs of transportation plans satisfying (1).

The completion of $\mathrm{TP}(X)$ with respect to $\|\cdot\|_{\mathrm{TC}}$ is called the *transportation cost space* $\mathrm{TC}(X)$. Note that though for finite metric spaces X, the spaces $\mathrm{TC}(X)$ and $\mathrm{TP}(X)$ coincide as sets, we mostly use the notation $\mathrm{TC}(X)$ to emphasize that we consider it as a normed space.

It has to be pointed out that, in our discussion, transportation plans are allowed to be *fake plans*, in the sense that it can happen that there is no product in x_i in order to make the delivery to y_i. Such fake plans will be used in the proof of Theorem 1.9.

We start with the next statement which is an easy consequence of the triangle inequality.

Proposition 1.1. *The infimum of costs of transportation plans for a finitely supported* $f \in \mathrm{TC}(X)$ *is attained, and this happens for some transportation plan with* $\{x_i\} = \{v : f(v) > 0\}$ *and* $\{y_i\} = \{v : f(v) < 0\}$.

A transportation plan for f of the minimal cost, that is, whose cost equals $\|f\|_{\mathrm{TC}}$ is called an *optimal transportation plan* for $f \in \mathrm{TP}(X)$.

In the sequel, a finite metric space X will be identified with an apposite weighted graph. This weighted graph will be specified uniquely according to the procedure described below and - following [AFGZ21] - will be called the *canonical graph associated with the metric space*. The graph is defined as follows. First, we consider a complete weighted graph with vertices in X, and the weight of an edge uv $(u, v \in X)$ defined as $d(u, v)$. After that we delete all edges uv for which there exists a vertex $w \notin \{u, v\}$ satisfying $d(u, w) + d(w, v) = d(u, v)$. It is easy to see that, for a finite metric space, it is a well-defined procedure leading to a uniquely determined by X weighted graph, which we denote $G(X, E)$, where $E = E(X)$ is the edge set of the canonical graph.

As a result, the constructed in this way graph $G(X, E)$ possesses the following feature, which is crucial for its application in the theory of the transportation cost spaces: the weighted graph distance of $G(X, E)$ on X coincides with the original metric on X. Note that if X is defined initially as a weighted graph with its weighted graph distance, then the corresponding canonical graph can be different because some edges can be dropped. However, if we start with a simple unweighted graph endowed with the graph distance, we will recover it as a canonical graph.

For a metric space X, let $\ell_{1,d}(E) = \ell_{1,d}(E(X))$ be the weighted ℓ_1-space on the edge set of the canonical graph $G(X, E)$ associated with X. The weight of an edge uv is $d(u, v)$. The space $\ell_{1,d}(E)$ consists of real-valued functions $\beta : E \to \mathbb{R}$ with the norm

$$\|\beta\|_{1,d} := \sum_{uv \in E(G)} |\beta_{uv}| d(u, v). \tag{2}$$

It is easy to see that if X is an unweighted graph with its graph distance, then $\ell_{1,d}(E)$ is just the vector space $\mathbb{R}^E$ of functions on the edge set of X with its ℓ_1-norm.

We introduce the space $\ell_{1,d}(E)$ because, for a weighted graph, it is natural and practical to consider the following, more detailed version of transportation plans.

In order to proceed, an orientation on the edge set E has to be specified. This means that we fix a *direction* on the edges by selecting the *head* and *tail* for each edge. The orientation can be chosen arbitrarily as the notions important for further reasoning do not depend on the choice of this orientation. We call this orientation a *reference orientation*.

Notation: if u is a tail and v is a head of a directed edge $e = uv$, we denote the directed edge either $\overrightarrow{uv}$ or $\overleftarrow{vu}$. We also write $u = \mathrm{tail}(e)$ and $v = \mathrm{head}(e)$.

Fix a reference orientation on $G(X, E)$ and consider $p \in \ell_{1,d}(E)$. We call p a *roadmap* for the reasons which will be explained below. For each such p, we

introduce the function

$$\text{problem}_p(v) = \sum_{\text{tail}(e)=v} p(e) - \sum_{\text{head}(e)=v} p(e) \in \text{TP}(X). \tag{3}$$

Note that the function is in $\text{TP}(X)$ because each edge is included in the right-hand side of (3) exactly twice - once for its head and once for its tail.

It can be readily seen that p provides an inherent transportation plan associated with problem_p, namely, the plan

$$\text{problem}_p = \sum_{e \in \text{supp} p} p(e) \left(\mathbf{1}_{\text{tail}(e)} - \mathbf{1}_{\text{head}(e)} \right), \tag{4}$$

Remark 1.2. It should be emphasized that the cost of this transportation plan is exactly $\|p\|_{1,d}$. This norm is also called the *cost of roadmap p*, because we can identify p and transportation plan (4) by virtue of the canonical correspondence between those objects.

Remark 1.3. The function $\text{problem}_p(v)$ is a negation of the well-known in combinatorial optimization (see [Sch03, p. 149]) *excess function*: $\text{excess}_p = - \text{problem}_p$.

We use the term *roadmap* because to organize the transportation (we use intuitive rather than mathematical language) according to a transportation plan, one has to identify edges used for transportation and to specify the direction on them.

Namely, we do the following for the transportation plan given by the right-hand side of (1), which we denote by P.

(1) If $x_i y_i$ is not an edge in E, we introduce a shortest path $u_{i,0} = x_i, u_{i,1}, \ldots, u_{i,m(i)} = y_i$ joining x_i and y_i in G, and replace $a_i(\mathbf{1}_{x_i} - \mathbf{1}_{y_i})$ in plan P by

$$\sum_{j=1}^{m(i)} a_i(\mathbf{1}_{u_{i,j-1}} - \mathbf{1}_{u_{i,j}}).$$

Observe that this step is uniquely determined if and only if for each transportation pair x_i, y_i of plan P there is one shortest path between x_i and y_i in $G(X, E)$.

As a result, we get a plan in which every transportation pair is the pair of ends of an edge, and so it is almost of the form shown in (4). We use the word "almost" because some pairs of ends of an edge can repeat in the sum.

(2) We combine all terms corresponding to the same edge in the obtained transportation plan. We get a transportation plan for the same problem whose cost does not exceed the original plan's cost (there can be some cancellations that will decrease the cost) and for which all transportation pairs are edges.

(3) If all transportation pairs are edges, we map the transportation plan

$$a_1(\mathbf{1}_{x_1} - \mathbf{1}_{y_1}) + a_2(\mathbf{1}_{x_2} - \mathbf{1}_{y_2}) + \cdots + a_n(\mathbf{1}_{x_n} - \mathbf{1}_{y_n})$$

onto the roadmap which on the edge $x_i y_i$ takes value a_i if the reference orientation of $x_i y_i$ is $\overrightarrow{x_i y_i}$ and value $-a_i$ if the reference orientation of $x_i y_i$ is $\overleftarrow{x_i y_i}$. For edges uv which are not transportation pairs the value of the roadmap is 0.

Remark 1.4 (Conclusion of the above construction). For each transportation plan P for $f \in \mathrm{TP}(X)$, there is a naturally defined (but not uniquely determined) roadmap p which can be regarded as an implementation of P and whose cost does not exceed the cost of P.

These definitions and remarks lead to what we consider very transparent and straightforward proof of the following proposition.

Proposition 1.5. *The quotient representation* $\mathrm{TC}(X) = \ell_{1,d}(E)/Z$, *where* Z *is the cycle space in* $\ell_{1,d}(E)$, *holds.*

As far as we know, Proposition 1.5 (in the case of unweighted graphs) was for the first time proved [Ost13, Proposition 10.10]. Later, different versions of the proof were published in [OO20], [AFGZ21], [DKO21]. Our version of the proof can be regarded as a generalization of the proof in [DKO21] for unweighted graphs to the case of metric spaces.

Prior to presenting this version, let us remind the notion of the cycle space and other related concepts of the algebraic graph theory, used in the subsequent parts of this paper.

The following definition, as a rule, is used for an arbitrary oriented finite graph $G = (V, E)$, not necessarily connected. We are going to use it for canonical graphs $G = (X, E)$ with reference orientation.

Denote by $\mathbb{R}^E$ and $\mathbb{R}^V$ the spaces of real-valued functions on the edge set and the vertex set, respectively.

Definition 1.6. The *incidence matrix* D of an oriented graph G is defined as a matrix whose rows are labelled using vertices of G, columns are labelled using edges of G, and the ve-entry is given by

$$d_{ve} = \begin{cases} 1, & \text{if } v = \mathrm{head}(e), \\ -1, & \text{if } v = \mathrm{tail}(e), \\ 0, & \text{if } v \text{ is not incident to } e. \end{cases}$$

Interpreting elements of $\mathbb{R}^E$ and $\mathbb{R}^V$ as column vectors, we may regard D as a matrix of a linear operator $D : \mathbb{R}^E \to \mathbb{R}^V$. We also consider the transpose matrix D^T and the corresponding operator $D^T : \mathbb{R}^V \to \mathbb{R}^E$.

It is easy to describe $\ker D^T$. In fact, for $f \in \mathbb{R}^V$ the value of $D^T(f) \in \mathbb{R}^E$ at an edge e is $f(\mathrm{head}(e)) - f(\mathrm{tail}(e))$, therefore $f \in \ker D^T$ if and only if it has the same value at the ends of each edge. It is clear that this happens if and only if f is constant on each of the connected components of G. Therefore the ranks of the operators D^T and D are equal to $|V| - c$, where c is the number of connected components of G.

Observe that
$$(Dp)(v) = - \operatorname{problem}_p(v) \tag{5}$$
for $p \in \mathbb{R}^E$. We let
$$Z = \ker D. \tag{6}$$
This subspace of $\mathbb{R}^E$ is called the *cycle space* or *cycle subspace*. The name was chosen because, as we explain below, the set of signed indicator functions of all cycles spans this subspace.

Now we consider a cycle C in G (in a graph-theoretical sense). We consider one of the two possible orientations of C satisfying the following condition: each vertex of C is a head of exactly one edge and a tail of exactly one edge. Now we introduce the *signed indicator function* $\chi_C \in \mathbb{R}^E$ of the oriented cycle C (in an oriented graph G) by

$$\chi_C(e) = \begin{cases} 1 & \text{if } e \in C \text{ and its orientations in } C \text{ and } G \text{ are the same} \\ -1 & \text{if } e \in C \text{ but its orientations in } C \text{ and } G \text{ are different} \\ 0 & \text{if } e \notin C. \end{cases} \tag{7}$$

It is immediate from (4) and (5) that $\chi_C \in \ker D$. Let us show that signed indicator functions of cycles span $\ker D$. To see this we observe that $\dim(\ker D) = |E(G)| - \operatorname{rank}(D) = |E(G)| - |V(G)| + c$. So it remains to show that there are $|E(G)| - |V(G)| + c$ cycles $\{C(i)\}_i$ such that the functions $\{\chi_{C(i)}\}_i$ are linearly independent.

To achieve this goal, consider a maximal subset of edges of G which does not form any cycles. Such collection of edges of G induces a subgraph H of G which is called a *spanning forest of G*; in the case where G is connected, it is called a *spanning tree of G*. Notice that we use these terms in a slightly different way than it can be found in the literature. This discrepancy stems from different possible treatments of the graph's isolated vertices. Since we do not need to consider graphs with isolated vertices, this issue is of no importance here. It is a well-known fact that the number of edges in a spanning forest is $|V(G)| - c$. Therefore, the number of the remaining edges is $|E(G)| - |V(G)| + c$. Now we construct the cycles $C(i)$, $i \in E(G) \backslash E(H)$, with linearly independent $\{\chi_{C(i)}\}$ as follows. Each $i \in E(G) \backslash E(H)$ forms a cycle together with some edges of H. One can show that such a cycle is uniquely determined, but at this point, it does not matter. We pick one such cycle and denote it $C(i)$. The fact that $\{\chi_{C(i)}\}$ are linearly independent is immediate because the value of the linear combination $\sum_i a_i \chi_{C(i)}$ at an edge i is equal to $\pm a_i$.

Proof of Proposition 1.5. Our idea of the proof is very straightforward: to apply the made above observation that using the reference orientation, any element p of $\ell_{1,d}(E)$ can be regarded as a roadmap for some transportation problem, namely, for $f = \operatorname{problem}_p$. In such a way, one gets a natural linear map $Q : \ell_{1,d}(E) \to \mathrm{TC}(X)$.

Remark 1.4 shows that this map is surjective and that it has norm ≤ 1. Finally, the fact that, for each $f \in \mathrm{TC}(X)$, there is an optimal transportation plan P with

cost $\|f\|_{\mathrm{TC}}$, implies that the roadmap p assigned to P - see Remark 1.4 - satisfies $\|p\|_{1,d} = \|f\|_{\mathrm{TC}}$ (the inequality $\|p\|_{1,d} < \|f\|_{\mathrm{TC}}$ would contradict the definition of an optimal transportation plan). Thus $Q : \ell_{1,d}(E) \to (\mathrm{TC}(X), \|\cdot\|_{\mathrm{TC}})$ is a quotient map. Finally, combining (5) with the definition (6) of Z, we conclude that $Z = \ker Q$. $\qquad\square$

Remark 1.7. There exist analogues of Proposition 1.5 for some infinite metric spaces, see [CKK17, GL18].

Proposition 1.5 implies that one can improve a non-optimal roadmap for a given transportation problem by adding an element of the cycle space. Our next target is to prove a ramification of this fact, stating that the roadmap can be improved by adding a multiple of a cycle. This is exactly the assertion of Theorem 1.9 in Section 1.2.

The idea of this result goes back to Tolstoi [Tol30, Tol39], see [Sch03, Theorem 12.1, Theorem 12.3, Theorem 21.12, and pp. 362–366] for interesting related information. Tolstoi [Tol30, Tol39] was the first one who proposed describing the optimality conditions for plans via nonexistence of "negative" cycles. The listed above theorems in [Sch03] describe such conditions of optimality in different settings.

1.1 An overview of the isometric theory of $\mathrm{TC}(X)$ for finite X

Since the isometric theory of transportation cost spaces on finite metric spaces is quite recent, it is beneficial not only for the presentation of this study but also for further developments in this area to precede the exposition of new findings with a short survey of the available ones.

(1) The space $\mathrm{TC}(X)$ is isometric to ℓ_1^n if and only if X is isometric to a weighted tree on $n + 1$ vertices. Godard [God10] proved the "if" part. In the "only if" direction, the paper [God10] contains only a partial result that X embeds into a tree. The "only if" result can be easily derived from the description of extreme points in the unit ball of $\mathrm{TC}(X)$ as vectors of the form $(\mathbf{1}_u - \mathbf{1}_v)/d(u,v)$, where u and v such that all triangle inequalities of the form

$$d(u,z) + d(z,v) \geq d(u,v)$$

are strict for $z \notin \{u,v\}$. The result is obtained by comparing the number of extreme points in ℓ_1^{n+1} and $\mathrm{TC}(X)$, see [DKO21, Proposition 26] for details. It should be remarked here that the study of extreme points in $\mathrm{TC}(X)$ was apparently initiated by Weaver in the first edition of his book [Wea99], and the mentioned above elementary description of extreme points in the unit ball of $\mathrm{TC}(X)$ for finite X is a part of folklore in this study.

Extreme points in the unit ball of the dual space $(\mathrm{TC}(X))^* = \mathrm{Lip}_0(X)$ (defined at the beginning of Section 2) were characterized in [Far94]. See [Wea18, Corollary 2.60] and [MP21, Theorem 2] for a more explicit description in the case of finite X.

(2) Isometric description of $\mathrm{TC}(X)$ as a quotient of $\ell_1(E)$ over the cycle space was obtained in [Ost13] in the case when X is an unweighted graph. This result was generalized to the case of weighted graphs in [OO20]. Alternative proofs of it were obtained in [AFGZ21], [DKO21], and this paper.

(3) The study of isometric embeddability of ℓ_1 into $\mathrm{TC}(X)$ for infinite X was initiated by Cúth and Johanis [CJ17]. However, the following seems to be the first isometric result in the case of finite X:

> **Theorem 1.8** (Khan, Mim, and Ostrovskii [KMO20]). *If a metric space M contains $2n$ elements, then $\mathrm{TC}(M)$ contains a 1-complemented subspace isometric to ℓ_1^n. If the space M is such that triangle inequalities for all distinct triples in M are strict, then $\mathrm{TC}(M)$ does not contain a subspace isometric to ℓ_1^{n+1}.*

(4) The paper [KMO20] also contains examples of finite metric spaces X_3 and X_4 such that $\mathrm{TC}(X_3)$ contains ℓ_∞^3 isometrically and $\mathrm{TC}(X_4)$ contains ℓ_∞^4 isometrically. The examples were simplified in [DKO21] where it was shown that one can take $X_3 = K_{2,4}$ (complete bipartite graph) and $X_4 = K_{4,4}$. Finally, in [AFGZ21] it was observed that $\mathrm{TC}(C_4) = \ell_\infty^3$.

Note that the problem on the existence of isometric copies of ℓ_∞^n in $\mathrm{TC}(X)$ can be regarded as an isometric version of the famous Bourgain's problem (see [Bou86] and the discussion in [Nao18, Section 1.2.2]) on the cotype of $\mathrm{TC}(\mathbb{R}^2)$.

(5) Alexander, Fradelizi, García-Lirola, and Zvavitch [AFGZ21] characterized isometries of $\mathrm{TC}(X)$ in terms of the canonical graph $G(X)$. They showed that isometries correspond to cycle-preserving bijections of edge sets. This allowed them to use the results of Whitney [Whi33] (see also [Oxl11, Section 5.3]) on classifications of cycle-preserving bijections of edge sets. In particular, they proved that, for a 3-connected graph, isometries correspond to homotheties of the graph - that is, the multiplication of all weights of edges by the same positive number. Earlier, a similar result was obtained for Sobolev spaces on graphs, see [Ost07].

(6) The result of [AFGZ21] on ℓ_1-decompositions: Decompositions of $\mathrm{TC}(X)$ of the form $\mathrm{TC}(X) = Z_1 \oplus_1 Z_2$ correspond to representability of the canonical graph X as a union of two graphs X_1 and X_2 with one common vertex, such that $Z_1 = \mathrm{TC}(X_1)$ and $Z_2 = \mathrm{TC}(X_2)$.

(7) The result of [AFGZ21] on ℓ_∞-decompositions: Decompositions of $\mathrm{TC}(X)$ of the form $\mathrm{TC}(X) = Z_1 \oplus_\infty Z_2$ imply that one of the summands of Z_1 and Z_2 is one-dimensional, while the other is ℓ_1^{n-1} and $X = K_{2,n}$ (unweighted complete bipartite graph).

(8) Findings of [AFGZ21] reveal that transportation cost spaces are significant examples for the studies related to the Mahler conjecture on the volume product of a symmetric convex body and its polar.

(9) Vershik [Ver15] suggested to study - within the framework of the theory of convex polytopes - combinatorial properties pertaining to the unit ball of $\mathrm{TC}(X)$ for a finite metric space (X, d). There, Vershik named the unit ball of $\mathrm{TC}(X)$ the *fundamental polytope of* (X, d) and used the term *Kantorovich-Rubinstein norm* for the transportation cost, because the continuous analogue of this norm was studied in [KR57, KR58]. Later, some authors started to use the term *Kantorovich-Rubinstein polytope* instead of *fundamental polytope*. In the case of the unweighted graphs, this study [DDM19] was related to the study of symmetric edge polytopes. Further developments in this area have been accomplished by several authors, see [DDM19, GP17, HJM19, JTZ21] and references therein.

(10) Montrucchio and Pistone [MP21] provided a differently-focused introductory survey on properties of $\mathrm{TC}(X)$ for finite metric spaces X.

1.2 A new Tolstoi-type theorem

The setting of our Tolstoi-type theorem is different from the ones in [Sch03] and the proof method for the "only if" direction is different as well. For this reason, we decided that its publication will be useful for further studies of transportation cost spaces. Meanwhile, the proof of the "if" direction is easily established in all cases.

For a roadmap p, we introduce the *induced by p direction (orientation) of edges which are in* $\mathrm{supp}\,p$ as follows: If $p(e) > 0$, the orientation of the edge coincides with the reference orientation; if $p(e) < 0$ - it is the opposite to the reference orientation.

Here comes our Tolstoi-type theorem.

Theorem 1.9. *Let p be a roadmap for some transportation problem f on a weighted graph $G = (X, E)$. The roadmap p is not optimal if and only if the graph G contains a directed cycle such that the total weight of the edges in it which are in $\mathrm{supp}\,p$ and whose induced by p direction is opposite to their direction in the cycle exceeds the total weight of other edges in the directed cycle.*

Proof of Theorem 1.9. The "if" direction. Assume that there is such an oriented cycle c. We identify it with χ_c defined in (7) and consider it as an element of $\ell_{1,d}(E)$. The assumption on c implies that, for sufficiently small $\alpha > 0$, one has $\|p + \alpha c\|_{1,d} < \|p\|_{1,d}$. On the other hand, since c is an oriented cycle, it is clear that $\mathrm{problem}_{p+\alpha c} = \mathrm{problem}_p = f$. Thus, the roadmap p is not optimal.

To prove the "only if" part of Theorem 1.9, assume that there are cheaper than p roadmaps for f. The set of such roadmaps includes optimal roadmaps for f. By compactness, among the optimal roadmaps, there is a roadmap $\tilde{p}$ which minimizes $\|\tilde{p} - p\|_{1,d}$. If such $\tilde{p}$ is not unique, we pick one of them.

As $d := \tilde{p} - p$ is a difference of two roadmaps for the same transportation problem, it is a roadmap for the null problem, that is, the transportation problem in which nothing is available and nothing is needed. Using the orientation induced by d, we create an oriented graph (X, E_d) with the same vertex set as G and with edge set being $\mathrm{supp}\,d$.

Since problem$_d = 0$, it is clear that each vertex with nontrivial indegree has a nontrivial outdegree in (X, E_d). Hence, this oriented graph has an oriented cycle c. We are going to prove that this cycle satisfies the condition in Theorem 1.9.

Assume the contrary. Then the following condition for c holds:

(*) The total weight of edges of c, whose directions are opposite to the induced by p does not exceed the total weight of the remaining edges in c.

We identify c with χ_c and consider it as a roadmap. The choice of c implies that for each $\delta \in (0, \min_{e \in c} d(e))$, there holds $\|d - \delta c\|_{1,d} < \|d\|_{1,d}$, or, equivalently,

$$\|(\tilde{p} - \delta c) - p\|_{1,d} < \|\tilde{p} - p\|_{1,d}. \tag{8}$$

Selecting arbitrary $\delta \in (0, \min_{e \in c} d(e))$, set $\hat{p} = \hat{p}_\delta := \tilde{p} - \delta c$. Now, we refine the choice of δ by showing that there exists $\delta \in (0, \min_{e \in c} d(e))$ for which $\|\hat{p}\|_{TC} \le \|\tilde{p}\|_{TC}$. Indeed, the definition of c implies that the signs of c and $\tilde{p}$, considered as vectors of $\ell_{1,d}(E)$, can differ only on edges where the sign of c coincides with the sign of $-p$. By condition (*), the total length of all such edges in c does not exceed half of length of c. Bearing this in mind, opt for $\delta \in (0, \min_{e \in c} d(e))$ which does not exceed the minimum absolute value of $\tilde{p}(e)$ over all of those edges e on which the signs of $\tilde{p}$ and c coincide. Then, for this δ, one has: $\|\hat{p}\|_{1,d} = \|\tilde{p} - \delta c\|_{1,d} \le \|\tilde{p}\|_{1,d}$.

Furthermore, for any such δ, roadmap $\hat{p}$ is a roadmap for the original transportation problem as we subtracted a roadmap for a null problem, and, by virtue of (8), one has: $\|\hat{p} - p\|_{1,d} < \|\tilde{p} - p\|_{1,d}$.

Thus, $\hat{p}$ is also an optimal roadmap for f, and it satisfies $\|\hat{p} - p\|_{1,d} < \|\tilde{p} - p\|_{1,d}$, contrary to our choice of $\tilde{p}$. $\qquad \square$

2 Directed graphs related to transportation problems and Lipschitz functions

The dual space of $\mathrm{TC}(X)$ is the space $\mathrm{Lip}_0(X)$ of Lipschitz functions on X which vanish at a base point (an arbitrarily chosen point in X), see [Ost13, Section 10.2]. Let $f \in \mathrm{TC}(X)$ and $l \in \mathrm{Lip}_0(X)$ be such that $\mathrm{Lip}(l) = 1$ and $l(f) = \|f\|_{TC}$, where by $l(f)$ we mean the result of the action of a functional l on a vector f. If $f = a_1(\mathbf{1}_{x_1} - \mathbf{1}_{y_1}) + a_2(\mathbf{1}_{x_2} - \mathbf{1}_{y_2}) + \cdots + a_n(\mathbf{1}_{x_n} - \mathbf{1}_{y_n}), \quad a_i > 0$, is an optimal transportation plan for f, then

$$\sum_{i=1}^{n} a_i(l(x_i) - l(y_i)) = \|f\|_{TC} = \sum_{i=1}^{n} a_i d(x_i, y_i),$$

whence

$$l(x_i) - l(y_i) = d(x_i, y_i) \tag{9}$$

for each transportation pair $x_i y_i$.

A transportation plan satisfying condition (9) for some 1-Lipschitz function l is called *potential* with *potential l*.

Observation 2.1 (Kantorovich and Gavurin [KG49]). *A transportation plan is optimal if and only if it is potential.*

According to the preceding discussion, we consider optimal (= minimal cost) roadmaps for $f \in \mathrm{TC}(X)$ as elements of $\ell_{1,d}(E)$ and define the *support* of a roadmap as the support of the corresponding element of $\ell_{1,d}(E)$. Let us show that there always exists an optimal roadmap p for f with the largest possible support among the roadmaps for f, whence the largest possible support of a roadmap is uniquely determined by f. In fact, any two optimal for f roadmaps, p_1 and p_2, cannot be represented by elements of $\ell_{1,d}(E)$ having different signs on any of the edges, because otherwise, the element $\frac{1}{2}(p_1 + p_2) \in \ell_{1,d}(E)$ would be also a roadmap for f with a strictly smaller cost. Thus, considering in $\ell_{1,d}(E)$ an average of a collection of optimal roadmaps covering all edges belonging to the support of at least one of the optimal roadmaps, one obtains an optimal roadmap with the maximal possible support.

For an optimal roadmap p with the maximal possible support, consider the following directed graph, which we call the *directed graph of f*. It contains only those edges of E that are in the support of p. Further, the edges with positive value of p have the reference orientation, while, for edges where the value of p is negative, the orientation is the opposite-to-reference one. To put differently, the directions of all edges coincide with the directions of transportation according to the roadmap p.

Our next purpose is to provide a convenient description for the obtained class of directed subgraphs of (X, E). This will be done in terms of the definition below.

Definition 2.2. Let l be a 1-Lipschitz function on $G = (X, E)$. Consider the set of edges $uv \in E$ for which $|l(u) - l(v)| = d(u, v)$, such an edge is called *supported by l*. A *downhill graph* for l is the directed graph consisting of all supported by l edges directed from the vertex where l is larger to the vertex where l is smaller. Each such directed edge is called a *downhill edge for l*.

By Observation 2.1 and the remark on roadmaps with the maximal possible support preceding Definition 2.2, the directed graph of a transportation problem is a subgraph of a downhill graph for some 1-Lipschitz function.

Theorem 2.3. *A directed subgraph of (X, E) with at least one edge is a downhill graph for a 1-Lipschitz function if and only if it is the directed graph for some non-zero transportation problem.*

Prior to proving the Theorem, some auxiliary results will be presented. The following definition comes in handy.

Definition 2.4. Given $0 \neq f \in \mathrm{TP}(X)$, a *supporting function s* for f is a function $s \in \mathrm{Lip}_0(X)$ satisfying $\mathrm{Lip}(s) = 1$ and $s(f) = \|f\|_{\mathrm{TC}}$.

Denote by T_f the subset of $E(X)$ consisting of all edges of $E(X)$ which are contained in supports of some optimal roadmaps for f. Trivially, (X, T_f) is a *spanning subgraph* of $(X, E(X))$, that is, a subgraph containing all vertices of the graph.

A necessary and sufficient condition for the uniqueness of a supporting function is given by the forthcoming statement.

Theorem 2.5. *A supporting function for f is unique if and only if the graph (X, T_f) is connected. If the graph (X, T_f) is disconnected, then a supporting function for f is uniquely determined on the connected component of (X, T_f) containing O, and is determined up to additive constants on other connected components of (X, T_f).*

Remark 2.6. The referee turned our attention to the fact that Theorem 2.5 can be derived from [AR20, Theorem 3.5] by means of the well-known fact (see [DGZ93, Corollary 1.5]) that, for every Banach space, the set of points of Gâteaux differentiability coincides with that possessing a unique supporting functional. Our proof of Theorem 2.5 is different and more direct.

Also, the referee drew our attention to the closely related result [GP17, Theorem 3].

Proof. Let p be an optimal roadmap for f whose support is T_f:

$$p = \sum_{j=1}^{m} a_j \mathbf{1}_{\overrightarrow{x_j y_j}}, \quad a_j > 0,$$

and $T_f = \{x_j y_j\}_{j=1}^{m}$. Then,

$$s(f) = \sum_{j=1}^{m} a_j (s(x_j) - s(y_j)).$$

On the other hand, the above sum is equal to $\|f\|_{\mathrm{TC}} = \sum_{j=1}^{m} a_j d(x_j, y_j)$. Since $\mathrm{Lip}(s) = 1$, this occurs if and only if $s(x_j) - s(y_j) = d(x_j, y_j)$ for each $x_j y_j \in T_f$ yielding that s is determined up to an additive constant on each of the components of (X, T_f). Taking into account that $s(O) = 0$, one concludes that s is uniquely determined on the component of (X, T_f) containing O.

Our proof of the "only if" part of Theorem 2.5 is based on the following two lemmas.

Lemma 2.7. *Downhill edges for s, whose both ends belong to the same component of T_f, are edges of T_f.*

Proof. Assume that uv is an edge with $s(u) - s(v) = d(u, v)$ satisfying the conditions of the lemma. Let $u = u_0, u_1, \ldots, u_k = v$ be a path in T_f joining u and v. Then, for each $i = 1, \ldots, k$, one has: $s(u_i) - s(u_{i-1}) = \pm d(u_i, u_{i-1})$. Now, consider the cycle $u_0, u_1, \ldots, u_k, u_0$ together with a maximal-support roadmap p for f. By the assumptions above, the support of p contains all edges of this cycle except $u_k u_0$. Denote by ε ($\varepsilon > 0$) the minimal amount of product transported according to the roadmap p along the edges $u_0 u_1, u_1 u_2, \ldots, u_{k-1} u_k$ in the direction of decrease of s.

Next, we create, for the same problem, a new roadmap $\tilde{p}$ from p as follows. In the roadmap $\tilde{p}$, we move ε units of product along the edge $u_0 u_k$ in the direction of

decrease of s. To modify p only for edges in the cycle $u_0, u_1, \ldots, u_k, u_0$, we proceed as follows: selecting the direction on the cycle which on $u_0 u_k$ coincides with the direction of decrease of s, we follow this orientation. From here on, for each edge on which the direction of the cycle and of the decrease of s coincide, we increase the amount of transported product by ε, while on each edge for which the direction of decrease of s and the direction of the cycle are opposite, the amount of transportation will be decreased by ε. It is easy to see that

$$\text{problem}_p = \text{problem}_{\tilde{p}} \quad \text{and} \quad \|p\|_{1,d} = \|\tilde{p}\|_{1,d}.$$

Therefore $\tilde{p}$ is a roadmap for the same problem and uv is in its support. $\qquad \square$

Lemma 2.8. *For each edge $e = uv$ joining different components of (X, T_f), there exists a supporting function s_e for f such that $|s_e(u) - s_e(v)| \neq d(u, v)$.*

Proof. Only edges which are downhill for s have to be considered. Let $X_1, \ldots, X_k$ be components of (X, T_f). If there is $u \in X_i$ and $v \in X_j$, $j \neq i$, such that $s(v) - s(u) = d(u, v)$, we write $X_i \prec X_j$.

We start by proving that there are no cycles in this ordering of components, that is, there are no finite subsets

$$\{n(1), \ldots, n(m)\} \subset \{1, \ldots, k\}, \quad \text{where} \quad m \geq 2,$$

such that

$$X_{n(1)} \prec X_{n(2)} \prec \cdots \prec X_{n(m)} \prec X_{n(1)}. \tag{10}$$

Assume the contrary, that is, that there exists such a cycle. Pick vertices $u_i \in X_{n(i)}$ and $v_i \in X_{n(i+1)}$ - bearing in mind that $v_m \in X_{n(1)}$ - in such a way that $s(v_i) - s(u_i) = d(u_i, v_i)$.

Denote by W_i a path in T_f joining v_{i-1} and u_i, where, by the agreement, W_1 is a path in T_f joining v_m and u_1. Then,

$$u_1 v_1 W_2 u_2 v_2 W_3 u_3 \ldots v_{m-1} W_m u_m v_m W_1 u_1 \tag{11}$$

is a cycle in $G(X, E)$. By our assumption, the edges $u_i v_i$ are not in T_f, and our goal is to show that this assumption leads to a contradiction.

To achieve this goal, it suffices to show that our assumptions imply the existence of an optimal roadmap for f, whose support contains edges $u_i v_i$. By the definition of T_f, there exists an optimal roadmap p for f, whose support is T_f. Let $\varepsilon > 0$ be the minimal value of p on edges of T_f. We add to p a roadmap $\tilde{p}$ which moves ε units along each edge of the cycle (11) in the direction of edges $\overrightarrow{v_i u_i}$. It should be emphasized that $\tilde{p}$ just follows the direction of the cycle and ignores all previous orientations on the edges involved. The important immediate consequence of this is the fact that the roadmap $\hat{p} = p + \tilde{p}$ is also a roadmap for f.

It remains only to verify that $\tilde{p}$ has the same cost as p. Our assumption on ε guarantees that the values of $\hat{p}$ and p do not have opposite signs on any of the edges. In addition, according to our definitions, the transportation in the roadmap p goes in the direction of decrease of s for each edge.

13

The roadmap $\tilde{p}$ has been introduced in such a way that its addition to p decreases the transported amount by ε for all edges for which the direction of the cycle (11) is the direction of increase of s, and increases the transported amount by ε for all edges for which the direction of the cycle (11) is the direction of decrease of s. At this point, recall that $d(u,v) = |s(u) - s(v)|$ for each edge in the cycle (11). Thus, in the directed cycle (11), the total length of the edges along which s decreases if the same as the total length of the edges along which s increases. This proves that the cost of p is the same as of $\tilde{p}$. So there is an optimal roadmap for f whose support contains edges $\{u_i, v_i\}_{i=1}^{\infty}$, This contradiction leads to the conclusion that there are no cycles in the ordering of components.

Now, let $\overrightarrow{uv}$ be a downhill edge for s, for which we are going to prove Lemma 2.8, and let X_1 and X_2 be the components of (X, T_f) joined by uv; $u \in X_1$ and $v \in X_2$.

To construct the desirable function s_e, we introduce the following definition. We say that a component X_i is *reachable down* from X_2 if there is a set $\{n(1), \dots, n(m)\} \subset \{1, \dots, k\}$ such that

$$X_i \prec X_{n(1)} \prec X_{n(2)} \prec \cdots \prec X_{n(m)} \prec X_2.$$

We split all components of (X, T_f) into the set U of components which are reachable down from X_2 and the set V of components which are not reachable down from X_2; we include X_2 in U. By the nonexistence of cycles of the form (10), X_1 is among those components which are not reachable down from X_2, whence X_1 is in V.

The definition of being reachable down implies that there is a $\delta > 0$ such that, for each edge with end w in V and end z in U, one has $s(z) - s(w) < d(z, w) - \delta$.

In the case when $O \in V$, put

$$s_e(x) = \begin{cases} s(x) & \text{if } x \in V, \\ s(x) + \frac{\delta}{2} & \text{if } x \in U, \end{cases} \tag{12}$$

and when $O \in U$, put

$$s_e(x) = \begin{cases} s(x) - \frac{\delta}{2} & \text{if } x \in V, \\ s(x) & \text{if } x \in U. \end{cases} \tag{13}$$

In either case, s_e is also a supporting function for f, while for s_e there are no downhill edges from U to V and from V to U. This completes the proof. $\square$

To derive Theorem 2.5 from Lemma 2.8, consider the case of disconnected (X, T_f), pick a supporting function s for f, and observe that we need to examine the two cases: (1) There exists an edge e between two distinct components of (X, T_f) which is downhill for s; (2) There are no such edges.

Lemma 2.8 proves Theorem 2.5 in Case (1), because s_e is obviously different from s.

Case (2). Since all considered graphs are finite, there exists $\omega > 0$ such that, for any edge uv between different components of (X, T_f), $|s(u) - s(v)| \leq d(u, v) - \omega$.

Let X_0 be the component of (X, T_f) containing O. Then the function $\tilde{s}$ given by

$$\tilde{s}(x) = \begin{cases} s(x) & \text{if } x \in X_0, \\ s(x) + \frac{\delta}{2} & \text{otherwise,} \end{cases} \tag{14}$$

is another supporting function for f. This proves Theorem 2.5. $\qquad\square$

After establishing all the necessary grounding, let us come back to the proof of Theorem 2.3.

Proof of Theorem 2.3. (a) Proof of "only if" part. Let (X, E_l) be the downhill graph for a 1-Lipschitz function l. For each edge $\overrightarrow{uv} \in E_l$ consider the transportation problem $\mathbf{1}_u - \mathbf{1}_v$. We add all such problems over all $\overrightarrow{uv} \in E_l$ and get a transportation problem

$$f = \sum_{\overrightarrow{uv} \in E_l} \left(\mathbf{1}_u - \mathbf{1}_v\right), \tag{15}$$

for which the right-hand side of (15) is an optimal transportation plan by Observation 2.1.

This plan has the largest support among optimal transportation plans for f, because, on one hand, $l(f) = \|f\|_{\mathrm{TC}}$, and, on the other hand $|l(\mathbf{1}_x - \mathbf{1}_y)| < d(x, y)$ for any edge xy which is not in the plan (15).

(b) Proof of "if" part. Taking average of support functions s_e constructed in Lemma 2.8 over all edges e between different components of T_f, we obtain a support function s_a, all of whose downhill edges are within one of the components of (X, T_f). By Lemma 2.7, T_f is the downhill graph of s_a. $\qquad\square$

3   Nonexistence of isometric copies of ℓ_∞^k in $\mathrm{TC}(X)$ for some finite graphs X

By the Maurey-Pisier theorem [MP76], the problem of presence of isometric copies of ℓ_∞^k in transportation cost spaces is closely related to the problem of determining the cotype of transportation cost spaces, which is a significant open problem for many metric spaces. See the survey [Nao18, Section 1.2.2], where this problem is restated in terms of universality.

Since ℓ_∞^2 is isometric to ℓ_1^2, it is easy to check that ℓ_∞^2 is isometrically contained in $\mathrm{TC}(X)$ for every X with at least 4 points. Alternatively, it can be derived as a consequence of Theorem 1.8. On the other hand, if X has 3 points, it is isometric to ℓ_∞^2 if and only if X is a tree. See the result (1) in Section 1.1.

If X contains C_4 with distortion 1, then, by virtue of the observation in [AFGZ21] saying that $\mathrm{TC}(C_4) = \ell_\infty^3$, $\mathrm{TC}(X)$ contains ℓ_∞^3 isometrically.

There are two known examples of finite metric spaces X for which $\mathrm{TC}(X)$ contains ℓ_∞^4 isometrically. Historically the first one is a discrete subset of the unit sphere

of ℓ_∞^4 in its ℓ_∞ metric discovered in [KMO20]. The second one is the complete bipartite graph $K_{4,4}$, see [DKO21].

This section purports to detect obstacles preventing an isometric containment of ℓ_∞^k in $\mathrm{TC}(X)$. To do so, we suppose that ℓ_∞^k admits an isometric embedding into $\mathrm{TC}(X)$ for a finite metric space X. Let $(X, E(X))$ be the canonical graph of X and let a sequence $\{e_i\}_{i=1}^k \subset \mathrm{TC}(X)$ be isometrically equivalent to the unit vector basis of ℓ_∞^k.

In the rest of this section, the elements $\{e_i\}$ will be presented by their optimal roadmaps with an understanding that such a presentation is not unique. As before, a reference orientation of $E(X)$ is taken to be fixed.

To elaborate more on the number of optimal roadmaps for elements e_i, we introduce the following definition.

Definition 3.1. Two transportation problems are called *strongly disjoint* if the maximal supports of their optimal roadmaps in $\ell_{1,d}(E)$ are disjoint. Equivalently, two transportation problems are *strongly disjoint* if any two roadmaps for them are disjoint as vectors in $\ell_{1,d}$.

It is not difficult to see that any two transportation problems $f, g \in \mathrm{TC}(X)$, where X is a finite weighted graph, which are isometrically equivalent to the unit vector basis of $\{\ell_1^2\}$ are strongly disjoint. Indeed, if this does not hold, then either $\|f + g\|_{\mathrm{TC}} < \|f\|_{\mathrm{TC}} + \|g\|_{\mathrm{TC}}$ or $\|f - g\|_{\mathrm{TC}} < \|f\|_{\mathrm{TC}} + \|g\|_{\mathrm{TC}}$.

This remark leads to the following statement.

Observation 3.2. *Any pair f_j, $j = 1, 2$ of transportation problems of the form $\sum_{i=1}^k a_{i,j} e_i$, for which $|a_{i,j}| \leq 1$ and also $1 = a_{l,1} a_{l,2} = -a_{m,1} a_{m,2}$ for some pair of indices l, m, is strongly disjoint.*

Proof. Since $\{e_i\}_{i=1}^k$ is isometrically equivalent to the unit vector basis of ℓ_∞^k, problems $\{f_j\}_{j=1}^2$ are isometrically equivalent to the unit vector basis of ℓ_1^2, the statement holds by the remark above. $\square$

The following result on the number of optimal roadmaps for elements $\{e_i\}$ holds.

Proposition 3.3. *For every e_j, there are at least 2^{k-2} disjoint optimal roadmaps in $\ell_{1,d}(E(X))$.*

Proof. Given real numbers $\{a_i\}_{i=1}^k$, set

$$f(a_1, a_2, \ldots, a_k) = \sum_{i=2}^k a_i e_i. \tag{16}$$

For each collection $\theta_1, \theta_2, \theta_3, \ldots, \theta_k$, $\theta_i = \pm 1$, we pick some optimal roadmap $p(\theta_1, \ldots, \theta_k)$ for $f(\theta_1, \ldots, \theta_k)$.

By Observation 3.2, problems

$$f(1, \theta_2, \ldots, \theta_{j-1}, 1, \theta_{j+1}, \ldots, \theta_k) \quad \text{and} \quad f(1, \theta_2, \ldots, \theta_{j-1}, -1, \theta_{j+1}, \ldots, \theta_k)$$

are strongly disjoint. Their sum equals $2f(1, \theta_2, \ldots, \theta_{j-1}, 0, \theta_{j+1}, \ldots, \theta_k)$, while their difference equals $2e_j$. This implies that

$$\frac{1}{2}\left(p(1, \theta_2, \ldots, \theta_{j-1}, 1, \theta_{j+1}, \ldots, \theta_k) - p(1, \theta_2, \ldots, \theta_{j-1}, -1, \theta_{j+1}, \ldots, \theta_k)\right) \qquad (17)$$

is an optimal roadmap for e_j and

$$\frac{1}{2}\left(p(1, \theta_2, \ldots, \theta_{j-1}, 1, \theta_{j+1}, \ldots, \theta_k) + p(1, \theta_2, \ldots, \theta_{j-1}, -1, \theta_{j+1}, \ldots, \theta_k)\right) \qquad (18)$$

is an optimal roadmap for $f(1, \theta_2, \ldots, \theta_{j-1}, 0, \theta_{j+1}, \ldots, \theta_k)$ Strong disjointness of $f(1, \theta_2, \ldots, \theta_{j-1}, 1, \theta_{j+1}, \ldots, \theta_k)$ and $f(1, \theta_2, \ldots, \theta_{j-1}, -1, \theta_{j+1}, \ldots, \theta_k)$ implies that the roadmaps (17) and (18) have the same support.

By Observation 3.2, problems

$$f(1, \theta_2, \ldots, \theta_{j-1}, 0, \theta_{j+1}, \ldots, \theta_k) \quad \text{and} \quad f(1, \tilde{\theta}_2, \ldots, \tilde{\theta}_{j-1}, 0, \tilde{\theta}_{j+1}, \ldots, \tilde{\theta}_k)$$

are strongly disjoint for any two distinct $(k-2)$-tuples $(\theta_2, \ldots, \theta_{j-1}, \theta_{j+1}, \ldots, \theta_k)$ and $(\tilde{\theta}_2, \ldots, \tilde{\theta}_{j-1}, \tilde{\theta}_{j+1}, \ldots, \tilde{\theta}_k)$ consisting of ± 1. Therefore roadmaps (17) are disjoint for distinct $(k-2)$-tuples $(\theta_2, \ldots, \theta_{j-1}, \theta_{j+1}, \ldots, \theta_k)$. $\qquad \square$

Employing Proposition 3.3 one arrives at:

Theorem 3.4. *If* $\mathrm{TC}(X)$ *contains an isometric copy of* ℓ_∞^k, *then the canonical graph* $G(X, E)$ *of* X *contains vertices whose degrees are at least* 2^{k-2}. *Furthermore, the number of such vertices should be nontrivially large, namely, if* $\{e_i\}_{i=1}^k \subset \mathrm{TC}(X)$ *are isometrically equivalent to the unit vector basis of* ℓ_∞^k, *then each vertex contained in a support of any of* $\{e_i\}_{i=1}^k$ *should have degree at least* 2^{k-2}.

Note that $\{e_i\}_{i=1}^k$ do not have to be disjointly supported, see [DKO21, Section 8].

It is worthy to state the following immediate consequence of Theorem 3.4 because it is related to the Bourgain's problem on cotype of $\mathrm{TC}(\mathbb{R}^2)$ which we mentioned in Section 1.1, see also [NS07] in this connection.

Corollary 3.5. *The space* $\mathrm{TC}(L_n)$, *where* L_n *is the plane* $n \times n$ *grid, does not contain* ℓ_∞^5 *isometrically.*

To formulate the next corollary, let us remind the definition of diamond graphs.

Definition 3.6. Diamond graphs $\{D_n\}_{n=0}^\infty$ are defined recursively: The *diamond graph* of level 0 has two vertices joined by an edge of length 1 and is denoted by D_0. The *diamond graph* D_n is obtained from D_{n-1} in the following way. Given an edge $uv \in E(D_{n-1})$, it is replaced by a quadrilateral u, a, v, b, with edges ua, av, vb, bu. (See Figure 1.)

Apparently Definition 3.6 was first introduced in [GNRS04]. We consider D_n as a weighted graph - the weight of each edge is 2^{-1}. It is clear that with these weights
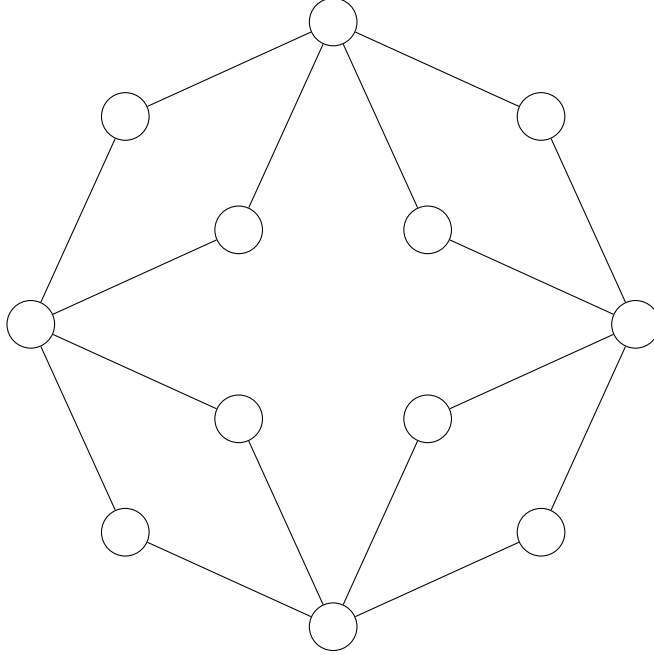
Figure 1: Diamond D_2.

the vertex set of D_{n-1} is naturally isometrically embeddable into the vertex set of D_n. We call those vertices of D_n which are in the set $V(D_n)\backslash V(D_{n-1})$ *vertices of the n-the generation*. Observe that all vertices of the n-th generation in D_n have degree 2.

Corollary 3.7. *The spaces $\{\mathrm{TC}(D_n)\}_{n=1}^{\infty}$ do not contain an isometric copy of ℓ_{∞}^4.*

Proof. Suppose that $\mathrm{TC}(D_n)$ contains ℓ_{∞}^4 isometrically, and $\{e_j\}_{j=1}^4 \subset \mathrm{TC}(D_n)$ are isometrically equivalent to the unit vector basis of ℓ_{∞}^4. By Proposition 3.3, each vertex belonging to support of one of $\{e_j\}_{j=1}^4$ has degree at least $4 = 2^{4-2}$. As observed in the paragraph above, none of these vertices is of the n-th generation, and all of them are in $V(D_{n-1})$. Therefore, $\mathrm{TC}(D_{n-1})$ also contains ℓ_{∞}^4 isometrically. Eventually, we arrive at a contradiction because the maximal degree of vertices of D_1 equals 2. $\qquad\square$

Remark 3.8. Corollary 3.7 is sharp: by the observation of [AFGZ21] mentioned at the beginning of this section, $\mathrm{TC}(D_n)$ contains an isometric copy of ℓ_{∞}^3 for every $n \in \mathbb{N}$.

It is easy to see that a similar argument can be used for recursive sequences of graphs introduced by Lee and Raghavendra [LR10]:

Definition 3.9. Let H and G be two finite connected directed graphs having distinguished vertices which we call *top* and *bottom*, respectively. The *composition* $H \oslash G$ is obtained by replacing each edge $\overrightarrow{uv} \in E(H)$ by a copy of G, the vertex

18

u is identified with the bottom of G and the vertex v is identified with the top of G. Directions of edges in $H \oslash G$ are inherited from G. The *top* and *bottom* of the obtained graph are defined as the top and bottom of H, respectively.

When we consider these graphs as metric spaces we use the graph distances of the underlying undirected weighted graphs - that is, we ignore the directions of edges.

Defining metrics on such graphs we make the following normalization. We assume that the distance between the top and bottom of G is equal to 1. When we replace an edge $e \in E(H)$ by a copy of G, we multiply all weights of edges in this copy of G by $w(e)$. This normalization is chosen because under this normalization the natural embedding of $V(H)$ into $V(H \oslash G)$ is isometric.

Let B be a connected weighted finite simple directed graph having two distinguished vertices, which we call *top* and *bottom*, respectively. Assume that the distance between the top and bottom in B is 1. We use B to construct a recursive family of graphs as follows:

Definition 3.10. We say that graphs $\{B_n\}_{n=0}^{\infty}$ are defined by *recursive composition* or that $\{B_n\}_{n=0}^{\infty}$ is a *recursive sequence* or *recursive family* of graphs if:

- A graph B_0 consists of one directed edge of length 1 with *bottom* being the initial vertex and *top* being the terminal vertex.

- $B_n = B_{n-1} \oslash B$.

The weights of edges in B_n are defined as described above. Now, one can formulate another corollary of Theorem 3.4:

Corollary 3.11. *Let Δ be the maximum degree of B. Then the spaces $\{\mathrm{TC}(B_n)\}_{n=0}^{\infty}$ do not contain ℓ_∞^k isometrically for $k > \log_2 \Delta + 2$.*

Proof. Suppose $\mathrm{TC}(B_n)$ contains ℓ_∞^k isometrically, and $\{e_j\}_{j=1}^{k} \subset \mathrm{TC}(B_n)$ are isometrically equivalent to the unit vector basis of ℓ_∞^k. By Proposition 3.3, each vertex belonging to support of one of $\{e_j\}_{j=1}^{k}$ has degree at least $2^{k-2} > 2^{(\log_2 \Delta + 2) - 2} = \Delta$. By the definition of B_n, none of these vertices can belong to any graph B replacing an edge of B_{n-1} in the last step of construction of B_n, except the top or bottom of B (degree can be $> \Delta$ only at the top and bottom of B). Thus all vertices in the support of $\{e_j\}_{j=1}^{k}$ belong to the vertex set of B_{n-1}. Therefore $\mathrm{TC}(B_{n-1})$ also contains ℓ_∞^k isometrically, and we can repeat the argument for B_{n-1}. Eventually, we arrive at a contradiction because the maximal degree of vertices of $B_1 = B$ equals Δ. $\square$

4   References

[AFGZ21]   M. Alexander, M. Fradelizi, L. C. García-Lirola, A. Zvavitch, Geometry and volume product of finite dimensional Lipschitz-free spaces. *J. Funct. Anal.* **280** (2021), no. 4, 108849.

[AR20]   R. J. Aliaga, A. Rueda Zoca, Points of differentiability of the norm in Lipschitz-free spaces. *J. Math. Anal. Appl.* **489** (2020), no. 2, 124171, 17 pp.

[ABS21]   L. Ambrosio, E. Brué, D. Semola, *Lectures on optimal transport.* Unitext, 130. Springer, Cham, 2021.

[AGS08]   L. Ambrosio, N. Gigli, G. Savaré, *Gradient flows in metric spaces and in the space of probability measures.* Second edition. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2008.

[BMSZ20+]   F. P. Baudier, P. Motakis, T. Schlumprecht, A. Zsák, Stochastic approximation of lamplighter metrics, `arXiv:2003.06093v2`.

[Bou86]   J. Bourgain, The metrical interpretation of superreflexivity in Banach spaces, *Israel J. Math.*, **56** (1986), no. 2, 222–230.

[Cha02]   M. S. Charikar, Similarity estimation techniques from rounding algorithms. *Proceedings of the Thirty-Fourth Annual ACM Symposium on Theory of Computing*, 380–388, ACM, New York, 2002.

[CJ17]   M. Cúth, M. Johanis, Isometric embedding of ℓ_1 into Lipschitz-free spaces and ℓ_∞ into their duals. *Proc. Amer. Math. Soc.* **145** (2017), no. 8, 3409–3421.

[CKK17]   M. Cúth, O. F. K. Kalenda, P. Kaplický, Isometric representation of Lipschitz-free spaces over convex domains in finite-dimensional spaces. *Mathematika* **63** (2017), no. 2, 538–552.

[DDM19]   A. D'Alì, E. Delucchi, M. Michałek. Many faces of symmetric edge polytopes, `arXiv:1910.05193` (2019).

[DGZ93]   R. Deville, G. Godefroy, V. Zizler, *Smoothness and renormings in Banach spaces*, Pitman Monographs and Surveys in Pure and Applied Mathematics, **64**, Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1993.

[DKO20]   S. J. Dilworth, D. Kutzarova, M. I. Ostrovskii, Lipschitz-free spaces on finite metric spaces, *Canad. J. Math.* **72** (2020), no. 3, 774–804.

[DKO21]   S. J. Dilworth, D. Kutzarova, M. I. Ostrovskii, Analysis on Laakso graphs with application to the structure of transportation cost spaces, *Positivity*, **25** (2021), no. 4, 1403–1435.

[Far94]   J. D. Farmer, Extreme points of the unit ball of the space of Lipschitz functions. *Proc. Amer. Math. Soc.* **121** (1994), no. 3, 807–813.

[FG21]   A. Figalli, F. Glaudo, *An invitation to optimal transport, Wasserstein distances, and gradient flows.* EMS Textbooks in Mathematics. EMS Press, Berlin, 2021.

[Gar18]   D. J. H. Garling, *Analysis on Polish spaces and an introduction to optimal transportation.* London Mathematical Society Student Texts, **89**. Cambridge University Press, Cambridge, 2018.

[God10]   A. Godard, Tree metrics and their Lipschitz-free spaces, *Proc. Amer. Math. Soc.*, **138** (2010), no. 12, 4311–4320.

[GL18]   G. Godefroy, N. Lerner, Some natural subspaces and quotient spaces of L_1. *Adv. Oper. Theory* **3** (2018), no. 1, 61–74.

[GP17]   J. Gordon, F. Petrov, Combinatorics of the Lipschitz polytope. *Arnold Math. J.* **3** (2017), no. 2, 205–218.

[GNRS04]   A. Gupta, I. Newman, Y. Rabinovich, A. Sinclair, Cuts, trees and ℓ_1-embeddings of graphs, *Combinatorica*, **24** (2004) 233–269; Conference version in: *40th Annual IEEE Symposium on Foundations of Computer Science*, 1999, pp. 399–408.

[HJM19]   A. Higashitani, K. Jochemko, M. Michalek, Arithmetic aspects of symmetric edge polytopes. *Mathematika* **65** (2019), no. 3, 763–784.

[IT03]   P. Indyk, N. Thaper, Fast image retrieval via embeddings, in: *ICCV 03: Proceedings of the 3rd International Workshop on Statistical and Computational Theories of Vision*, 2003.

[JTZ21]   F. D. Jevtić, M. Timotijevic, R. T. Živaljević, Polytopal Bier spheres and Kantorovich-Rubinstein polytopes of weighted cycles. *Discrete Comput. Geom.* **65** (2021), no. 4, 1275–1286.

[Kan42]   L. V. Kantorovich, On mass transportation (Russian), *Doklady Akad. Nauk SSSR*, (N.S.) **37**, (1942), 199–201; English transl.: *J. Math. Sci.* (N. Y.), **133** (2006), no. 4, 1381–1382.

[Kan48]   L. V. Kantorovich, On a problem of Monge (Russian), *Uspekhi Mat. Nauk,* **3** (1948), no. 2, 225–226; English transl.: *J. Math. Sci.* (N. Y.), **133** (2006), no. 4, 1383.

[KA82]   L. V. Kantorovich, G. P. Akilov, *Functional analysis.* Translated from the Russian by Howard L. Silcock. Second edition. Pergamon Press, Oxford-Elmsford, N.Y., 1982.

[KG49]   L. V. Kantorovich, M. K. Gavurin, Application of mathematical methods in the analysis of cargo flows (Russian), in: *Problems of improving of transport efficiency*, USSR Academy of Sciences Publishers, Moscow, 1949, pp. 110–138.

[KR57]   L. V. Kantorovich, G. S. Rubinstein, On a functional space and certain extremum problems (Russian), *Dokl. Akad. Nauk SSSR* (N.S.), **115** (1957), 1058–1061.

[KR58]   L. V. Kantorovich, G. S. Rubinstein, On a space of completely additive functions (Russian), *Vestnik Leningrad. Univ.*, **13** (1958), no. 7, 52–59.

[KKMR09]   H. Karloff, S. Khot, A. Mehta, Y. Rabani, On earthmover distance, metric labeling, and 0-extension. *SIAM J. Comput.* **39** (2009), no. 2, 371–387.

[KMO20]   S. S. Khan, M. Mim, M. I. Ostrovskii, Isometric copies of ℓ_∞^n and ℓ_1^n in transportation cost spaces on finite metric spaces, in: *The Mathematical Legacy of Victor Lomonosov. Operator Theory*, pp. 189–203, De Gruyter, 2020.

[KN06]   S. Khot, A. Naor, Nonembeddability theorems via Fourier analysis, *Math. Ann.*, **334** (2006), 821–852.

[LR10]   J.R. Lee and P. Raghavendra, *Coarse differentiation and multi-flows in planar graphs*, Discrete Comput. Geom. 43 (2) (2010), 346–362.

[MP76]   B. Maurey, G. Pisier, Séries de variables aléatoires vectorielles indépendantes et propriétés géométriques des espaces de Banach, *Studia Math.*, **58** (1976), no. 1, 45–90.

[MP21]   L. Montrucchio, G. Pistone, Kantorovich distance on finite metric spaces: Arens-Eells norm and CUT norms. *Information Geometry* (2021), 1–37; `DOI:10.1007/s41884-021-00050-w`.

[Nao18]   A. Naor, Metric dimension reduction: a snapshot of the Ribe program, *Proc. Int. Cong. of Math. - 2018*, Rio de Janeiro, Vol. **1**, 759–838.

[Nao21]   A. Naor, Impossibility of almost extension, *Adv. Math.*, **384** (2021), 107761.

[NR17]   A. Naor, Y. Rabani, On Lipschitz extension from finite subsets. *Israel J. Math.* **219** (2017), no. 1, 115–161.

[NS07]   A. Naor, G. Schechtman, Planar Earthmover is not in L_1, *SIAM J. Computing*, **37** (2007), 804–826.

[OO19]   S. Ostrovska, M. I. Ostrovskii, Generalized transportation cost spaces, *Mediterr. J. Math.* **16** (2019), no. 6, Paper No. 157.

[OO20]   S. Ostrovska, M. I. Ostrovskii, On relations between transportation cost spaces and ℓ_1, *J. Math. Anal. Appl.*, **491** (2020), no. 2, 124338.

[Ost07]   M. I. Ostrovskii, Isometric classification of Sobolev spaces on graphs. *Colloq. Math.* **109** (2007), no. 2, 287–295.

[Ost13]   M. I. Ostrovskii, *Metric Embeddings: Bilipschitz and Coarse Embeddings into Banach Spaces*, de Gruyter Studies in Mathematics, **49**. Walter de Gruyter & Co., Berlin, 2013.

[Oxl11] J. Oxley, *Matroid theory.* Second edition. Oxford Graduate Texts in Mathematics, **21**. Oxford University Press, Oxford, 2011.

[RTG98] Y. Rubner, C. Tomasi, L. J. Guibas, A metric for distributions with applications to image databases, *Proceedings ICCV 1998*, pp. 59–66; doi:10.1109/ICCV.1998.710701.

[Sch03] A. Schrijver, *Combinatorial Optimization. Polyhedra and Efficiency*, Berlin, Springer, 2003.

[Tol30] A. N. Tolstoi, Methods of finding the minimal total kilometrage in cargo-transportation planning in space (Russian), in: Transportation Planning, Volume I, TransPress of the National Commissariat of Transportation, Moscow, 1930, pp. 23–55.

[Tol39] A. N. Tolstoi, Methods of removing irrational transportation in planning (Russian), Socialist Transport **9** (1939) 28–51.

[Ver15] A. M. Vershik, Classification of finite metric spaces and combinatorics of convex polytopes. *Arnold Math. J.* **1** (2015), no. 1, 75–81.

[Vil03] C. Villani, *Topics in optimal transportation.* Graduate Studies in Mathematics, **58**. American Mathematical Society, Providence, RI, 2003.

[Vil09] C. Villani, *Optimal transport.* Old and new. Grundlehren der Mathematischen Wissenschaften, **338**. Springer-Verlag, Berlin, 2009.

[Wea99] N. Weaver, *Lipschitz algebras*, First Edition, World Scientific Publishing Co., Inc., River Edge, NJ, 1999.

[Wea18] N. Weaver, *Lipschitz algebras*, Second Edition, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2018.

[Whi33] H. Whitney, 2-isomorphic graphs, *Amer. J. Math.*, **55** (1933), 73–84.

DEPARTMENT OF MATHEMATICS, ATILIM UNIVERSITY, 06830 INCEK, ANKARA, TURKEY

E-mail address: sofia.ostrovska@atilim.edu.tr

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, ST. JOHN'S UNIVERSITY, 8000 UTOPIA PARKWAY, QUEENS, NY 11439, USA

E-mail address: ostrovsm@stjohns.edu