Finite-dimensional Transportation Cost Spaces

Mikhail Ostrovskii
St. John's University
New York, USA
https://facpub.stjohns.edu/ostrovsm/

Summer School
"Topics in Banach Space Theory"
Castro Urdiales, Spain
July 6-10, 2026

On terminology

- In this course, we study Transportation Cost Spaces.

On terminology

- In this course, we study Transportation Cost Spaces.
- It is important to know that notions we study in this course are known by many other names.

On terminology

- In this course, we study Transportation Cost Spaces.
- It is important to know that notions we study in this course are known by many other names.
- I would like to mention the other names of these notions and topics in which they appeared.

On terminology

- In this course, we study Transportation Cost Spaces.
- It is important to know that notions we study in this course are known by many other names.
- I would like to mention the other names of these notions and topics in which they appeared.
- Topic 1: Distance between probability measures. Names: Earth mover distance, Gini distance, Kantorovich distance, Wasserstein distance.

On terminology

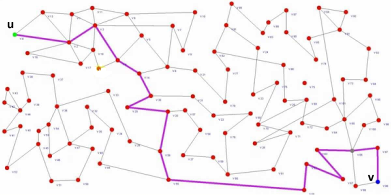
- In this course, we study Transportation Cost Spaces.
- It is important to know that notions we study in this course are known by many other names.
- I would like to mention the other names of these notions and topics in which they appeared.
- Topic 1: Distance between probability measures. Names: Earth mover distance, Gini distance, Kantorovich distance, Wasserstein distance.
- Topic 2: Banach space of transportation problems with the norm induced by optimal transport: Names: Transportation cost, Kantorovich-Rubinstein norm, Optimal transport, Wasserstein norm.

On terminology

- In this course, we study Transportation Cost Spaces.
- It is important to know that notions we study in this course are known by many other names.
- I would like to mention the other names of these notions and topics in which they appeared.
- Topic 1: Distance between probability measures. Names: Earth mover distance, Gini distance, Kantorovich distance, Wasserstein distance.
- Topic 2: Banach space of transportation problems with the norm induced by optimal transport: Names: Transportation cost, Kantorovich-Rubinstein norm, Optimal transport, Wasserstein norm.
- Topic 3: Free linear spaces generated by metric spaces with the norm induced by the metric: Arens-Eells spaces, free Banach spaces, Lipschitz-free spaces.

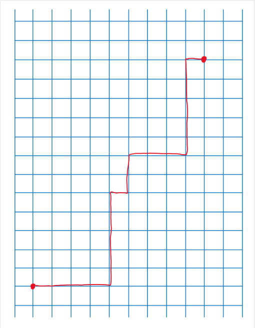
Finite metric spaces

We study finite metric spaces (X, d).



Without loss of generality, we may assume:

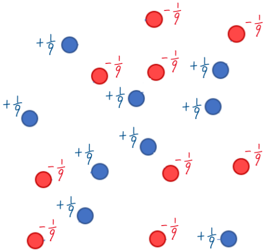
- X is a finite weighted graph
- $d(u, v)$ is shortest-path distance



Important example: planar grid $(X = [n]^2)$ with Manhattan metric

1. Transportation problem

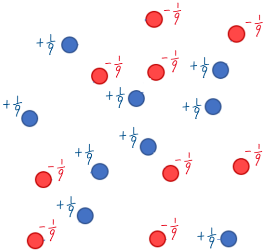
A signed measure μ on (X, d) with total mass 0.

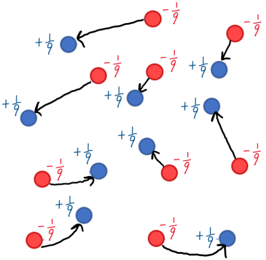


1. Transportation problem
A signed measure μ on (X, d) with total mass 0.

2. Transportation plan
A representation of the form

$$\mu = \sum_{i=1}^{n} a_i(\mathbf{1}_{x_i} - \mathbf{1}_{y_i}), \quad a_i \geq 0.$$





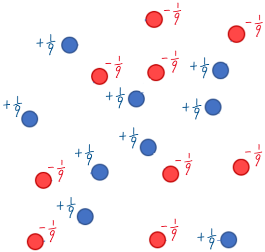
Transportation cost

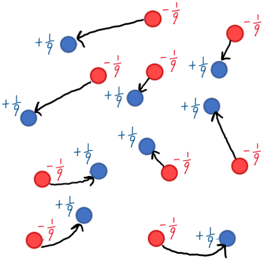
1. Transportation problem
A signed measure μ on (X, d) with total mass 0.

2. Transportation plan
A representation of the form

$$\mu = \sum_{i=1}^{n} a_i(\mathbf{1}_{x_i} - \mathbf{1}_{y_i}), \quad a_i \geq 0.$$

3. Transportation cost (TC) of the plan $\sum_{i=1}^{n} a_i\, d(x_i, y_i)$;
$\|\mu\|_{\mathrm{TC}} = \inf_{\text{plans}} \sum_{i=1}^{n} a_i\, d(x_i, y_i)$





Transportation cost (more formal presentation)

- Let (X, d) be a finite metric space. Consider a real-valued measure μ on X satisfying $\mu(X) = 0$. A natural and important interpretation of such a real measure is considering it as a *transportation problem*: one needs to transport certain product from points where $\mu(v) > 0$ to points where $\mu(v) < 0$.

Transportation cost (more formal presentation)

- Let (X, d) be a finite metric space. Consider a real-valued measure μ on X satisfying $\mu(X) = 0$. A natural and important interpretation of such a real measure is considering it as a *transportation problem*: one needs to transport certain product from points where $\mu(v) > 0$ to points where $\mu(v) < 0$.
- One can easily see that a transportation problem μ can be represented as

$$\mu = a_1(\mathbf{1}_{x_1} - \mathbf{1}_{y_1}) + a_2(\mathbf{1}_{x_2} - \mathbf{1}_{y_2}) + \dots + a_n(\mathbf{1}_{x_n} - \mathbf{1}_{y_n}), \quad (1)$$

where $a_i \geq 0$, $x_i, y_i \in X$, and $\mathbf{1}_u(x)$ for $u \in X$ is the unit measure at u.

Transportation cost (more formal presentation)

- Let (X, d) be a finite metric space. Consider a real-valued measure μ on X satisfying $\mu(X) = 0$. A natural and important interpretation of such a real measure is considering it as a *transportation problem*: one needs to transport certain product from points where $\mu(v) > 0$ to points where $\mu(v) < 0$.
- One can easily see that a transportation problem μ can be represented as

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

where $a_i \geq 0$, $x_i, y_i \in X$, and $\mathbf{1}_u(x)$ for $u \in X$ is the unit measure at u.
- We call each such representation a *transportation plan* for μ, it can be interpreted as plan of moving a_i units of the product from x_i to y_i.

Transportation cost

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

Transportation cost

- The *cost* of the transportation plan (1) is defined as $\sum_{i=1}^{n} a_i d(x_i, y_i)$.
- We introduce the *transportation cost* norm $\|\mu\|_{\mathrm{TC}}$ on the linear space of all transportation problems on X as the infimum of costs of transportation plans satisfying (1). We get a finite-dimensional Banach space which we denote $\mathrm{TC}(X)$. This definition goes back to Kantorovich (1942).

Transportation cost

- The *cost* of the transportation plan (1) is defined as $\sum_{i=1}^{n} a_i d(x_i, y_i)$.
- We introduce the *transportation cost* norm $\|\mu\|_{\mathrm{TC}}$ on the linear space of all transportation problems on X as the infimum of costs of transportation plans satisfying (1). We get a finite-dimensional Banach space which we denote $\mathrm{TC}(X)$. This definition goes back to Kantorovich (1942).
- **Exercise:** Show that $\|\cdot\|_{\mathrm{TC}}$ is a norm.

Transportation cost

- The *cost* of the transportation plan (1) is defined as $\sum_{i=1}^{n} a_i d(x_i, y_i)$.
- We introduce the *transportation cost* norm $\|\mu\|_{\mathrm{TC}}$ on the linear space of all transportation problems on X as the infimum of costs of transportation plans satisfying (1). We get a finite-dimensional Banach space which we denote $\mathrm{TC}(X)$. This definition goes back to Kantorovich (1942).
- **Exercise:** Show that $\|\cdot\|_{\mathrm{TC}}$ is a norm.
- **Remark:** I advise you to do the exercises if you want to **start** doing research in this area.

- For transportation problems defined above there is always an *optimal transportation plan*, that is a transportation plan for which $\sum_{i=1}^{n} a_i d(x_i, y_i) = \|\mu\|_{\mathrm{TC}}$.

- For transportation problems defined above there is always an *optimal transportation plan*, that is a transportation plan for which $\sum_{i=1}^{n} a_i d(x_i, y_i) = \|\mu\|_{\mathrm{TC}}$.
- This is done in steps (exercise). **Step 1:** Eliminate terms $a_i(\mathbf{1}_{x_i} - \mathbf{1}_{y_i})$ for which $\mu(x_i) = 0$ or $\mu(y_i) = 0$.

- For transportation problems defined above there is always an *optimal transportation plan*, that is a transportation plan for which $\sum_{i=1}^{n} a_i d(x_i, y_i) = \|\mu\|_{\mathrm{TC}}$.
- This is done in steps (exercise). **Step 1:** Eliminate terms $a_i(\mathbf{1}_{x_i} - \mathbf{1}_{y_i})$ for which $\mu(x_i) = 0$ or $\mu(y_i) = 0$.
- **Step 2:** After that do cancellation to achieve: all transportation is from positive support to negative support.

- For transportation problems defined above there is always an *optimal transportation plan*, that is a transportation plan for which $\sum_{i=1}^{n} a_i d(x_i, y_i) = \|\mu\|_{\mathrm{TC}}$.
- This is done in steps (exercise). **Step 1:** Eliminate terms $a_i(\mathbf{1}_{x_i} - \mathbf{1}_{y_i})$ for which $\mu(x_i) = 0$ or $\mu(y_i) = 0$.
- **Step 2:** After that do cancellation to achieve: all transportation is from positive support to negative support.
- **Step 3:** Check that the resulting set of transportation plans is compact in the natural topology. This implies the existence of an optimal transportation plan.

- For transportation problems defined above there is always an *optimal transportation plan*, that is a transportation plan for which $\sum_{i=1}^{n} a_i d(x_i, y_i) = \|\mu\|_{\mathrm{TC}}$.
- This is done in steps (exercise). **Step 1:** Eliminate terms $a_i(\mathbf{1}_{x_i} - \mathbf{1}_{y_i})$ for which $\mu(x_i) = 0$ or $\mu(y_i) = 0$.
- **Step 2:** After that do cancellation to achieve: all transportation is from positive support to negative support.
- **Step 3:** Check that the resulting set of transportation plans is compact in the natural topology. This implies the existence of an optimal transportation plan.
- **Disclaimer:** In this mini-course I do not even touch (at least) three important things: (a) Isometric classification developed by Alexander, Fradelizi, García-Lirola, and Zvavitch [AFGZ21], (b) Stochastic bases studied by Medina and Tresch [MT26+], and (c) Combinatorial theory initiated by Vershik [Ver15] (see [OO22] for more references).

- **Question:** Can $\mathrm{TC}(X)$ be isometric ℓ_p^n for some $1 \le p \le \infty$, $n \ge 2$?

- **Question:** Can $\mathrm{TC}(X)$ be isometric ℓ_p^n for some $1 \le p \le \infty$, $n \ge 2$?
- **Answer:** Can be isometric to ℓ_1^n for any n. Can be isometric to ℓ_∞^2 and ℓ_∞^3. Cannot be isometric to any other of the listed spaces.

- **Question:** Can $\mathrm{TC}(X)$ be isometric ℓ_p^n for some $1 \leq p \leq \infty$, $n \geq 2$?
- **Answer:** Can be isometric to ℓ_1^n for any n. Can be isometric to ℓ_∞^2 and ℓ_∞^3. Cannot be isometric to any other of the listed spaces.
- **Positive answers:**

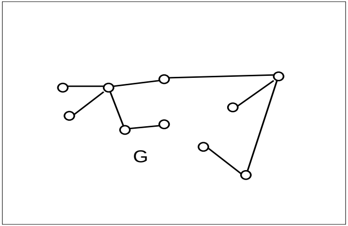
- **Question:** Can $\mathrm{TC}(X)$ be isometric ℓ_p^n for some $1 \le p \le \infty$, $n \ge 2$?
- **Answer:** Can be isometric to ℓ_1^n for any n. Can be isometric to ℓ_∞^2 and ℓ_∞^3. Cannot be isometric to any other of the listed spaces.
- **Positive answers:**
- **Exercise 1:** $\mathrm{TC}(C_4) = \ell_\infty^3$ (here $=$ means "isometric"). C_4 is a cycle of length 4. Find three transportation problems f_1, f_2, and f_3, that satisfy $\|\sum a_i f_i\|_{\mathrm{TC}} = \max_{1\le i\le 3} |a_i|$. [AFGZ21].

- **Question:** Can $\mathrm{TC}(X)$ be isometric ℓ_p^n for some $1 \le p \le \infty$, $n \ge 2$?
- **Answer:** Can be isometric to ℓ_1^n for any n. Can be isometric to ℓ_∞^2 and ℓ_∞^3. Cannot be isometric to any other of the listed spaces.
- **Positive answers:**
- **Exercise 1:** $\mathrm{TC}(C_4) = \ell_\infty^3$ (here $=$ means "isometric"). C_4 is a cycle of length 4. Find three transportation problems f_1, f_2, and f_3, that satisfy $\|\sum a_i f_i\|_{\mathrm{TC}} = \max_{1\le i\le 3} |a_i|$. [AFGZ21].
- **Exercise 2:** The space $\mathrm{TC}(\{0, 1, 2, \dots, n\})$ where $\{0, 1, 2, \dots, n\}$ is endowed with the topology inherited from the real line, is isometric to ℓ_1^n.

- **Question:** Can $\mathrm{TC}(X)$ be isometric ℓ_p^n for some $1 \le p \le \infty$, $n \ge 2$?
- **Answer:** Can be isometric to ℓ_1^n for any n. Can be isometric to ℓ_∞^2 and ℓ_∞^3. Cannot be isometric to any other of the listed spaces.
- **Positive answers:**
- **Exercise 1:** $\mathrm{TC}(C_4) = \ell_\infty^3$ (here $=$ means "isometric"). C_4 is a cycle of length 4. Find three transportation problems f_1, f_2, and f_3, that satisfy $\|\sum a_i f_i\|_{\mathrm{TC}} = \max_{1\le i\le 3} |a_i|$. [AFGZ21].
- **Exercise 2:** The space $\mathrm{TC}(\{0,1,2,\dots,n\})$ where $\{0,1,2,\dots,n\}$ is endowed with the topology inherited from the real line, is isometric to ℓ_1^n.
- **Exercise 3:** The spaces ℓ_1^2 and ℓ_∞^2 are isometric.

- **Definition:** A *tree* is a connected graph without cycles.

- **Definition:** A *tree* is a connected graph without cycles.



-

Figure: A tree.

- **Definition:** A *tree* is a connected graph without cycles.

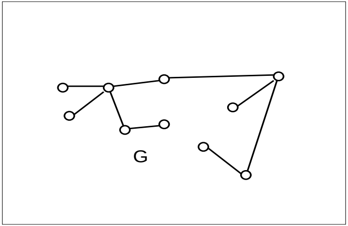


Figure: A tree.

- **Exercise 4:** Let T_n be a tree with n vertices. We allow tree to be *weighted*. This means that edges can be of any positive lengths. We endow the set of vertices of a tree with the metric defined as the length of the path between vertices. In a tree such path is unique.
  Prove that $\mathrm{TC}(T_n)$ is isometric to ℓ_1^{n-1}.
  (We shall need this fact.)

Additional Problems and Exercises

- **More difficult Exercise:** (Converse to the exercise on trees). If a finite metric space X is such that $\mathrm{TC}(X) = \ell_1^{n-1}$, then X is isometric to a (possibly weighted) tree with n vertices.

Additional Problems and Exercises

- **More difficult Exercise:** (Converse to the exercise on trees). If a finite metric space X is such that $\mathrm{TC}(X) = \ell_1^{n-1}$, then X is isometric to a (possibly weighted) tree with n vertices.
- **Open Problem:** Does there exist a finite metric space X such that $\mathrm{TC}(X)$ contains isometrically ℓ_∞^n for $n \geq 5$?

Additional Problems and Exercises

- **More difficult Exercise:** (Converse to the exercise on trees). If a finite metric space X is such that $\mathrm{TC}(X) = \ell_1^{n-1}$, then X is isometric to a (possibly weighted) tree with n vertices.
- **Open Problem:** Does there exist a finite metric space X such that $\mathrm{TC}(X)$ contains isometrically ℓ_∞^n for $n \geq 5$?
- **(Relevant) Theorem:** Let G be a weighted graph for which $\mathrm{TC}(G)$ contains an isometric copy of ℓ_∞^k. Then G 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}. [OO22]

Additional Problems and Exercises

- **More difficult Exercise:** (Converse to the exercise on trees). If a finite metric space X is such that $\mathrm{TC}(X) = \ell_1^{n-1}$, then X is isometric to a (possibly weighted) tree with n vertices.
- **Open Problem:** Does there exist a finite metric space X such that $\mathrm{TC}(X)$ contains isometrically ℓ_∞^n for $n \geq 5$?
- **(Relevant) Theorem:** Let G be a weighted graph for which $\mathrm{TC}(G)$ contains an isometric copy of ℓ_∞^k. Then G 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}. [OO22]
- **Theorem [AFGZ21, Corollary 3.8]:** The space $\mathrm{TC}(X)$ cannot be isometric to ℓ_∞^n for $n \geq 4$.

- **Exercise (with a hint):** The space $\mathrm{TC}(K_{4,4})$ contains ℓ_∞^4 isometrically [DKO21].

- **Exercise (with a hint):** The space $\mathrm{TC}(K_{4,4})$ contains ℓ_∞^4 isometrically [DKO21].
- **Hint:** Let $\{a, b, c, d\}$ and $\{e, f, g, h\}$ be a bipartition. Use the basis:

	a	b	c	d	e	f	g	h
f_1	1/2	-1/2	1/2	-1/2	0	0	0	0
f_2	1/2	1/2	-1/2	-1/2	0	0	0	0
f_3	0	0	0	0	1/2	-1/2	1/2	-1/2
f_4	0	0	0	0	1/2	1/2	-1/2	-1/2

- The following description of $\mathrm{TC}(G)$ is very useful:

$$\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G),$$

where G is a connected weighted graph, $\ell_{1,d}(E(G))$ is a weighted ℓ_1 on $E(G)$, the subspace $Z(G)$ is a cycle space in $\ell_{1,d}(E(G))$, and Y/Z denotes the quotient space of a Banach space over its subspace. The first version of this result is [Ost13, Proposition 10.10]; further versions are in [OO20, AFGZ21, DKO21, OO22, SV26].

- The following description of $\mathrm{TC}(G)$ is very useful:
$$\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G),$$
where G is a connected weighted graph, $\ell_{1,d}(E(G))$ is a weighted ℓ_1 on $E(G)$, the subspace $Z(G)$ is a cycle space in $\ell_{1,d}(E(G))$, and Y/Z denotes the quotient space of a Banach space over its subspace. The first version of this result is [Ost13, Proposition 10.10]; further versions are in [OO20, AFGZ21, DKO21, OO22, SV26].
- Proving the equality above, we assume that the graph G is *connected*, *oriented* and *weighted*.

- The following description of $\mathrm{TC}(G)$ is very useful:

$$\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G),$$

where G is a connected weighted graph, $\ell_{1,d}(E(G))$ is a weighted ℓ_1 on $E(G)$, the subspace $Z(G)$ is a cycle space in $\ell_{1,d}(E(G))$, and Y/Z denotes the quotient space of a Banach space over its subspace. The first version of this result is [Ost13, Proposition 10.10]; further versions are in [OO20, AFGZ21, DKO21, OO22, SV26].
- Proving the equality above, we assume that the graph G is *connected*, *oriented* and *weighted*.
    - *Connected* means that any two vertices are connected by a finite sequence of edges.

- The following description of $\mathrm{TC}(G)$ is very useful:

$$\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G),$$

where G is a connected weighted graph, $\ell_{1,d}(E(G))$ is a weighted ℓ_1 on $E(G)$, the subspace $Z(G)$ is a cycle space in $\ell_{1,d}(E(G))$, and Y/Z denotes the quotient space of a Banach space over its subspace. The first version of this result is [Ost13, Proposition 10.10]; further versions are in [OO20, AFGZ21, DKO21, OO22, SV26].
- Proving the equality above, we assume that the graph G is *connected*, *oriented* and *weighted*.
  - *Connected* means that any two vertices are connected by a finite sequence of edges.
  - *Oriented* means that each edge is assigned a direction (arrow) and writing $\vec{e} = uv$ we mean that v is the head and u is the tail of the arrow.

- The following description of $\mathrm{TC}(G)$ is very useful:

$$\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G),$$

where G is a connected weighted graph, $\ell_{1,d}(E(G))$ is a weighted ℓ_1 on $E(G)$, the subspace $Z(G)$ is a cycle space in $\ell_{1,d}(E(G))$, and Y/Z denotes the quotient space of a Banach space over its subspace. The first version of this result is [Ost13, Proposition 10.10]; further versions are in [OO20, AFGZ21, DKO21, OO22, SV26].

- Proving the equality above, we assume that the graph G is *connected*, *oriented* and *weighted*.
  - *Connected* means that any two vertices are connected by a finite sequence of edges.
  - *Oriented* means that each edge is assigned a direction (arrow) and writing $\vec{e} = uv$ we mean that v is the head and u is the tail of the arrow.
  - *Weighted* means that each edge uv is assigned the *length function* $d_G(u, v) > 0$, and that the formula $d_G(u, v) =$ (the total weight of the shortest (no direction) path joining u and v) extends the length function to all pairs of vertices .

- A graph is called *unweighted* if the length function is equal to 1 on all edges.

- **Definition:** For finite-dimensional Banach spaces X and Y, we say that X is a *quotient space of* Y, if there is a linear operator $Q : Y \to X$ such that $Q(B(Y)) = B(X)$, where $B(X)$ denotes the *unit ball*, that is, $B(X) = \{x \in X : \|x\| \leq 1\}$. Let U be the kernel of Q, we write $X = Y/U$. The operator Q is called a *quotient map*.

- **Definition:** For finite-dimensional Banach spaces X and Y, we say that X is a *quotient space of* Y, if there is a linear operator $Q : Y \to X$ such that $Q(B(Y)) = B(X)$, where $B(X)$ denotes the *unit ball*, that is, $B(X) = \{x \in X : \|x\| \leq 1\}$. Let U be the kernel of Q, we write $X = Y/U$. The operator Q is called a *quotient map*.
- We say that a finite dimensional Banach space is *polyhedral* if its unit ball is a convex hull of a finite set.

- **Definition:** For finite-dimensional Banach spaces X and Y, we say that X is a *quotient space of* Y, if there is a linear operator $Q : Y \to X$ such that $Q(B(Y)) = B(X)$, where $B(X)$ denotes the *unit ball*, that is, $B(X) = \{x \in X : \|x\| \leq 1\}$. Let U be the kernel of Q, we write $X = Y/U$. The operator Q is called a *quotient map*.
- We say that a finite dimensional Banach space is *polyhedral* if its unit ball is a convex hull of a finite set.
- **Important observations (A):** The quotient of a (finite-dimensional) polyhedral space is also polyhedral.

- **Definition:** For finite-dimensional Banach spaces X and Y, we say that X is a *quotient space of* Y, if there is a linear operator $Q: Y \to X$ such that $Q(B(Y)) = B(X)$, where $B(X)$ denotes the *unit ball*, that is, $B(X) = \{x \in X : \|x\| \le 1\}$. Let U be the kernel of Q, we write $X = Y/U$. The operator Q is called a *quotient map*.
- We say that a finite dimensional Banach space is *polyhedral* if its unit ball is a convex hull of a finite set.
- **Important observations (A):** The quotient of a (finite-dimensional) polyhedral space is also polyhedral.
- **(B)** For each $x \in X$,

$$\|x\| = \inf\{\|y\| : Qy = x.\}$$

This is true because the equality $Q(B(Y)) = B(X)$ implies that the equality in the previous line holds if $\|x\| = 1$.

- We introduce the *weighted* ℓ_1*-space* $\ell_{1,d}(E(G))$ on the edge set of a weighted graph $E(G)$. It consists of real-valued functions $F : E(G) \to \mathbb{R}$ with the norm

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

- We introduce the *weighted ℓ_1-space* $\ell_{1,d}(E(G))$ on the edge set of a weighted graph $E(G)$. It consists of real-valued functions $F : E(G) \to \mathbb{R}$ with the norm

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

- The space $\ell_{1,d}(E(G))$ is polyhedral because its unit ball is the convex hull of functions $\pm f_e$, where f_e is equal to 0 on all edges except e, and equal to $\frac{1}{d_G(u,v)}$ if $e = uv$.

- We introduce the *weighted ℓ_1-space* $\ell_{1,d}(E(G))$ on the edge set of a weighted graph $E(G)$. It consists of real-valued functions $F : E(G) \to \mathbb{R}$ with the norm

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

- The space $\ell_{1,d}(E(G))$ is polyhedral because its unit ball is the convex hull of functions $\pm f_e$, where f_e is equal to 0 on all edges except e, and equal to $\frac{1}{d_G(u,v)}$ if $e = uv$.
- It is well-known that ℓ_p^n, $1 < p < \infty$, $n \geq 2$ and their subspaces of dimension ≥ 2, are not polyhedral.

- We introduce the *weighted ℓ_1-space* $\ell_{1,d}(E(G))$ on the edge set of a weighted graph $E(G)$. It consists of real-valued functions $F : E(G) \to \mathbb{R}$ with the norm

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

- The space $\ell_{1,d}(E(G))$ is polyhedral because its unit ball is the convex hull of functions $\pm f_e$, where f_e is equal to 0 on all edges except e, and equal to $\frac{1}{d_G(u,v)}$ if $e = uv$.
- It is well-known that ℓ_p^n, $1 < p < \infty$, $n \geq 2$ and their subspaces of dimension ≥ 2, are not polyhedral.
- Therefore the equality $\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G)$ implies that $\mathrm{TC}(X)$ cannot contain isometric copies of ℓ_p^n, $1 < p < \infty$, $n \geq 2$ and their subspaces of dimension ≥ 2.

- The complete formal proof of $\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G)$ looks rather complicated, but the main idea is simple: one can easily guess a suitable linear map $Q : \ell_{1,d}(E(G))) \to \mathrm{TC}(X)$, and then verify that it is actually a quotient map, and its kernel is the *cycle space* as it is defined in Algebraic Graph Theory, usually denoted $Z(G)$.

- The complete formal proof of $\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G)$ looks rather complicated, but the main idea is simple: one can easily guess a suitable linear map $Q : \ell_{1,d}(E(G))) \to \mathrm{TC}(X)$, and then verify that it is actually a quotient map, and its kernel is the *cycle space* as it is defined in Algebraic Graph Theory, usually denoted $Z(G)$.
- It looks like there is only one natural candidate for Q: We derive from any real-valued function F on $E(G)$ the following transportation plan.

- The complete formal proof of $\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G)$ looks rather complicated, but the main idea is simple: one can easily guess a suitable linear map $Q : \ell_{1,d}(E(G))) \to \mathrm{TC}(X)$, and then verify that it is actually a quotient map, and its kernel is the *cycle space* as it is defined in Algebraic Graph Theory, usually denoted $Z(G)$.
- It looks like there is only one natural candidate for Q: We derive from any real-valued function F on $E(G)$ the following transportation plan.
- Let $\vec{e} = uv$. If $F(e) > 0$, we transport $F(e)$ units of the product from u to v. If $F(e) < 0$, we transport $-F(e) = |F(e)|$ unit of the product from v to u. The corresponding plan (as any transportation plan) can be regarded as a transportation problem. We let this transportation problem be $Q(F)$.

- The complete formal proof of $\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G)$ looks rather complicated, but the main idea is simple: one can easily guess a suitable linear map $Q : \ell_{1,d}(E(G))) \to \mathrm{TC}(X)$, and then verify that it is actually a quotient map, and its kernel is the *cycle space* as it is defined in Algebraic Graph Theory, usually denoted $Z(G)$.
- It looks like there is only one natural candidate for Q: We derive from any real-valued function F on $E(G)$ the following transportation plan.
- Let $\vec{e} = uv$. If $F(e) > 0$, we transport $F(e)$ units of the product from u to v. If $F(e) < 0$, we transport $-F(e) = |F(e)|$ unit of the product from v to u. The corresponding plan (as any transportation plan) can be regarded as a transportation problem. We let this transportation problem be $Q(F)$.
- **Important:** One can easily verify that the cost of this transportation plan is $\|F\|_{1,d} := \sum_{uv \in E(G)} |F(uv)| d_G(u, v)$.

- To complete the proof, we need to prove two things: **(a)** Q is a quotient map. **(b)** Its kernel is the cycle space (defined below).

- To complete the proof, we need to prove two things: **(a)** Q is a quotient map. **(b)** Its kernel is the cycle space (defined below).
- **(a)** We need to show that for each transportation problem μ there is a function $F \in \ell_{1,d}(E(G))$ such that $QF = \mu$ and $\|F\|_{1,d} = \|\mu\|_{\mathrm{TC}}$. This is done as follows. We know that there exists an optimal plan P for μ: $\mu = a_1(\mathbf{1}_{x_1} - \mathbf{1}_{y_1}) + \cdots + a_n(\mathbf{1}_{x_n} - \mathbf{1}_{y_n})$.

- To complete the proof, we need to prove two things: **(a)** Q is a quotient map. **(b)** Its kernel is the cycle space (defined below).
- **(a)** We need to show that for each transportation problem μ there is a function $F \in \ell_{1,d}(E(G))$ such that $QF = \mu$ and $\|F\|_{1,d} = \|\mu\|_{\mathrm{TC}}$. This is done as follows. We know that there exists an optimal plan P for μ: $\mu = a_1(\mathbf{1}_{x_1} - \mathbf{1}_{y_1}) + \cdots + a_n(\mathbf{1}_{x_n} - \mathbf{1}_{y_n})$.
- If $x_i y_i$ were edges, it would immediately lead to the corresponding function F. For non-adjacent pairs x_i, y_i, we replace $a_i(\mathbf{1}_{x_i} - \mathbf{1}_{y_i})$ with a more detailed plan along the shortest path joining x_i and y_i. Such plans (combined for all x_i, y_i) lead $F \in \ell_{1,d}(E(G))$ satisfying $QF = \mu$ and $\|F\|_{1,d} = \|\mu\|_{\mathrm{TC}}$.

- To complete the proof, we need to prove two things: **(a)** Q is a quotient map. **(b)** Its kernel is the cycle space (defined below).
- **(a)** We need to show that for each transportation problem μ there is a function $F \in \ell_{1,d}(E(G))$ such that $QF = \mu$ and $\|F\|_{1,d} = \|\mu\|_{\mathrm{TC}}$. This is done as follows. We know that there exists an optimal plan P for μ: $\mu = a_1(\mathbf{1}_{x_1} - \mathbf{1}_{y_1}) + \cdots + a_n(\mathbf{1}_{x_n} - \mathbf{1}_{y_n})$.
- If $x_i y_i$ were edges, it would immediately lead to the corresponding function F. For non-adjacent pairs x_i, y_i, we replace $a_i(\mathbf{1}_{x_i} - \mathbf{1}_{y_i})$ with a more detailed plan along the shortest path joining x_i and y_i. Such plans (combined for all x_i, y_i) lead $F \in \ell_{1,d}(E(G))$ satisfying $QF = \mu$ and $\|F\|_{1,d} = \|\mu\|_{\mathrm{TC}}$.
- So Q is the quotient map.

- **(b)** To determine the kernel of Q, we observe the following formula for Q. Let $F \in \ell_{1,d}(E(G))$, we have

$$(QF)(v) = \sum_{\{e,\ v \text{ is the head of } e\}} F(e) - \sum_{\{e,\ v \text{ is the tail of e}\}} F(e)$$

- **(b)** To determine the kernel of Q, we observe the following formula for Q. Let $F \in \ell_{1,d}(E(G))$, we have

$$(QF)(v) = \sum_{\{e,\ v \text{ is the head of } e\}} F(e) - \sum_{\{e,\ v \text{ is the tail of e}\}} F(e)$$

- **Definition:** Let $Z = \ker Q$. This subspace of $\ell_{1,d}(E(G))$ is called the *cycle space*. Actually, the cycle spaces is usually considered as the linear space of functions on $E(G)$, without any norm.

- **(b)** To determine the kernel of Q, we observe the following formula for Q. Let $F \in \ell_{1,d}(E(G))$, we have

$$(QF)(v) = \sum_{\{e,\ v \text{ is the head of } e\}} F(e) - \sum_{\{e,\ v \text{ is the tail of e}\}} F(e)$$

- **Definition:** Let $Z = \ker Q$. This subspace of $\ell_{1,d}(E(G))$ is called the *cycle space*. Actually, the cycle spaces is usually considered as the linear space of functions on $E(G)$, without any norm.
- The name 'cycle space' is chosen because this subspace is spanned by the set of signed indicator functions of all cycles (see the next slide).

- 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

$$\chi_C(e) = \begin{cases} 1 & \text{if } e \in C \text{ and its orientations in } C \text{ and } G \text{ coincide} \\ -1 & \text{if } e \in C \text{ but its orientations in } C \text{ and } G \text{ differ} \\ 0 & \text{if } e \notin C. \end{cases} \quad (3)$$

- 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

$$\chi_C(e) = \begin{cases} 1 & \text{if } e \in C \text{ and its orientations in } C \text{ and } G \text{ coincide} \\ -1 & \text{if } e \in C \text{ but its orientations in } C \text{ and } G \text{ differ} \\ 0 & \text{if } e \notin C. \end{cases} \tag{3}$$

- It is easy to see that $Q\chi_C = 0$ (for each vertex of the cycle we deliver and take out the same amount of product)

- 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

$$\chi_C(e) = \begin{cases} 1 & \text{if } e \in C \text{ and its orientations in } C \text{ and } G \text{ coincide} \\ -1 & \text{if } e \in C \text{ but its orientations in } C \text{ and } G \text{ differ} \\ 0 & \text{if } e \notin C. \end{cases} \tag{3}$$

- It is easy to see that $Q\chi_C = 0$ (for each vertex of the cycle we deliver and take out the same amount of product)
- The fact that the cycle space is spanned by the signed indicator functions of cycles is based on Linear Algebra and some results on spanning trees. Namely, we prove that the number of linearly independent signed indicator functions of cycles is equal to the dimension of the kernel of Q.

- The representation $\mathrm{TC}(G) = \ell_{1,d}(E(G)))/Z(G)$ implies much more than the polyhedrality of $\mathrm{TC}(G)$ and the fact that the dimension of the needed $\ell_{1,d}(E(G)))$ does not exceed $\frac{|V(G)|(|V(G)|-1)}{2}$.

- The representation $\mathrm{TC}(G) = \ell_{1,d}(E(G)))/Z(G)$ implies much more than the polyhedrality of $\mathrm{TC}(G)$ and the fact that the dimension of the needed $\ell_{1,d}(E(G)))$ does not exceed $\frac{|V(G)|(|V(G)|-1)}{2}$.
- The most important feature of the representation $\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G)$ is the explicit description of the kernel.

- The representation $\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G)$ implies much more than the polyhedrality of $\mathrm{TC}(G)$ and the fact that the dimension of the needed $\ell_{1,d}(E(G))$ does not exceed $\frac{|V(G)|(|V(G)|-1)}{2}$.
- The most important feature of the representation $\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G)$ is the explicit description of the kernel.
- There is a hope to use this description to find more information on $\mathrm{TC}(X)$.

- The representation $\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G)$ implies much more than the polyhedrality of $\mathrm{TC}(G)$ and the fact that the dimension of the needed $\ell_{1,d}(E(G))$ does not exceed $\frac{|V(G)|(|V(G)|-1)}{2}$.
- The most important feature of the representation $\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G)$ is the explicit description of the kernel.
- There is a hope to use this description to find more information on $\mathrm{TC}(X)$.
- This description was used to estimate the Banach-Mazur distance of TC to ℓ_1 for many graphs with large groups of cycle-preserving bijections (see [DKO20]).

- The representation $\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G)$ implies much more than the polyhedrality of $\mathrm{TC}(G)$ and the fact that the dimension of the needed $\ell_{1,d}(E(G))$ does not exceed $\frac{|V(G)|(|V(G)|-1)}{2}$.
- The most important feature of the representation $\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G)$ is the explicit description of the kernel.
- There is a hope to use this description to find more information on $\mathrm{TC}(X)$.
- This description was used to estimate the Banach-Mazur distance of TC to ℓ_1 for many graphs with large groups of cycle-preserving bijections (see [DKO20]).
- This study continued in [DKO21] and [DKO24].

What about geometry of cycle spaces?

- In connection with the discussed representation $\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G)$, it would be very important to study the Banach space structure of $Z(G)$ and its "location" in $\ell_{1,d}(E(G))$.

What about geometry of cycle spaces?

- In connection with the discussed representation $\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G)$, it would be very important to study the Banach space structure of $Z(G)$ and its "location" in $\ell_{1,d}(E(G))$.
- It would be very interesting to connect this study with the line of research due to Kislyakov (1976), Pisier (1978), Bourgain (1984), ..., devoted to dependence of the properties of L_1/Z on the geometry of Z.

What about geometry of cycle spaces?

- In connection with the discussed representation $\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G)$, it would be very important to study the Banach space structure of $Z(G)$ and its "location" in $\ell_{1,d}(E(G))$.
- It would be very interesting to connect this study with the line of research due to Kislyakov (1976), Pisier (1978), Bourgain (1984), ..., devoted to dependence of the properties of L_1/Z on the geometry of Z.
- The results of [DKO20, DKO21, DKO24] seem to go in a different direction.

What about geometry of cycle spaces?

- In connection with the discussed representation $\mathrm{TC}(G) = \ell_{1,d}(E(G)))/Z(G)$, it would be very important to study the Banach space structure of $Z(G)$ and its "location" in $\ell_{1,d}(E(G)))$.
- It would be very interesting to connect this study with the line of research due to Kislyakov (1976), Pisier (1978), Bourgain (1984), ..., devoted to dependence of the properties of L_1/Z on the geometry of Z.
- The results of [DKO20, DKO21, DKO24] seem to go in a different direction.
- The only result which provide some Banach-space-theoretical consequences of the desired type is the following result of P. Erdős, L. Pósa [EP62]: The number of edge disjoint cycles in a cycle space of dimension d is $\Omega\left(\frac{d}{\log d}\right)$ and it is the best possible estimate. Note that disjoint cycles in $\ell_{1,d}(E(G))$ span an isometric to ℓ_1^k 1-complemented subspace.

Comments

- The representation $\mathrm{TC}(G) = \ell_{1,d}(E(G)))/Z(G)$ implies that if a transportation plan for μ (described as an element of $\ell_{1,d}(E(G))))$ is not optimal, then we can add to it an element of the cycle space to get a cheaper plan.

Comments

- The representation $\mathrm{TC}(G) = \ell_{1,d}(E(G)))/Z(G)$ implies that if a transportation plan for μ (described as an element of $\ell_{1,d}(E(G)))$) is not optimal, then we can add to it an element of the cycle space to get a cheaper plan.
- Many years ago (1939), Tolstoi noticed that this statement can be strengthened to the following **Tolstoi type theorem**: if a transportation plan for μ (described as an element of $\ell_{1,d}(E(G)))$) is not optimal, then we can add to it a multiple of the signed indicator function of a cycle to get a cheaper plan.

Comments

- The representation $\mathrm{TC}(G) = \ell_{1,d}(E(G))/Z(G)$ implies that if a transportation plan for μ (described as an element of $\ell_{1,d}(E(G))$) is not optimal, then we can add to it an element of the cycle space to get a cheaper plan.
- Many years ago (1939), Tolstoi noticed that this statement can be strengthened to the following **Tolstoi type theorem**: if a transportation plan for μ (described as an element of $\ell_{1,d}(E(G))$) is not optimal, then we can add to it a multiple of the signed indicator function of a cycle to get a cheaper plan.
- This Tolstoi was an economist, and he never proved this theorem. Different people published proofs of some versions. It seems that the most suitable for Banach space theory Tolstoi type theorem was proved in [OO22].

Geometric meaning of Tolstoi type theorem

- Recall the statement: if a transportation plan for μ (described as an element of $\ell_{1,d}(E(G))$)) is not optimal, then we can subtract from it a multiple of the signed indicator function of a cycle to get a cheaper plan.

Geometric meaning of Tolstoi type theorem

- Recall the statement: if a transportation plan for μ (described as an element of $\ell_{1,d}(E(G))$) is not optimal, then we can subtract from it a multiple of the signed indicator function of a cycle to get a cheaper plan.
- We describe the plan as a set of edges with directions and quantities of the product which are to be moved along them.

Geometric meaning of Tolstoi type theorem

- Recall the statement: if a transportation plan for μ (described as an element of $\ell_{1,d}(E(G)))$ is not optimal, then we can subtract from it a multiple of the signed indicator function of a cycle to get a cheaper plan.
- We describe the plan as a set of edges with directions and quantities of the product which are to be moved along them.
- The geometric meaning is that the plan is not optimal if and only if there exists a directed cycle, for which the total length (weight) of edges which follow the transportation plan is more than half of the total length.

Optional: Another way of getting $\mathrm{TC}(X)$ as a quotient

- **Definition:** A point x in a nonempty convex set A of a vector space is called *extreme* if whenever $x = \alpha u + (1-\alpha)v$ with $\alpha \in (0,1)$ and with $u, v \in A$, then $x = u = v$.

Optional: Another way of getting $\mathrm{TC}(X)$ as a quotient

- **Definition:** A point x in a nonempty convex set A of a vector space is called *extreme* if whenever $x = \alpha u + (1-\alpha)v$ with $\alpha \in (0,1)$ and with $u, v \in A$, then $x = u = v$.
- **Minkowski theorem:** A compact convex set in $\mathbb{R}^n$ is a closed convex hull of its extreme points.

Optional: Another way of getting $\mathrm{TC}(X)$ as a quotient

- **Definition:** A point x in a nonempty convex set A of a vector space is called *extreme* if whenever $x = \alpha u + (1-\alpha)v$ with $\alpha \in (0,1)$ and with $u, v \in A$, then $x = u = v$.
- **Minkowski theorem:** A compact convex set in $\mathbb{R}^n$ is a closed convex hull of its extreme points.
- **Observation** If a finite-dimensional normed space Y is such that its unit ball has exactly $2m$ extreme points, then Y is a quotient of ℓ_1^m and Y is not a quotient of ℓ_1^k for $k < m$.

- **Proposition:** If X is a finite metric space, extreme points of the unit ball of $\mathrm{TC}(X)$ are transportation problems of the form

$$m_{u,v} := \frac{1}{d(u,v)}(\mathbf{1}_u - \mathbf{1}_v), \tag{4}$$

where $u, v \in X$ are such that uv is an edge in the canonical graph of (X, d), that is, there are no points $z \in X$, $z \notin \{u, v\}$ such that $d(u,z) + d(z,v) = d(u,v)$. (This proposition implies the desired estimate for the number of extreme points from above.)

- **Proposition:** If X is a finite metric space, extreme points of the unit ball of $\mathrm{TC}(X)$ are transportation problems of the form

$$m_{u,v} := \frac{1}{d(u,v)}(\mathbf{1}_u - \mathbf{1}_v), \tag{4}$$

  where $u, v \in X$ are such that uv is an edge in the canonical graph of (X, d), that is, there are no points $z \in X$, $z \notin \{u, v\}$ such that $d(u,z) + d(z,v) = d(u,v)$. (This proposition implies the desired estimate for the number of extreme points from above.)
- **Remark (Extreme points of $\mathrm{TC}(X)$ for infinite X)** The statement of Proposition remains true for infinite metric spaces (Aliaga, Pernecká, Smith [APS26+]). This result is much more difficult than its finite version.

Another geometric result on $\mathrm{TC}(M)$

- **Theorem (Khan, Mim, M.O., 2020):** If a metric space M contains $2n$ elements, then $\mathrm{TC}(M)$ contains a 1-complemented subspace isometric to ℓ_1^n. (1-*complemented* means that there exists a linear map from $\mathrm{TC}(M)$ onto this subspace which does not increase norms.)
  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}.

Another geometric result on $\mathrm{TC}(M)$

- **Theorem (Khan, Mim, M.O., 2020):** If a metric space M contains $2n$ elements, then $\mathrm{TC}(M)$ contains a 1-complemented subspace isometric to ℓ_1^n. (1-*complemented* means that there exists a linear map from $\mathrm{TC}(M)$ onto this subspace which does not increase norms.)
  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}.
- The most difficult part of this result is that the subspace is 1-complemented. We shall not prove it because it uses a rather lengthy argument based on the analysis of the Edmonds algorithm for finding a minimum weight perfect matching.

Another geometric result on $\mathrm{TC}(M)$

- **Theorem (Khan, Mim, M.O., 2020):** If a metric space M contains $2n$ elements, then $\mathrm{TC}(M)$ contains a 1-complemented subspace isometric to ℓ_1^n. (1-*complemented* means that there exists a linear map from $\mathrm{TC}(M)$ onto this subspace which does not increase norms.)
  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}.
- The most difficult part of this result is that the subspace is 1-complemented. We shall not prove it because it uses a rather lengthy argument based on the analysis of the Edmonds algorithm for finding a minimum weight perfect matching.
- I should note that an earlier result in [DKO20] proving that there exists a 2-complemented 2-isomorphic ℓ_1^n has a simpler proof.

Presence of isometric ℓ_1^n

- Consider M as a weighted graph K_{2n}. The weight of an edge is the distance between its ends. Pick in this graph a minimum weight perfect matching, that is partition M into a set of pairs $\{u_i, v_i\}_{i=1}^n$ such that the sum $\sum_{i=1}^n d(u_i, v_i)$ is the least possible.

Presence of isometric ℓ_1^n

- Consider M as a weighted graph K_{2n}. The weight of an edge is the distance between its ends. Pick in this graph a minimum weight perfect matching, that is partition M into a set of pairs $\{u_i, v_i\}_{i=1}^n$ such that the sum $\sum_{i=1}^n d(u_i, v_i)$ is the least possible.
- We claim that $f_i := \dfrac{\mathbf{1}_{u_i} - \mathbf{1}_{v_i}}{d(u_i, v_i)}$, $i = 1, \dots, n$ satisfy $\|\sum_{i=1}^n a_i f_i\|_{\mathrm{TC}} = \sum_{i=1}^n |a_i|$, so these vectors span a subspace isometric to ℓ_1^n.

Presence of isometric ℓ_1^n

- Consider M as a weighted graph K_{2n}. The weight of an edge is the distance between its ends. Pick in this graph a minimum weight perfect matching, that is partition M into a set of pairs $\{u_i, v_i\}_{i=1}^n$ such that the sum $\sum_{i=1}^n d(u_i, v_i)$ is the least possible.
- We claim that $f_i := \dfrac{\mathbf{1}_{u_i} - \mathbf{1}_{v_i}}{d(u_i, v_i)}, \ i = 1, \dots, n$ satisfy $\|\sum_{i=1}^n a_i f_i\|_{\mathrm{TC}} = \sum_{i=1}^n |a_i|$, so these vectors span a subspace isometric to ℓ_1^n.
- We need estimate the norm from below only (see next slide).

- Assume the contrary, then

$$\sum_{i=1}^{n} \frac{a_i}{d(u_i, v_i)} (\mathbf{1}_{u_i} - \mathbf{1}_{v_i})$$

is not an optimal transportation plan for $\sum_{i=1}^{n} a_i f_i$.

- Assume the contrary, then

$$\sum_{i=1}^{n} \frac{a_i}{d(u_i, v_i)} (\mathbf{1}_{u_i} - \mathbf{1}_{v_i})$$

  is not an optimal transportation plan for $\sum_{i=1}^{n} a_i f_i$.
- In the next step, the proof is more clear if we assume that $a_i > 0$ for $i = 1, \dots, n$ (we can re-denote some u_i and v_i in other cases).

- Assume the contrary, then

$$\sum_{i=1}^{n} \frac{a_i}{d(u_i, v_i)} (\mathbf{1}_{u_i} - \mathbf{1}_{v_i})$$

  is not an optimal transportation plan for $\sum_{i=1}^{n} a_i f_i$.
- In the next step, the proof is more clear if we assume that $a_i > 0$ for $i = 1, \ldots, n$ (we can re-denote some u_i and v_i in other cases).
- By the Tolstoi-type theorem, there exists a cycle which includes some of the edges $\overrightarrow{v_i u_i}$ in which the total weight of the edges $\{v_i u_i\}_{i=1}^{n}$ is strictly greater than the weight of the rest of the cycle.

- Assume the contrary, then

$$\sum_{i=1}^{n} \frac{a_i}{d(u_i, v_i)} (\mathbf{1}_{u_i} - \mathbf{1}_{v_i})$$

  is not an optimal transportation plan for $\sum_{i=1}^{n} a_i f_i$.
- In the next step, the proof is more clear if we assume that $a_i > 0$ for $i = 1, \ldots, n$ (we can re-denote some u_i and v_i in other cases).
- By the Tolstoi-type theorem, there exists a cycle which includes some of the edges $\overrightarrow{v_i u_i}$ in which the total weight of the edges $\{v_i u_i\}_{i=1}^{n}$ is strictly greater than the weight of the rest of the cycle.
- This implies that $\{u_i, v_i\}_{i=1}^{n}$ is not a minimum weight perfect matching (sketch), we get a contradiction.

No ℓ_1^{n+1} if all triangles are "strict"

- **Enough:** If all triangle inequalities in M are strict and $\{f_i\}_{i=1}^{n+1}$ are isometrically equivalent to the unit vector basis of ℓ_1^{n+1}, then $\{f_i\}$ have disjoint supports.

No ℓ_1^{n+1} if all triangles are "strict"

- **Enough:** If all triangle inequalities in M are strict and $\{f_i\}_{i=1}^{n+1}$ are isometrically equivalent to the unit vector basis of ℓ_1^{n+1}, then $\{f_i\}$ have disjoint supports.
- Proof. Assume the contrary, let $v \in M$ be in the supports of both f_i and f_j, $i \neq j$. Without loss of generality we assume that $f_i(v) > 0$ and $f_j(v) < 0$.

No ℓ_1^{n+1} if all triangles are "strict"

- **Enough:** If all triangle inequalities in M are strict and $\{f_i\}_{i=1}^{n+1}$ are isometrically equivalent to the unit vector basis of ℓ_1^{n+1}, then $\{f_i\}$ have disjoint supports.
- Proof. Assume the contrary, let $v \in M$ be in the supports of both f_i and f_j, $i \neq j$. Without loss of generality we assume that $f_i(v) > 0$ and $f_j(v) < 0$.
- To get a contradiction it suffices to show that $\|f_i + f_j\|_{\mathrm{TC}} < \|f_i\|_{\mathrm{TC}} + \|f_j\|_{\mathrm{TC}}$. This can be done in the following way. In an optimal plan for f_i some amount of units, denote it $\alpha > 0$, is moved from v to some $u \in M$. In an optimal plan for f_j some amount of units, denote it $\beta > 0$, is moved to v from some $w \in M$ (w can be the same as u).

No ℓ_1^{n+1} if all triangles are "strict"

- **Enough:** If all triangle inequalities in M are strict and $\{f_i\}_{i=1}^{n+1}$ are isometrically equivalent to the unit vector basis of ℓ_1^{n+1}, then $\{f_i\}$ have disjoint supports.
- Proof. Assume the contrary, let $v \in M$ be in the supports of both f_i and f_j, $i \neq j$. Without loss of generality we assume that $f_i(v) > 0$ and $f_j(v) < 0$.
- To get a contradiction it suffices to show that $\|f_i + f_j\|_{\mathrm{TC}} < \|f_i\|_{\mathrm{TC}} + \|f_j\|_{\mathrm{TC}}$. This can be done in the following way. In an optimal plan for f_i some amount of units, denote it $\alpha > 0$, is moved from v to some $u \in M$. In an optimal plan for f_j some amount of units, denote it $\beta > 0$, is moved to v from some $w \in M$ (w can be the same as u).
- Let $\gamma = \min\{\alpha, \beta\}$. Now we combine the optimal plans for f_i and f_j with the following exception: we move γ units of the product directly from w to u. Since, by our assumption, $d(w, u) < d(w, v) + d(v, u)$, the cost of the obtained plan is $< \|f_i\|_{\mathrm{TC}} + \|f_j\|_{\mathrm{TC}}$.

- So we know that $\mathrm{TC}(X)$ is a rather special space. (1) It is an $(|X|-1)$-dimensional quotient of $\ell_1^{|E(X)|}$, (2) It contains an isometric 1-complemented ℓ_1^n with $n \ge |X|/2$.

- So we know that $\mathrm{TC}(X)$ is a rather special space. (1) It is an $(|X|-1)$-dimensional quotient of $\ell_1^{|E(X)|}$, (2) It contains an isometric 1-complemented ℓ_1^n with $n \ge |X|/2$.
- Now we show that $\mathrm{TC}(X)$ has some amount of freedom.

- So we know that $\mathrm{TC}(X)$ is a rather special space. (1) It is an $(|X|-1)$-dimensional quotient of $\ell_1^{|E(X)|}$, (2) It contains an isometric 1-complemented ℓ_1^n with $n \ge |X|/2$.
- Now we show that $\mathrm{TC}(X)$ has some amount of freedom.
- In the proof below we need the following definition. Let (M, d) be a metric space and $\delta > 0$. A subset $N_\delta \subset M$ is called a δ-*net* in M if

- So we know that $\mathrm{TC}(X)$ is a rather special space. (1) It is an $(|X|-1)$-dimensional quotient of $\ell_1^{|E(X)|}$, (2) It contains an isometric 1-complemented ℓ_1^n with $n \geq |X|/2$.
- Now we show that $\mathrm{TC}(X)$ has some amount of freedom.
- In the proof below we need the following definition. Let (M, d) be a metric space and $\delta > 0$. A subset $N_\delta \subset M$ is called a δ-*net* in M if
  - $\forall x \in M \quad \exists n \in N_\delta \quad d(x, n) \leq \delta,$

- So we know that $\mathrm{TC}(X)$ is a rather special space. (1) It is an $(|X|-1)$-dimensional quotient of $\ell_1^{|E(X)|}$, (2) It contains an isometric 1-complemented ℓ_1^n with $n \ge |X|/2$.
- Now we show that $\mathrm{TC}(X)$ has some amount of freedom.
- In the proof below we need the following definition. Let (M, d) be a metric space and $\delta > 0$. A subset $N_\delta \subset M$ is called a δ-*net* in M if
  - $\forall x \in M \quad \exists n \in N_\delta \quad d(x, n) \le \delta$,
  - $\forall n, m \in N_\delta \quad d(n, m) \ge \delta$.

- **Theorem:** For every n-dimensional Banach space B and any $\varepsilon > 0$, there is $N = N(n, \varepsilon)$ and a metric space X with $|X| = N$ such that there is linear embedding $T : B \to \mathrm{TC}(X)$ satisfying $\forall x \in B \quad \|x\| \le \|Tx\| \le (1 + \varepsilon)\|x\|$. (*)

- **Theorem:** For every n-dimensional Banach space B and any $\varepsilon > 0$, there is $N = N(n, \varepsilon)$ and a metric space X with $|X| = N$ such that there is linear embedding $T : B \to \mathrm{TC}(X)$ satisfying $\forall x \in B \quad \|x\| \le \|Tx\| \le (1+\varepsilon)\|x\|$. (*)
- Two things are needed: (1) observation of Arens-Eells [AE56]: for a pointed metric space (X, O) the mapping $x \mapsto \mathbf{1}_x - \mathbf{1}_O$ is an isometric embedding of X into $\mathrm{TC}(X)$.

- **Theorem:** For every n-dimensional Banach space B and any $\varepsilon > 0$, there is $N = N(n, \varepsilon)$ and a metric space X with $|X| = N$ such that there is linear embedding $T : B \to \mathrm{TC}(X)$ satisfying $\forall x \in B \quad \|x\| \le \|Tx\| \le (1 + \varepsilon)\|x\|$. (*)
- Two things are needed: (1) observation of Arens-Eells [AE56]: for a pointed metric space (X, O) the mapping $x \mapsto \mathbf{1}_x - \mathbf{1}_O$ is an isometric embedding of X into $\mathrm{TC}(X)$.
- (2) Deep result of Bourgain [Bou87]: For every n and $\varepsilon > 0$, there is $\delta > 0$ such that for every n-dimensional Banach space X and for any Banach space Y we have: if there is an isometric embedding of a δ-net N_δ of the unit ball of B into a Banach space Y, then there is an embedding $T : B \to Y$ satisfying (*)

- **Theorem:** For every n-dimensional Banach space B and any $\varepsilon > 0$, there is $N = N(n, \varepsilon)$ and a metric space X with $|X| = N$ such that there is linear embedding $T : B \to \mathrm{TC}(X)$ satisfying $\forall x \in B \quad \|x\| \le \|Tx\| \le (1+\varepsilon)\|x\|$. (*)
- Two things are needed: (1) observation of Arens-Eells [AE56]: for a pointed metric space (X, O) the mapping $x \mapsto \mathbf{1}_x - \mathbf{1}_O$ is an isometric embedding of X into $\mathrm{TC}(X)$.
- (2) Deep result of Bourgain [Bou87]: For every n and $\varepsilon > 0$, there is $\delta > 0$ such that for every n-dimensional Banach space X and for any Banach space Y we have: if there is an isometric embedding of a δ-net N_δ of the unit ball of B into a Banach space Y, then there is an embedding $T : B \to Y$ satisfying (*)
- To complete the proof, we denote by $N(n, \varepsilon)$ the estimate for $|N_\delta|$, and let $X = N_\delta$.

Earth Mover Distance

- In computer science, $\|\mu - \nu\|_{\mathrm{TC}}$, where μ and ν are two positive measures of the same total mass, has ample applications in comparison of complex data objects. For instance, it has been recognized to capture well the differences between images.

Earth Mover Distance

- In computer science, $\|\mu - \nu\|_{\mathrm{TC}}$, where μ and ν are two positive measures of the same total mass, has ample applications in comparison of complex data objects. For instance, it has been recognized to capture well the differences between images.
- Note that in Computer Science, $\|\mu - \nu\|_{\mathrm{TC}}$ is usually denoted $\mathrm{EMD}(\mu, \nu)$, and is called the *Earth Mover Distance*.

Earth Mover Distance

- In computer science, $\|\mu - \nu\|_{\mathrm{TC}}$, where μ and ν are two positive measures of the same total mass, has ample applications in comparison of complex data objects. For instance, it has been recognized to capture well the differences between images.
- Note that in Computer Science, $\|\mu - \nu\|_{\mathrm{TC}}$ is usually denoted $\mathrm{EMD}(\mu, \nu)$, and is called the *Earth Mover Distance*.
- The term *Earth Mover Distance* was introduced by experts in computer vision who recalled that the first known mathematical publication on a transportation problem was about soil transportation (Monge 1781)

Earth Mover Distance

- In computer science, $\|\mu - \nu\|_{\mathrm{TC}}$, where μ and ν are two positive measures of the same total mass, has ample applications in comparison of complex data objects. For instance, it has been recognized to capture well the differences between images.
- Note that in Computer Science, $\|\mu - \nu\|_{\mathrm{TC}}$ is usually denoted $\mathrm{EMD}(\mu, \nu)$, and is called the *Earth Mover Distance*.
- The term *Earth Mover Distance* was introduced by experts in computer vision who recalled that the first known mathematical publication on a transportation problem was about soil transportation (Monge 1781)
- EMD is also applied for comparisons between text documents.

Earth Mover Distance

- In computer science, $\|\mu - \nu\|_{\mathrm{TC}}$, where μ and ν are two positive measures of the same total mass, has ample applications in comparison of complex data objects. For instance, it has been recognized to capture well the differences between images.
- Note that in Computer Science, $\|\mu - \nu\|_{\mathrm{TC}}$ is usually denoted $\mathrm{EMD}(\mu, \nu)$, and is called the *Earth Mover Distance*.
- The term *Earth Mover Distance* was introduced by experts in computer vision who recalled that the first known mathematical publication on a transportation problem was about soil transportation (Monge 1781)
- EMD is also applied for comparisons between text documents.
- EMD in high-dimensional spaces is often used to compare documents and even two-dimensional color images.

Complexity of TC and EMD computations

- Exact computation of $\|\mu - \nu\|_{\mathrm{TC}}$ is based on Linear Programming.

Complexity of TC and EMD computations

- Exact computation of $\|\mu - \nu\|_{\mathrm{TC}}$ is based on Linear Programming.
- Let $\{u_i\}_{i=1}^{s} = \{v : \ \mu(\{v\}) > 0\}$ and $\{v_j\}_{j=1}^{t}$ $= \{v : \ \mu(\{v\}) < 0\}$. We consider only transportation plans of the form $\sum_{i=1}^{s}\sum_{j=1}^{t} p_{ij}(\mathbf{1}_{u_i} - \mathbf{1}_{v_j})$, where by p_{ij} $(p_{ij} \geq 0)$ we denote the amount which we plan to deliver from u_i to v_j.

Complexity of TC and EMD computations

- Exact computation of $\|\mu - \nu\|_{\mathrm{TC}}$ is based on Linear Programming.
- Let $\{u_i\}_{i=1}^s = \{v : \mu(\{v\}) > 0\}$ and $\{v_j\}_{j=1}^t = \{v : \mu(\{v\}) < 0\}$. We consider only transportation plans of the form $\sum_{i=1}^s \sum_{j=1}^t p_{ij}(\mathbf{1}_{u_i} - \mathbf{1}_{v_j})$, where by p_{ij} $(p_{ij} \geq 0)$ we denote the amount which we plan to deliver from u_i to v_j.
- The cost of such a transportation plan is

$$\sum_{i=1}^{s} \sum_{j=1}^{t} p_{ij} d(u_i, v_j). \tag{5}$$

Complexity of TC and EMD computations

- Exact computation of $\|\mu - \nu\|_{\mathrm{TC}}$ is based on Linear Programming.
- Let $\{u_i\}_{i=1}^s = \{v : \ \mu(\{v\}) > 0\}$ and $\{v_j\}_{j=1}^t = \{v : \ \mu(\{v\}) < 0\}$. We consider only transportation plans of the form $\sum_{i=1}^s \sum_{j=1}^t p_{ij}(\mathbf{1}_{u_i} - \mathbf{1}_{v_j})$, where by p_{ij} ($p_{ij} \geq 0$) we denote the amount which we plan to deliver from u_i to v_j.
- The cost of such a transportation plan is

$$\sum_{i=1}^{s} \sum_{j=1}^{t} p_{ij} d(u_i, v_j). \tag{5}$$

- Also the plan has to satisfy the restrictions $\sum_{i=1}^s p_{ij} = -\mu(v_j)$, $\sum_{j=1}^t p_{ij} = \mu(u_i)$, and $p_{ij} \geq 0$.

Complexity of TC and EMD computations

- Exact computation of $\|\mu - \nu\|_{\mathrm{TC}}$ is based on Linear Programming.
- Let $\{u_i\}_{i=1}^s = \{v : \mu(\{v\}) > 0\}$ and $\{v_j\}_{j=1}^t = \{v : \mu(\{v\}) < 0\}$. We consider only transportation plans of the form $\sum_{i=1}^s \sum_{j=1}^t p_{ij}(\mathbf{1}_{u_i} - \mathbf{1}_{v_j})$, where by p_{ij} ($p_{ij} \geq 0$) we denote the amount which we plan to deliver from u_i to v_j.
- The cost of such a transportation plan is

$$\sum_{i=1}^{s} \sum_{j=1}^{t} p_{ij} d(u_i, v_j). \tag{5}$$

- Also the plan has to satisfy the restrictions $\sum_{i=1}^s p_{ij} = -\mu(v_j)$, $\sum_{j=1}^t p_{ij} = \mu(u_i)$, and $p_{ij} \geq 0$.
- So, we need to solve this linear programming problem.

- LP problems can be solved reasonably fast. The simplex method, although in general is not polynomial, usually is fast.

- LP problems can be solved reasonably fast. The simplex method, although in general is not polynomial, usually is fast.
- There also exist polynomial algorithms. The most well-known are Khachiyan's and Karmarkar's. The second seems to be reasonably fast. There are many newer methods which are worth looking at.

- LP problems can be solved reasonably fast. The simplex method, although in general is not polynomial, usually is fast.
- There also exist polynomial algorithms. The most well-known are Khachiyan's and Karmarkar's. The second seems to be reasonably fast. There are many newer methods which are worth looking at.
- Nevertheless, for Data Science, where people need quickly identify images that are closest to the query image, LP is not fast enough, because the search requires to compare many thousands of images fast.

- LP problems can be solved reasonably fast. The simplex method, although in general is not polynomial, usually is fast.
- There also exist polynomial algorithms. The most well-known are Khachiyan's and Karmarkar's. The second seems to be reasonably fast. There are many newer methods which are worth looking at.
- Nevertheless, for Data Science, where people need quickly identify images that are closest to the query image, LP is not fast enough, because the search requires to compare many thousands of images fast.
- Charikar [Cha02] and Indyk-Thaper [IT03] suggested to do the search using low-distortion embeddings into ℓ_1^d (of a suitable dimension), where fast nearest neighbor algorithms are known.

- LP problems can be solved reasonably fast. The simplex method, although in general is not polynomial, usually is fast.
- There also exist polynomial algorithms. The most well-known are Khachiyan's and Karmarkar's. The second seems to be reasonably fast. There are many newer methods which are worth looking at.
- Nevertheless, for Data Science, where people need quickly identify images that are closest to the query image, LP is not fast enough, because the search requires to compare many thousands of images fast.
- Charikar [Cha02] and Indyk-Thaper [IT03] suggested to do the search using low-distortion embeddings into ℓ_1^d (of a suitable dimension), where fast nearest neighbor algorithms are known.
- It makes solutions approximate, which is OK for many applications, especially when it is known that the exact solution takes much more time.

The problem of L_1-embeddings of $\mathrm{TC}(X)$

- The L_1-**distortion** $c_1(Z)$ of a metric space (Z, d) is the **infimal** D for which
$$d(x,y) \le \|f(x) - f(y)\|_1 \le Dd(x,y)$$
for some $d \in \mathbb{N}$ and $f : Z \to \ell_1^d$.

The problem of L_1-embeddings of $\mathrm{TC}(X)$

- The L_1-**distortion** $c_1(Z)$ of a metric space (Z, d) is the **infimal** D for which
$$d(x,y) \le \|f(x) - f(y)\|_1 \le Dd(x,y)$$
for some $d \in \mathbb{N}$ and $f : Z \to \ell_1^d$.
- One should mention that low-distortion embeddings of $\mathrm{TC}(X)$ into "better" spaces, like ℓ_2 or ℓ_p $(1 < p \le 2)$ would be even more helpful.

The problem of L_1-embeddings of $\mathrm{TC}(X)$

- The L_1-**distortion** $c_1(Z)$ of a metric space (Z, d) is the **infimal** D for which
$$d(x,y) \le \|f(x) - f(y)\|_1 \le Dd(x,y)$$
for some $d \in \mathbb{N}$ and $f : Z \to \ell_1^d$.
- One should mention that low-distortion embeddings of $\mathrm{TC}(X)$ into "better" spaces, like ℓ_2 or ℓ_p $(1 < p \le 2)$ would be even more helpful.
- However, the result on $|X|/2$-dimensional isometric of ℓ_1^d (which we have studied), combined with the classical results on embeddings of ℓ_1^d into $\ell_p, (1 < p \le 2)$ implies that embeddings of $\mathrm{TC}(X)$ into ℓ_p have distortion $\Omega(|X|^{1-\frac{1}{p}})$, which is too large for the intended applications.

The problem of L_1-embeddings of $\mathrm{TC}(X)$

- The L_1-**distortion** $c_1(Z)$ of a metric space (Z, d) is the **infimal** D for which
$$d(x,y) \leq \|f(x) - f(y)\|_1 \leq Dd(x,y)$$
for some $d \in \mathbb{N}$ and $f : Z \to \ell_1^d$.
- One should mention that low-distortion embeddings of $\mathrm{TC}(X)$ into "better" spaces, like ℓ_2 or ℓ_p $(1 < p \leq 2)$ would be even more helpful.
- However, the result on $|X|/2$-dimensional isometric of ℓ_1^d (which we have studied), combined with the classical results on embeddings of ℓ_1^d into $\ell_p, (1 < p \leq 2)$ implies that embeddings of $\mathrm{TC}(X)$ into ℓ_p have distortion $\Omega(|X|^{1-\frac{1}{p}})$, which is too large for the intended applications.
- **Consequence.** This activated the study of embeddings:
$$\mathrm{TC}(X),\ \mathrm{EMD}(X) \hookrightarrow \ell_1^d$$

What is known on L_1-distortions

- We shall use the following general fact for finite X [NS07]: $c_1(\mathrm{TC}(X)) = c_1(\mathrm{EMD}(X))$, and for $\mathrm{TC}(X)$ we may restrict our attention to linear maps f. **Note:** This result is proved in [NS07] is a slightly weaker form and only for planar grids. Modifications needed to a general case were added in [GO26].

What is known on L_1-distortions

- We shall use the following general fact for finite X [NS07]: $c_1(\mathrm{TC}(X)) = c_1(\mathrm{EMD}(X))$, and for $\mathrm{TC}(X)$ we may restrict our attention to linear maps f. **Note:** This result is proved in [NS07] is a slightly weaker form and only for planar grids. Modifications needed to a general case were added in [GO26].
- **Approximation by trees** ([AKPW95], [Bar96, Bar98], [Cha02], [FRT04])

What is known on L_1-distortions

- We shall use the following general fact for finite X [NS07]: $c_1(\mathrm{TC}(X)) = c_1(\mathrm{EMD}(X))$, and for $\mathrm{TC}(X)$ we may restrict our attention to linear maps f. **Note:** This result is proved in [NS07] is a slightly weaker form and only for planar grids. Modifications needed to a general case were added in [GO26].
- **Approximation by trees** ([AKPW95], [Bar96, Bar98], [Cha02], [FRT04])
- A key method for L_1-embeddability of $\mathrm{TC}(X)$ is to approximate the metric space (X, d) by random dominating tree metrics.

- **Definition:** We say that a finite metric space (X, d) is *embeds into a convex combination of dominating tree metrics with distortion* D or *is probabilistically (stochastically) approximated by dominating tree metrics with distortion* D if there is a finite collection of weighted trees $\{(T_i, d_i)\}_{i=1}^{m}$, such that

- **Definition:** We say that a finite metric space (X, d) is *embeds into a convex combination of dominating tree metrics with distortion D* or *is probabilistically (stochastically) approximated by dominating tree metrics with distortion D* if there is a finite collection of weighted trees $\{(T_i, d_i)\}_{i=1}^m$, such that
  - for each i, the vertex set of X is a subset of the vertex set of T_i, and the following conditions are satisfied:

- **Definition:** We say that a finite metric space (X, d) is *embeds into a convex combination of dominating tree metrics with distortion* D or *is probabilistically (stochastically) approximated by dominating tree metrics with distortion* D if there is a finite collection of weighted trees $\{(T_i, d_i)\}_{i=1}^m$, such that
    - for each i, the vertex set of X is a subset of the vertex set of T_i, and the following conditions are satisfied:
    - $\forall u, v \in X \quad d_i(u, v) \geq d(u, v)$. We describe this condition as d_i *is dominating* for d.

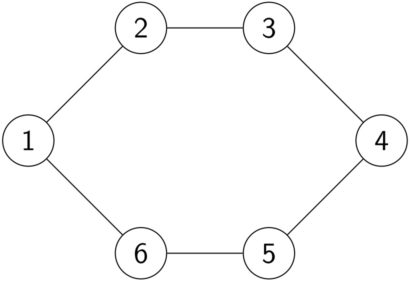
- **Definition:** We say that a finite metric space (X, d) is *embeds into a convex combination of dominating tree metrics with distortion D* or *is probabilistically (stochastically) approximated by dominating tree metrics with distortion D* if there is a finite collection of weighted trees $\{(T_i, d_i)\}_{i=1}^m$, such that
  - for each i, the vertex set of X is a subset of the vertex set of T_i, and the following conditions are satisfied:
  - $\forall u, v \in X \quad d_i(u, v) \geq d(u, v)$. We describe this condition as d_i *is dominating* for d.
  - There is a collection of positive numbers $\{\alpha_i\}$ with $\sum_{i=1}^m \alpha_i = 1$ such that $\forall u, v \in X \quad \sum_{i=1}^m \alpha_i d_i(u, v) \leq Dd(u, v)$.

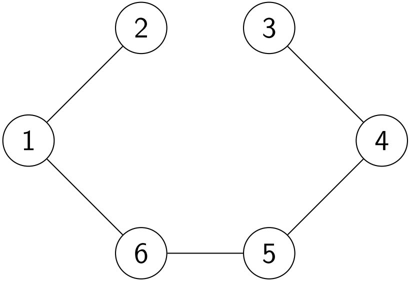
This direction started with Karp's observation

- The first nontrivial example of this type (due to Karp) is a cycle C_n (with an arbitrarily large n) with vertices $\{1, \dots, n\}$ and its graph distance. It can be approximated by the trees $\{T_i\}_{i=1}^n$, each of which is obtained by deleting one edge in the cycle. Below you see C_6 and its T_2.

This direction started with Karp's observation

- The first nontrivial example of this type (due to Karp) is a cycle C_n (with an arbitrarily large n) with vertices $\{1, \dots, n\}$ and its graph distance. It can be approximated by the trees $\{T_i\}_{i=1}^n$, each of which is obtained by deleting one edge in the cycle. Below you see C_6 and its T_2.

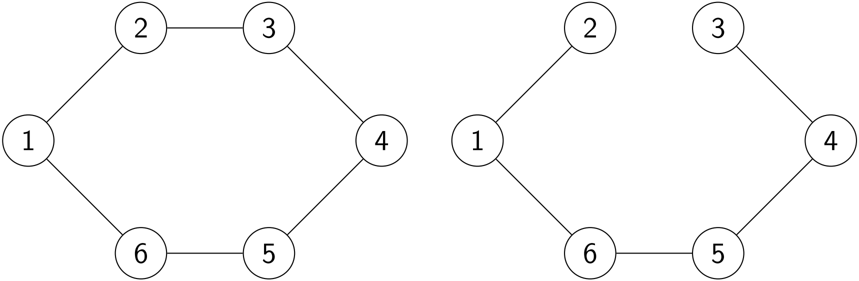




-

This direction started with Karp's observation

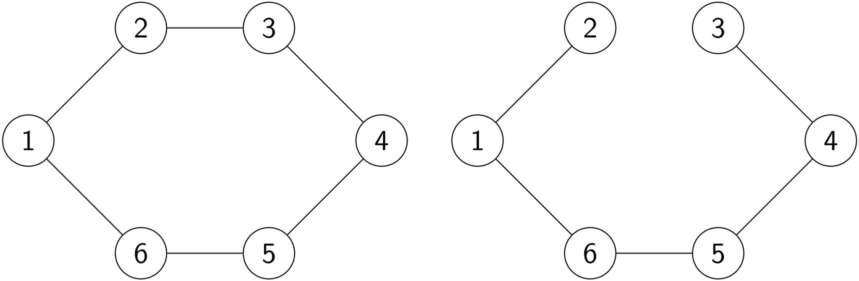
- The first nontrivial example of this type (due to Karp) is a cycle C_n (with an arbitrarily large n) with vertices $\{1, \ldots, n\}$ and its graph distance. It can be approximated by the trees $\{T_i\}_{i=1}^n$, each of which is obtained by deleting one edge in the cycle. Below you see C_6 and its T_2.



-
- It is trivial that $d_{C_n}(u, v) \leq d_{T_i}(u, v)$.

This direction started with Karp's observation

- The first nontrivial example of this type (due to Karp) is a cycle C_n (with an arbitrarily large n) with vertices $\{1, \dots, n\}$ and its graph distance. It can be approximated by the trees $\{T_i\}_{i=1}^n$, each of which is obtained by deleting one edge in the cycle. Below you see C_6 and its T_2.



-
- It is trivial that $d_{C_n}(u, v) \leq d_{T_i}(u, v)$.
- And easy that $\forall u, v \in C_n \quad \sum_{i=1}^n \frac{1}{n} d_{T_i}(u, v) \leq 2d_{C_n}(u, v)$.

- Approximation by dominating trees was introduced by Alon, Karp, Peleg, West [AKPW95] for spanning trees (which are always dominating) and by Bartal [Bar96] for general dominating trees.

- Approximation by dominating trees was introduced by Alon, Karp, Peleg, West [AKPW95] for spanning trees (which are always dominating) and by Bartal [Bar96] for general dominating trees.
- These authors, especially Bartal, initiated systematic usage of such approximations in the theory of Approximation Algorithms.

- Approximation by dominating trees was introduced by Alon, Karp, Peleg, West [AKPW95] for spanning trees (which are always dominating) and by Bartal [Bar96] for general dominating trees.
- These authors, especially Bartal, initiated systematic usage of such approximations in the theory of Approximation Algorithms.
- An important step in this development was the following result of Fakcharoenphol, Rao, Talwar [FRT04] (Bartal [Bar96, Bar98] proved slightly weaker estimates).

- Approximation by dominating trees was introduced by Alon, Karp, Peleg, West [AKPW95] for spanning trees (which are always dominating) and by Bartal [Bar96] for general dominating trees.
- These authors, especially Bartal, initiated systematic usage of such approximations in the theory of Approximation Algorithms.
- An important step in this development was the following result of Fakcharoenphol, Rao, Talwar [FRT04] (Bartal [Bar96, Bar98] proved slightly weaker estimates).
- **Theorem ([FRT04]).** There exists an absolute constant $0 < C < \infty$ such that any n-point metric space can be probabilistically approximated by dominated tree metrics with distortion $\leq C \log n$.

- Approximation by dominating trees was introduced by Alon, Karp, Peleg, West [AKPW95] for spanning trees (which are always dominating) and by Bartal [Bar96] for general dominating trees.
- These authors, especially Bartal, initiated systematic usage of such approximations in the theory of Approximation Algorithms.
- An important step in this development was the following result of Fakcharoenphol, Rao, Talwar [FRT04] (Bartal [Bar96, Bar98] proved slightly weaker estimates).
- **Theorem ([FRT04]).** There exists an absolute constant $0 < C < \infty$ such that any n-point metric space can be probabilistically approximated by dominated tree metrics with distortion $\leq C \log n$.
- The strongest similar result for spanning trees is $\leq C \log n (\log \log n)^2$ [ABN08].

- Approximation by dominating trees was introduced by Alon, Karp, Peleg, West [AKPW95] for spanning trees (which are always dominating) and by Bartal [Bar96] for general dominating trees.
- These authors, especially Bartal, initiated systematic usage of such approximations in the theory of Approximation Algorithms.
- An important step in this development was the following result of Fakcharoenphol, Rao, Talwar [FRT04] (Bartal [Bar96, Bar98] proved slightly weaker estimates).
- **Theorem ([FRT04]).** There exists an absolute constant $0 < C < \infty$ such that any n-point metric space can be probabilistically approximated by dominated tree metrics with distortion $\leq C \log n$.
- The strongest similar result for spanning trees is $\leq C \log n (\log \log n)^2$ [ABN08].
- One can show that constant degree expanders provide the lower bound $c \log n$ in both cases. For spanning trees there is still a gap between upper and lower bounds.

- Approximation by dominating trees is connected to our goal - estimating $c_1(\mathrm{TC}(X))$ through the following result:

- Approximation by dominating trees is connected to our goal - estimating $c_1(\mathrm{TC}(X))$ through the following result:
- **Theorem:** If a finite metric space X is probabilistically approximated by dominating tree metrics with distortion $\le D$, then $\mathrm{TC}(X)$ admits a linear embedding into L_1 whose distortion does not exceed D.

- Approximation by dominating trees is connected to our goal - estimating $c_1(\mathrm{TC}(X))$ through the following result:
- **Theorem:** If a finite metric space X is probabilistically approximated by dominating tree metrics with distortion $\le D$, then $\mathrm{TC}(X)$ admits a linear embedding into L_1 whose distortion does not exceed D.
- Charikar [Cha02] was the first to observe this result (its proof is rather straightforward). Since Charikar did not state and explain it clearly, several other people proved and published this theorem somewhat later.

- Our sketch of the proof that for a probabilistically D-approximated by dominating tree metrics X, there is a linear embedding $\mathrm{TC}(X) \mapsto L_1$ whose distortion does not exceed D:

- Our sketch of the proof that for a probabilistically D-approximated by dominating tree metrics X, there is a linear embedding $\mathrm{TC}(X) \mapsto L_1$ whose distortion does not exceed D:
- The condition implies the existence of a set $\{(T_i, d_i)\}_{i=1}^m$ of weighted trees, such that for each i X onto a subset of vertices of T_i, such that for all $u, v \in X$, $d_i(u,v) \geq d(u,v)$, and there is a collection of positive numbers $\sum_{i=1}^m \alpha_i = 1$ such that $\forall u, v$ $\sum_{i=1}^m \alpha_i d_i(u,v) \leq Dd(u,v)$.

- Our sketch of the proof that for a probabilistically D-approximated by dominating tree metrics X, there is a linear embedding $\mathrm{TC}(X) \mapsto L_1$ whose distortion does not exceed D:
- The condition implies the existence of a set $\{(T_i, d_i)\}_{i=1}^m$ of weighted trees, such that for each i X onto a subset of vertices of T_i, such that for all $u, v \in X$, $d_i(u, v) \geq d(u, v)$, and there is a collection of positive numbers $\sum_{i=1}^m \alpha_i = 1$ such that $\forall u, v$ $\sum_{i=1}^m \alpha_i d_i(u, v) \leq Dd(u, v)$.
- Let $\mu \in \mathrm{TC}(X, d)$. For each $i = 1, \ldots, m$, we introduce $\mu_i \in \mathrm{TC}(T_i, d_i)$ as a transportation problem with the same support and the same values as μ, but now this support is considered as a subset of the vertex set of T_i.

- Our sketch of the proof that for a probabilistically D-approximated by dominating tree metrics X, there is a linear embedding $\mathrm{TC}(X) \mapsto L_1$ whose distortion does not exceed D:
- The condition implies the existence of a set $\{(T_i, d_i)\}_{i=1}^m$ of weighted trees, such that for each i X onto a subset of vertices of T_i, such that for all $u, v \in X$, $d_i(u, v) \geq d(u, v)$, and there is a collection of positive numbers $\sum_{i=1}^m \alpha_i = 1$ such that $\forall u, v$ $\sum_{i=1}^m \alpha_i d_i(u, v) \leq Dd(u, v)$.
- Let $\mu \in \mathrm{TC}(X, d)$. For each $i = 1, \ldots, m$, we introduce $\mu_i \in \mathrm{TC}(T_i, d_i)$ as a transportation problem with the same support and the same values as μ, but now this support is considered as a subset of the vertex set of T_i.
- Let $F : \mathrm{TC}(X, d) \to (\oplus_{i=1}^m \mathrm{TC}(T_i, d_i))_1$ be given by $F\mu = \{\alpha_i \mu_i\}_{i=1}^m$. Earlier exercise: each $\mathrm{TC}(T_i, d_i)$, and thus the sum $(\oplus_{i=1}^m \mathrm{TC}(T_i, d_i))_1$, is isometric to ℓ_1^d of the corresponding dimension. Verification of the inequality $\|\mu\| \leq \|F\mu\| \leq D\|\mu\|$ is straightforward (summation order).

- We sketched the proof of $c_1(\mathrm{TC}(X)) = O(\log n)$ for any n-element metric space X.

- We sketched the proof of $c_1(\mathrm{TC}(X)) = O(\log n)$ for any n-element metric space X.
- It is easy to see that expanders achieve the lower bound $\Omega(\log n)$ for L_1-distortion, so, in general, the estimate $O(\log n)$ cannot be improved. Note, that this result follows from the fact that the distortion of embedding expanders into L_1 is $\Omega(\log n)$.

- We sketched the proof of $c_1(\mathrm{TC}(X)) = O(\log n)$ for any n-element metric space X.
- It is easy to see that expanders achieve the lower bound $\Omega(\log n)$ for L_1-distortion, so, in general, the estimate $O(\log n)$ cannot be improved. Note, that this result follows from the fact that the distortion of embedding expanders into L_1 is $\Omega(\log n)$.
- It became natural to investigate whether the upper estimate $O(\log n)$ can be improved for common classes of finite metric spaces **known to admit low-distortion embeddings into L_1**.

- We sketched the proof of $c_1(\mathrm{TC}(X)) = O(\log n)$ for any n-element metric space X.
- It is easy to see that expanders achieve the lower bound $\Omega(\log n)$ for L_1-distortion, so, in general, the estimate $O(\log n)$ cannot be improved. Note, that this result follows from the fact that the distortion of embedding expanders into L_1 is $\Omega(\log n)$.
- It became natural to investigate whether the upper estimate $O(\log n)$ can be improved for common classes of finite metric spaces **known to admit low-distortion embeddings into L_1**.
- It is known, Alon–Karp–Peleg–West [AKPW95], Konjevod–Ravi–Salman [KRS01], Gupta–Newman, Rabinovich–Sinclair [GNRS04], Schlumprecht-Tresch [ST24] that the distortion of approximation by convex combination of dominating trees has the maximal possible order $\log n$ for many relatively simple spaces X: planar grids, diamond graphs admitting low distortion embeddings into L_1.

- We sketched the proof of $c_1(\mathrm{TC}(X)) = O(\log n)$ for any n-element metric space X.
- It is easy to see that expanders achieve the lower bound $\Omega(\log n)$ for L_1-distortion, so, in general, the estimate $O(\log n)$ cannot be improved. Note, that this result follows from the fact that the distortion of embedding expanders into L_1 is $\Omega(\log n)$.
- It became natural to investigate whether the upper estimate $O(\log n)$ can be improved for common classes of finite metric spaces **known to admit low-distortion embeddings into L_1.**
- It is known, Alon–Karp–Peleg–West [AKPW95], Konjevod–Ravi–Salman [KRS01], Gupta–Newman, Rabinovich–Sinclair [GNRS04], Schlumprecht-Tresch [ST24] that the distortion of approximation by convex combination of dominating trees has the maximal possible order $\log n$ for many relatively simple spaces X: planar grids, diamond graphs admitting low distortion embeddings into L_1.
- It became important to know: what is the order of $c_1(X)$ for these spaces?

- Bourgain [Bou86] for Hamming cubes. The most difficult. It is based on presence of high-dimensional ℓ_∞^k in $\mathrm{TC}(X)$. Could be applicable to other spaces listed below, but I would expect that these proofs would be harder than the existing.

Methods of proving $c_1(\mathrm{TC}(X)) = \Omega(\log |X|)$

- Bourgain [Bou86] for Hamming cubes. The most difficult. It is based on presence of high-dimensional ℓ_∞^k in $\mathrm{TC}(X)$. Could be applicable to other spaces listed below, but I would expect that these proofs would be harder than the existing.
- Khot-Naor [KN06] for Hamming cubes. The most elegant. Harmonic analysis based. Difficult to transfer to other X.

Methods of proving $c_1(\mathrm{TC}(X)) = \Omega(\log |X|)$

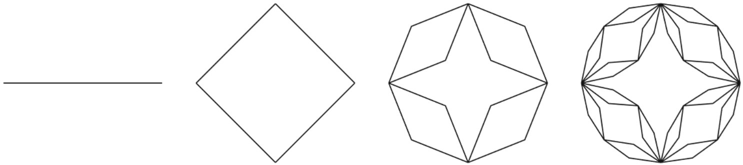
- Bourgain [Bou86] for Hamming cubes. The most difficult. It is based on presence of high-dimensional ℓ_∞^k in $\mathrm{TC}(X)$. Could be applicable to other spaces listed below, but I would expect that these proofs would be harder than the existing.
- Khot-Naor [KN06] for Hamming cubes. The most elegant. Harmonic analysis based. Difficult to transfer to other X.
- Gartland-M. O. [GO26] for planar grids and many slash powers (diamond graphs, Laakso graphs, $\ldots$). Based on customized mixed Sobolev inequalities. It is relatively easy to imagine its applications to other spaces. (Supersedes earlier results of [NS07] and [BGS23].)

Methods of proving $c_1(\mathrm{TC}(X)) = \Omega(\log |X|)$

- Bourgain [Bou86] for Hamming cubes. The most difficult. It is based on presence of high-dimensional ℓ_∞^k in $\mathrm{TC}(X)$. Could be applicable to other spaces listed below, but I would expect that these proofs would be harder than the existing.
- Khot-Naor [KN06] for Hamming cubes. The most elegant. Harmonic analysis based. Difficult to transfer to other X.
- Gartland-M. O. [GO26] for planar grids and many slash powers (diamond graphs, Laakso graphs, ...). Based on customized mixed Sobolev inequalities. It is relatively easy to imagine its applications to other spaces. (Supersedes earlier results of [NS07] and [BGS23].)
- Gartland-M. O.-Rabani-Young [GORY26] for high-dimensional grids. Based on Poincaré inequalities for random measures on grids. Could be transferred to other discrete structures by people who could understand how to handle random measures there.

- The rest of this lecture will be devoted to a description of the method of Gartland-M. O. [GO26] in the easiest for this method case - diamond graphs.

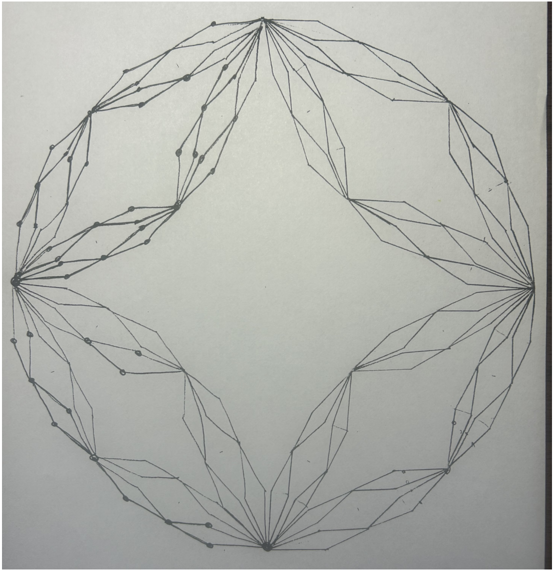
- The rest of this lecture will be devoted to a description of the method of Gartland-M. O. [GO26] in the easiest for this method case - diamond graphs.
- Diamond graphs $\{D_n\}_{n=0}^{\infty}$ are defined recursively: The *diamond graph* of level 0 has two vertices joined by an edge of length 1 and is denoted by D_0. The *diamond graph* D_n is obtained from D_{n-1} in the following way. Each edge $uv \in E(D_{n-1})$, it is replaced by a quadrilateral u, a, v, b, with edges ua, av, vb, bu.

- The rest of this lecture will be devoted to a description of the method of Gartland-M. O. [GO26] in the easiest for this method case - diamond graphs.
- 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. Each edge $uv \in E(D_{n-1})$, it is replaced by a quadrilateral u, a, v, b, with edges ua, av, vb, bu.
- 

D_0 D_1 D_2 D_3

Figure: The first four diamond graphs

Diamond D_4



Proof description

- According to the result which I stated (but did not prove) in this mini-course, it suffices to estimate the distortion of linear embeddings into ℓ_1^d for an arbitrary d.

Proof description

- According to the result which I stated (but did not prove) in this mini-course, it suffices to estimate the distortion of linear embeddings into ℓ_1^d for an arbitrary d.
- With this result, our plan is the following: to find a collection of measures $\{\tau_\alpha\}_\alpha \in \mathrm{TC}(D_n)$ such that any linear operator $T : \mathrm{TC}(D_n) \to \ell_1^d$ which does not decrease the norms of all measures of the collection $\{\tau_\alpha\}_\alpha$, has the operator norm of order n.

Proof description

- According to the result which I stated (but did not prove) in this mini-course, it suffices to estimate the distortion of <u>linear</u> embeddings into ℓ_1^d for an arbitrary d.
- With this result, our plan is the following: to find a collection of measures $\{\tau_\alpha\}_\alpha \in \mathrm{TC}(D_n)$ such that any linear operator $T : \mathrm{TC}(D_n) \to \ell_1^d$ which does not decrease the norms of all measures of the collection $\{\tau_\alpha\}_\alpha$, has the operator norm of order n.
- The following things are the main difficulties: (1) Guess the right measures $\{\tau_\alpha\}_\alpha$, (2) Prove the statement of the previous paragraph.

Proof description

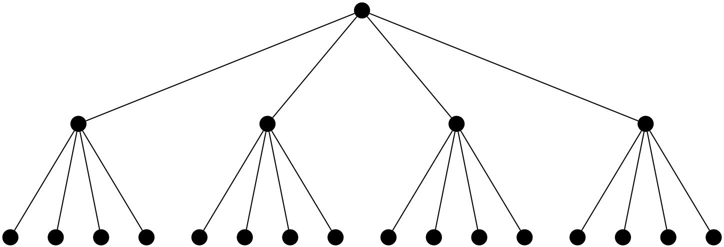
- According to the result which I stated (but did not prove) in this mini-course, it suffices to estimate the distortion of linear embeddings into ℓ_1^d for an arbitrary d.
- With this result, our plan is the following: to find a collection of measures $\{\tau_\alpha\}_\alpha \in \mathrm{TC}(D_n)$ such that any linear operator $T : \mathrm{TC}(D_n) \to \ell_1^d$ which does not decrease the norms of all measures of the collection $\{\tau_\alpha\}_\alpha$, has the operator norm of order n.
- The following things are the main difficulties: (1) Guess the right measures $\{\tau_\alpha\}_\alpha$, (2) Prove the statement of the previous paragraph.
- Experience showed us that a useful collection of measures can be built using "rooted trees" of measures.

Tree of measures

- In the case of diamond graphs, we use a 4-branching rooted tree. For each of its vertices t we introduce a measure μ_t.

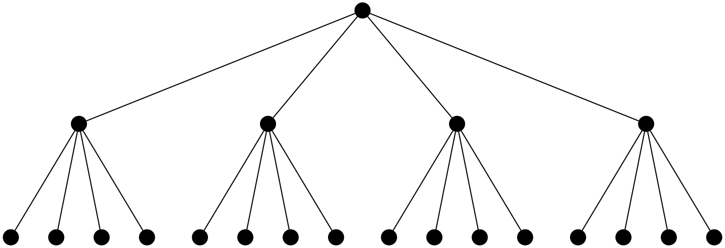
Tree of measures

- In the case of diamond graphs, we use a 4-branching rooted tree. For each of its vertices t we introduce a measure μ_t.



- On this figure you see a 4-branching tree of depth 2. The top vertex is a *root*.

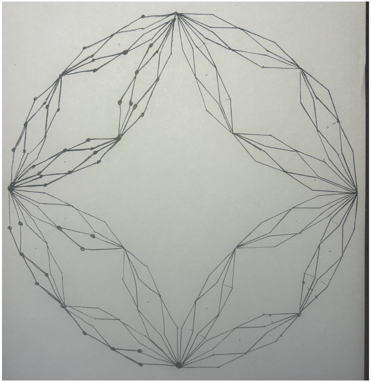
Tree of measures

- In the case of diamond graphs, we use a 4-branching rooted tree. For each of its vertices t we introduce a measure μ_t.
- 
  On this figure you see a 4-branching tree of depth 2. The top vertex is a *root*.
- For D_n, we need a tree of depth $n-1$, we denote it T. We introduce *level* k $(0 \le k \le n-1)$ of T, denoted T_k, as the set of vertices at distance k from the root. So T_0 is consists of the root only.

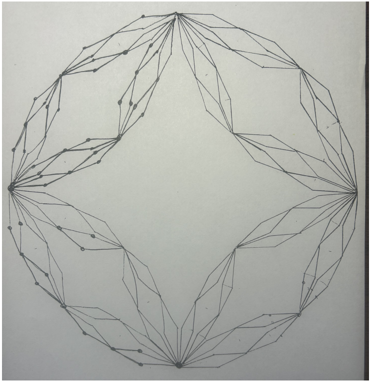
- **Terminology for diamonds:** To describe the measures, we introduce some terminology for diamonds.

- **Terminology for diamonds:** To describe the measures, we introduce some terminology for diamonds.

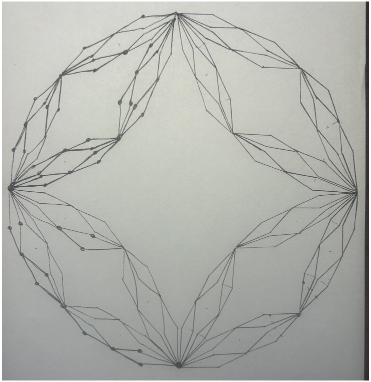
- D_4



- **Terminology for diamonds:** To describe the measures, we introduce some terminology for diamonds.
- D_n is its own *subdiamond of level* 0. It has 4 *subdiamonds of level* 1 (they are isomorphic to D_{n-1}), and 16 *subdiamonds of level* 2 (isomorphic to D_{n-2}), and so on.



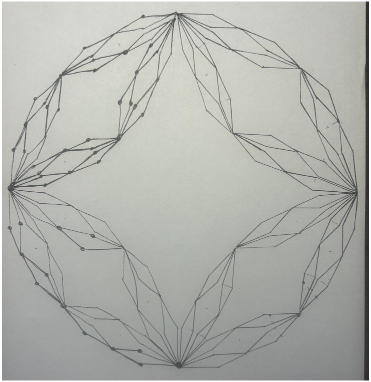
- D_4

- **Terminology for diamonds:** To describe the measures, we introduce some terminology for diamonds.
- D_n is its own *subdiamond of level* 0. It has 4 *subdiamonds of level* 1 (they are isomorphic to D_{n-1}), and 16 *subdiamonds of level* 2 (isomorphic to D_{n-2}), and so on.
- For each of the subdiamonds we define the corresponding midpoints, top and bottom.
- D_4 

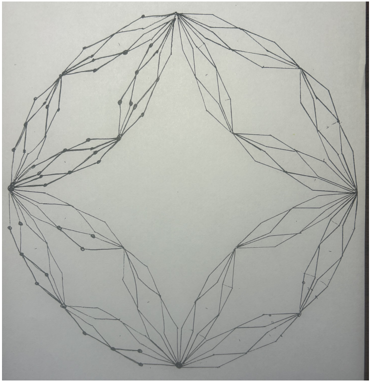
Observations, measures of level k

- The length of each edge in D_n is set to be 1. The distance between two midpoints in a subdiamond of level k is 2^{n-k}.

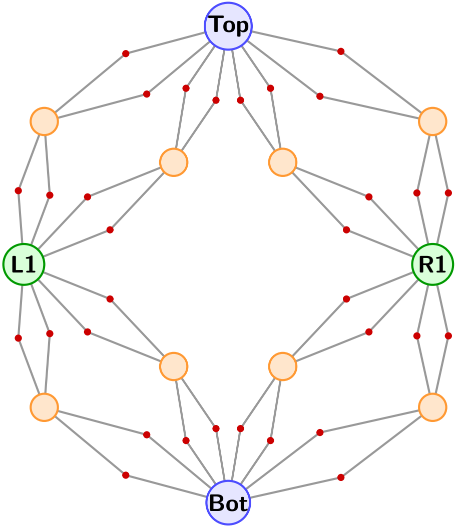
Observations, measures of level k

- The length of each edge in D_n is set to be 1. The distance between two midpoints in a subdiamond of level k is 2^{n-k}.
- D_4 

Observations, measures of level k

- The length of each edge in D_n is set to be 1. The distance between two midpoints in a subdiamond of level k is 2^{n-k}.
- D_4 
- In each of the 4^k subdiamonds of level k ($k \in \{0\dots,n-1\}$) we introduce a measure supported on the pair of midpoints of the subdiamond, with the masses $\pm 2^{n-k}\cdot 4^{-n}$, so that $\mu_t(D_n)=0$ and $\|\mu_t\|_{\mathrm{TC}}=4^{-k}$ for each $t\in T_k$.

Supports of measures: Level 0 - green, Level 1 - yellow, Level 2 - red



Tree of measures and multiplication by ± 1

- We say that several measures μ_t are on the same *level* if the corresponding vertices t of the tree are at the same distance from the root.

Tree of measures and multiplication by ± 1

- We say that several measures μ_t are on the same *level* if the corresponding vertices t of the tree are at the same distance from the root.
- We have introduced *measures of level k* supported on midpoints of subdiamonds of level k. There are 4^k of them and each subdiamond of level $k-1$ contains 4 subdiamonds of level k, so there is a natural tree structure.

Tree of measures and multiplication by ± 1

- We say that several measures μ_t are on the same *level* if the corresponding vertices t of the tree are at the same distance from the root.
- We have introduced *measures of level k* supported on midpoints of subdiamonds of level k. There are 4^k of them and each subdiamond of level $k-1$ contains 4 subdiamonds of level k, so there is a natural tree structure.
- It is easy to verify that the cost of the sum of measures on level k is equal to $1 = 4^k \cdot 4^{-k}$ for all levels from 0 to $n-1$.

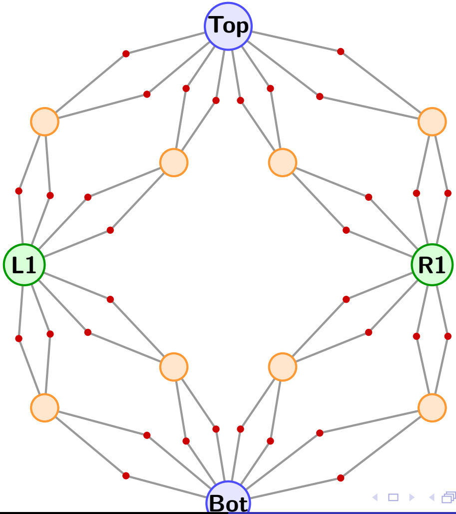
Tree of measures and multiplication by ± 1

- We say that several measures μ_t are on the same *level* if the corresponding vertices t of the tree are at the same distance from the root.
- We have introduced *measures of level* k supported on midpoints of subdiamonds of level k. There are 4^k of them and each subdiamond of level $k-1$ contains 4 subdiamonds of level k, so there is a natural tree structure.
- It is easy to verify that the cost of the sum of measures on level k is equal to $1 = 4^k \cdot 4^{-k}$ for all levels from 0 to $n-1$.
- **Reason:** Any measure of level k is supported on two midpoints, u and v, of a subdiamond of level k. The distance from these midpoints to any other midpoints of subdiamonds of level k cannot be smaller than $d(u, v)$.

Tree of measures and multiplication by ± 1

- We say that several measures μ_t are on the same *level* if the corresponding vertices t of the tree are at the same distance from the root.
- We have introduced *measures of level* k supported on midpoints of subdiamonds of level k. There are 4^k of them and each subdiamond of level $k-1$ contains 4 subdiamonds of level k, so there is a natural tree structure.
- It is easy to verify that the cost of the sum of measures on level k is equal to $1 = 4^k \cdot 4^{-k}$ for all levels from 0 to $n-1$.
- **Reason:** Any measure of level k is supported on two midpoints, u and v, of a subdiamond of level k. The distance from these midpoints to any other midpoints of subdiamonds of level k cannot be smaller than $d(u, v)$.
- One of the important conditions for the tree of measures: the cost of the sum of measures of each level remains equal 1 after multiplying measures μ_t by ± 1 arbitrarily.

Any measure of level k is supported on two midpoints, u and v, of a subdiamond of level k. The distance from these midpoints to any other midpoints of subdiamonds of level k cannot be smaller than $d(u, v)$.



The inequality implying the main theorem

- We do not need all possible ± 1. The following set of signs is sufficient: Let T_k be the set of measures of level k. Let $\{\theta_s(t)\}_{s,t\in T_k}$ be any orthogonal matrix with entries ± 1. A sufficient set of signs is $\{\sum_{t\in T_k}\theta_s(t)\mu_t\}_{s\in T_k}$ (Hadamard).

The inequality implying the main theorem

- We do not need all possible ± 1. The following set of signs is sufficient: Let T_k be the set of measures of level k. Let $\{\theta_s(t)\}_{s,t\in T_k}$ be any orthogonal matrix with entries ± 1. A sufficient set of signs is $\{\sum_{t\in T_k}\theta_s(t)\mu_t\}_{s\in T_k}$ (Hadamard).
- Let $f : \mathrm{TC}(D_n) \to \ell_1^d$, where $d \in \mathbb{N}$ is sufficiently large, be an embedding of $\mathrm{TC}(D_n)$ with almost optimal distortion. We also may and shall assume that f does not decrease norms.

The inequality implying the main theorem

- We do not need all possible ± 1. The following set of signs is sufficient: Let T_k be the set of measures of level k. Let $\{\theta_s(t)\}_{s,t\in T_k}$ be any orthogonal matrix with entries ± 1. A sufficient set of signs is $\{\sum_{t\in T_k}\theta_s(t)\mu_t\}_{s\in T_k}$ (Hadamard).
- Let $f:\mathrm{TC}(D_n)\to \ell_1^d$, where $d\in\mathbb{N}$ is sufficiently large, be an embedding of $\mathrm{TC}(D_n)$ with almost optimal distortion. We also may and shall assume that f does not decrease norms.
- We prove that choosing μ_t as above, we can achieve, for some absolute constant C,

$$\sum_{k=0}^{n-1}\left(\frac{1}{|T_k|}\sum_{s\in T_k}\left\|f\left(\sum_{t\in T_k}\theta_s(t)\mu_t\right)\right\|_{\ell_1^d}^2\right)^{1/2} \qquad (*)$$
$$\le \frac{C}{|E(D_n)|}\sum_{e\in E(D_n)}\|f(\mathbf{1}_{e^+}-\mathbf{1}_{e^-})\|_{\ell_1^d}.$$

The inequality (*) implies the main theorem

- We claim that the inequality (*), together with all the assumptions (which include that condition that f is non-contractive), imply that $\|f\| \geq n/C$, thus completing the proof.

The inequality (*) implies the main theorem

- We claim that the inequality (*), together with all the assumptions (which include that condition that f is non-contractive), imply that $\|f\| \ge n/C$, thus completing the proof.
- We repeat the inequality (*):

$$\sum_{k=0}^{n-1}\left(\frac{1}{|T_k|}\sum_{s\in T_k}\left\|f\left(\sum_{t\in T_k}\theta_s(t)\mu_t\right)\right\|_{\ell_1^d}^2\right)^{1/2} \le \frac{C}{|E(D_n)|}\sum_{e\in E(D_n)}\|f(\mathbf{1}_{e^+}-\mathbf{1}_{e^-})\|_{\ell_1^d}.$$

The inequality (*) implies the main theorem

- We claim that the inequality (*), together with all the assumptions (which include that condition that f is non-contractive), imply that $\|f\| \geq n/C$, thus completing the proof.
- We repeat the inequality (*):

$$\sum_{k=0}^{n-1}\left(\frac{1}{|T_k|}\sum_{s\in T_k}\left\|f\left(\sum_{t\in T_k}\theta_s(t)\mu_t\right)\right\|_{\ell_1^d}^2\right)^{1/2} \leq \frac{C}{|E(D_n)|}\sum_{e\in E(D_n)}\|f(\mathbf{1}_{e^+}-\mathbf{1}_{e^-})\|_{\ell_1^d}.$$

- Our assumptions imply that the norm in the left-hand side is at least 1. Thus, the left-hand side is at least n.

The inequality (*) implies the main theorem

- We claim that the inequality (*), together with all the assumptions (which include that condition that f is non-contractive), imply that $\|f\| \geq n/C$, thus completing the proof.
- We repeat the inequality (*):

$$\sum_{k=0}^{n-1}\left(\frac{1}{|T_k|}\sum_{s\in T_k}\left\|f\left(\sum_{t\in T_k}\theta_s(t)\mu_t\right)\right\|_{\ell_1^d}^2\right)^{1/2}$$
$$\leq \frac{C}{|E(D_n)|}\sum_{e\in E(D_n)}\|f(\mathbf{1}_{e^+}-\mathbf{1}_{e^-})\|_{\ell_1^d}.$$

- Our assumptions imply that the norm in the left-hand side is at least 1. Thus, the left-hand side is at least n.
- As for the right-hand side. We use the fact that for each edge e, we have $\|\mathbf{1}_{e^+}-\mathbf{1}_{e^-}\|_{\mathrm{TC}}=1$. Therefore the right-hand side does not exceed $C\|f\|$, this completes the proof.

Proving the inequality

- **Step 1:** Using Parseval's identity, reduce to a Sobolev inequality for a scalar function.

Proving the inequality

- **Step 1:** Using Parseval's identity, reduce to a Sobolev inequality for a scalar function.
- **Step 2:** Proof of the Sobolev inequality (for the defined above measures μ_t).

Step 1

- $$\left(\sum_{k=0}^{n-1}\left(\frac{1}{|T_k|}\sum_{s\in T_k}\left\|f\left(\sum_{t\in T_k}\theta_s(t)\mu_t\right)\right\|_{\ell_1^d}^2\right)\right)^{1/2}$$
$$\stackrel{\text{def}\ \|\cdot\|_1}{=}\left(\sum_{k=0}^{n-1}\left(\frac{1}{|T_k|}\sum_{s\in T_k}\left(\sum_{j=1}^{d}\left|f_j\left(\sum_{t\in T_k}\theta_s(t)\mu_t\right)\right|\right)^2\right)\right)^{1/2}$$
$$\stackrel{\text{triangle}}{\leq}\sum_{j=1}^{d}\left(\sum_{k=0}^{n-1}\left(\frac{1}{|T_k|}\sum_{s\in T_k}\left|f_j\left(\sum_{t\in T_k}\theta_s(t)\mu_t\right)\right|^2\right)\right)^{1/2}$$
$$\stackrel{\text{Parseval \& linearity}}{=}\sum_{j=1}^{d}\left(\sum_{k=0}^{n-1}\left(\sum_{t\in T_k}|f_j(\mu_t)|^2\right)\right)^{1/2}$$

Moving to Step 2

- So we got
$$\leq \sum_{j=1}^{d} \sum_{k=0}^{n-1} \left(\sum_{t \in T_k} |f_j(\mu_t)|^2 \right)^{1/2}$$
in the rightmost side.

Moving to Step 2

- So we got
$$\leq \sum_{j=1}^{d} \sum_{k=0}^{n-1} \left(\sum_{t \in T_k} |f_j(\mu_t)|^2 \right)^{1/2}$$
in the rightmost side.
- We want to continue in order to get
$$\leq \frac{C}{|E(D_n)|} \sum_{e \in E(D_n)} \|f(\mathbf{1}_{e^+} - \mathbf{1}_{e^-})\|_{\ell_1^d}$$
$$= \sum_{j=1}^{d} \frac{C}{|E(D_n)|} \sum_{e \in E(D_n)} |f_j(\mathbf{1}_{e^+} - \mathbf{1}_{e^-})|$$

Moving to Step 2

- So we got
$$\le \sum_{j=1}^{d} \sum_{k=0}^{n-1} \left(\sum_{t \in T_k} |f_j(\mu_t)|^2 \right)^{1/2}$$
in the rightmost side.
- We want to continue in order to get
$$\le \frac{C}{|E(D_n)|} \sum_{e \in E(D_n)} \| f(\mathbf{1}_{e^+} - \mathbf{1}_{e^-}) \|_{\ell_1^d}$$
$$= \sum_{j=1}^{d} \frac{C}{|E(D_n)|} \sum_{e \in E(D_n)} |f_j(\mathbf{1}_{e^+} - \mathbf{1}_{e^-})|$$
- So, it remains to prove the following Sobolev inequality
$$\sum_{k=0}^{n-1} \left(\sum_{t \in T_k} |f_j(\mu_t)|^2 \right)^{1/2} \le \frac{C}{|E(D_n)|} \sum_{e \in E(D_n)} |f_j(\mathbf{1}_{e^+} - \mathbf{1}_{e^-})|$$

Proving the Sobolev inequality

- First step is to show that we can reduce the general case to the case where f_j is the indicator function of a simply connected set A in $V(D_n)$.

Proving the Sobolev inequality

- First step is to show that we can reduce the general case to the case where f_j is the indicator function of a simply connected set A in $V(D_n)$.
- By a *simply connected set* we mean a set A such that both A and A^c are connected.

Proving the Sobolev inequality

- First step is to show that we can reduce the general case to the case where f_j is the indicator function of a simply connected set A in $V(D_n)$.
- By a *simply connected set* we mean a set A such that both A and A^c are connected.
- The reduction is a rather straightforward consequence of the proved by M. O.[Ost05] characterization of extreme points in unit balls of Sobolev spaces on graphs.

Proving the Sobolev inequality

- First step is to show that we can reduce the general case to the case where f_j is the indicator function of a simply connected set A in $V(D_n)$.
- By a *simply connected set* we mean a set A such that both A and A^c are connected.
- The reduction is a rather straightforward consequence of the proved by M. O.[Ost05] characterization of extreme points in unit balls of Sobolev spaces on graphs.
- After that, it remains to prove that there is an absolute constant C such that for each simply connected subset A of $V(D_n)$ the following inequality holds (we use the integral for $\mathbf{1}_A(\mu_t)$).

$$\sum_{k=0}^{n-1}\left(\sum_{t\in T_k}\left|\int \mathbf{1}_A d\mu_t\right|^2\right)^{\frac{1}{2}} \le \frac{C}{|E(D_n)|}\sum_{(u,v)\in E(D_n)}|\mathbf{1}_A(v)-\mathbf{1}_A(u)|.$$

Some terminology

- We say that A *splits* the measure μ_t if exactly one of the two points of support of μ_t belongs to A.

Some terminology

- We say that A *splits* the measure μ_t if exactly one of the two points of support of μ_t belongs to A.
- We say that A *splits* a subdiamond if only one vertex of the pair $\{$top of the subdiamond, bottom of the subdiamond$\}$ is in A.

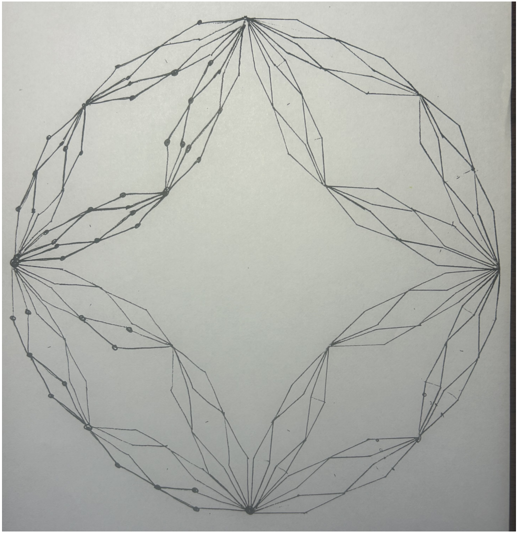
Some terminology

- We say that A *splits* the measure μ_t if exactly one of the two points of support of μ_t belongs to A.
- We say that A *splits* a subdiamond if only one vertex of the pair {top of the subdiamond, bottom of the subdiamond} is in A.
- We start with the case where A either splits the only measure in T_0 or the subdiamond of order 0. We describe this condition as A splits D_1 (because both A and A^c intersect $V(D_1)$).

Some terminology

- We say that A *splits* the measure μ_t if exactly one of the two points of support of μ_t belongs to A.
- We say that A *splits* a subdiamond if only one vertex of the pair {top of the subdiamond, bottom of the subdiamond} is in A.
- We start with the case where A either splits the only measure in T_0 or the subdiamond of order 0. We describe this condition as A splits D_1 (because both A and A^c intersect $V(D_1)$).
- We claim that in such a case A splits at most 2^k measures in T_k, $1 \leq k \leq n-1$.

Ways to split the measure in T_0



▶

- (Repeated) We claim that if A splits D_1, then A splits at most 2^k measures in T_k, $1 \leq k \leq n-1$.

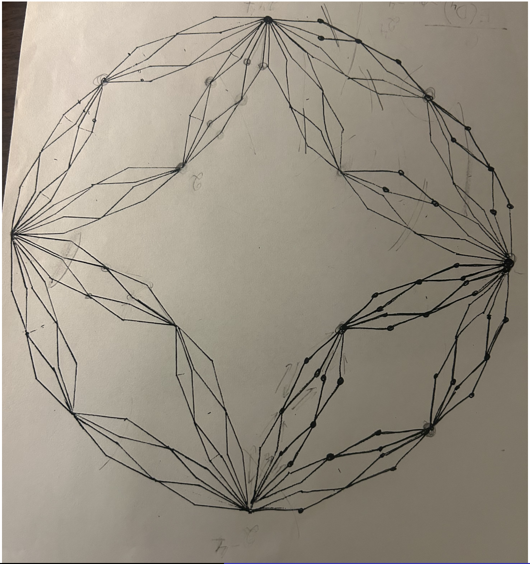
- (Repeated) We claim that if A splits D_1, then A splits at most 2^k measures in T_k, $1 \leq k \leq n-1$.
- Note that simple connectedness implies that if A does not split a subdiamond S and both A and A^c have nonempty intersection with S^c, then either all vertices of S are in A or in A^c.

- (Repeated) We claim that if A splits D_1, then A splits at most 2^k measures in T_k, $1 \le k \le n-1$.
- Note that simple connectedness implies that if A does not split a subdiamond S and both A and A^c have nonempty intersection with S^c, then either all vertices of S are in A or in A^c.
- A splits D_1 (and is simply connected) $\Rightarrow$ A splits 2 out 4 subdiamonds of level 1 $\Rightarrow$ splits at most 2 measures of level 1.

- (Repeated) We claim that if A splits D_1, then A splits at most 2^k measures in T_k, $1 \le k \le n-1$.
- Note that simple connectedness implies that if A does not split a subdiamond S and both A and A^c have nonempty intersection with S^c, then either all vertices of S are in A or in A^c.
- A splits D_1 (and is simply connected) $\Rightarrow$ A splits 2 out 4 subdiamonds of level 1 $\Rightarrow$ splits at most 2 measures of level 1.
- A splits 2 subdiamonds of level 1 (and is simply connected) $\Rightarrow$ A splits 4 subdiamonds of level 2 $\Rightarrow$ splits at most 4 measures of level 2.

- (Repeated) We claim that if A splits D_1, then A splits at most 2^k measures in T_k, $1 \leq k \leq n-1$.
- Note that simple connectedness implies that if A does not split a subdiamond S and both A and A^c have nonempty intersection with S^c, then either all vertices of S are in A or in A^c.
- A splits D_1 (and is simply connected) $\Rightarrow$ A splits 2 out 4 subdiamonds of level 1 $\Rightarrow$ splits at most 2 measures of level 1.
- A splits 2 subdiamonds of level 1 (and is simply connected) $\Rightarrow$ A splits 4 subdiamonds of level 2 $\Rightarrow$ splits at most 4 measures of level 2.
- So on. It proves the statement repeated above.

Proving the statement on measures splitting $V(D_1)$

- 

- Easy counting argument implies that the number of edges belonging to $\partial_E A$ (the set of edges with only one end in A; such edges contribute to the right-hand side of the Sobolev inequality) needed to split D_1 is 2^n.

- Easy counting argument implies that the number of edges belonging to $\partial_E A$ (the set of edges with only one end in A; such edges contribute to the right-hand side of the Sobolev inequality) needed to split D_1 is 2^n.
- So the inequality

$$\sum_{k=0}^{n-1}\left(\sum_{t\in T_k}\left|\int \mathbf{1}_A d\mu_t\right|^2\right)^{\frac{1}{2}} \le \frac{C}{|E(D_n)|}\sum_{(u,v)\in E(D_n)}|\mathbf{1}_A(v)-\mathbf{1}_A(u)|$$

becomes (if A splits the measure in T_0)

- Easy counting argument implies that the number of edges belonging to $\partial_E A$ (the set of edges with only one end in A; such edges contribute to the right-hand side of the Sobolev inequality) needed to split D_1 is 2^n.
- So the inequality

$$\sum_{k=0}^{n-1}\left(\sum_{t\in T_k}\left|\int \mathbf{1}_A d\mu_t\right|^2\right)^{\frac{1}{2}} \le \frac{C}{|E(D_n)|}\sum_{(u,v)\in E(D_n)}|\mathbf{1}_A(v)-\mathbf{1}_A(u)|$$

becomes (if A splits the measure in T_0)

- $((2^{-n})^2)^{1/2}+(2\cdot(2^{-n-1})^2)^{1/2}+(2^2(2^{-n-2})^2)^{1/2}+\cdots\le\frac{C}{4^n}\cdot 2^n.$

- Easy counting argument implies that the number of edges belonging to $\partial_E A$ (the set of edges with only one end in A; such edges contribute to the right-hand side of the Sobolev inequality) needed to split D_1 is 2^n.
- So the inequality

$$\sum_{k=0}^{n-1}\left(\sum_{t\in T_k}\left|\int \mathbf{1}_A d\mu_t\right|^2\right)^{\frac{1}{2}} \le \frac{C}{|E(D_n)|}\sum_{(u,v)\in E(D_n)}|\mathbf{1}_A(v)-\mathbf{1}_A(u)|$$

  becomes (if A splits the measure in T_0)
- $((2^{-n})^2)^{1/2}+(2\cdot(2^{-n-1})^2)^{1/2}+(2^2(2^{-n-2})^2)^{1/2}+\cdots\le\frac{C}{4^n}\cdot 2^n$.
- It remains to observe that the LHS can be transformed to $(2^{-n})(1+2^{-1/2}+\cdots+2^{-n/2})\le\frac{\sqrt{2}}{\sqrt{2}-1}\cdot 2^{-n}$. The inequality follows.

The complementary case

- Let $k \in \{1, \ldots, n-1\}$ be the largest for which vertices of $V(D_k)$ are either all in A or all in A^c. WLOG assume that all of them are in A.

The complementary case

- Let $k \in \{1, \ldots, n-1\}$ be the largest for which vertices of $V(D_k)$ are either all in A or all in A^c. WLOG assume that all of them are in A.
- Then there is exactly one subdiamond of level k containing vertices of A^c (otherwise, A^c would be disconnected).

The complementary case

- Let $k \in \{1, \ldots, n-1\}$ be the largest for which vertices of $V(D_k)$ are either all in A or all in A^c. WLOG assume that all of them are in A.
- Then there is exactly one subdiamond of level k containing vertices of A^c (otherwise, A^c would be disconnected).
- Then A splits exactly two subdiamonds of level $k+1$. This implies that $|\partial_E A| = 2^{n-k}$.

The complementary case

- Let $k \in \{1, \ldots, n-1\}$ be the largest for which vertices of $V(D_k)$ are either all in A or all in A^c. WLOG assume that all of them are in A.
- Then there is exactly one subdiamond of level k containing vertices of A^c (otherwise, A^c would be disconnected).
- Then A splits exactly two subdiamonds of level $k+1$. This implies that $|\partial_E A| = 2^{n-k}$.
- Also this implies that A splits at most one problem of level k. As in the previous case, we get that A splits at most two problems of level $k+1$, 4 problems of level $k+2$, so on.

The complementary case

- Let $k \in \{1, \ldots, n-1\}$ be the largest for which vertices of $V(D_k)$ are either all in A or all in A^c. WLOG assume that all of them are in A.
- Then there is exactly one subdiamond of level k containing vertices of A^c (otherwise, A^c would be disconnected).
- Then A splits exactly two subdiamonds of level $k+1$. This implies that $|\partial_E A| = 2^{n-k}$.
- Also this implies that A splits at most one problem of level k. As in the previous case, we get that A splits at most two problems of level $k+1$, 4 problems of level $k+2$, so on.
- The Sobolev inequality in this case is $((2^{-n-k})^2)^{1/2} + (2 \cdot (2^{-n-k-1})^2)^{1/2} + (2^2(2^{-n-k-2})^2)^{1/2} + \cdots \leq \frac{C}{4^n} \cdot 2^{n-k}$.

The complementary case

- Let $k \in \{1, \dots, n-1\}$ be the largest for which vertices of $V(D_k)$ are either all in A or all in A^c. WLOG assume that all of them are in A.
- Then there is exactly one subdiamond of level k containing vertices of A^c (otherwise, A^c would be disconnected).
- Then A splits exactly two subdiamonds of level $k+1$. This implies that $|\partial_E A| = 2^{n-k}$.
- Also this implies that A splits at most one problem of level k. As in the previous case, we get that A splits at most two problems of level $k+1$, 4 problems of level $k+2$, so on.
- The Sobolev inequality in this case is $((2^{-n-k})^2)^{1/2} + (2 \cdot (2^{-n-k-1})^2)^{1/2} + (2^2(2^{-n-k-2})^2)^{1/2} + \cdots \leq \frac{C}{4^n} \cdot 2^{n-k}$.
- As above, the inequality can be proved summing the geometric series.

I. Abraham, Y. Bartal, O. Neiman, Nearly tight low stretch spanning trees, 49th Annual IEEE Symposium on Foundations of Computer Science, 2008; `arXiv:0808.2017`.

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.

R. J. Aliaga, E. Pernecká, R. Smith, A solution to the extreme point problem and other applications of Choquet theory to Lipschitz-free spaces, *Trans. Amer. Math. Soc.* `DOI: https://doi.org/10.1090/tran/9778`.

N. Alon, R. Karp, D. Peleg, D. West, A graph-theoretic game and its application to the k-server problem. SIAM J. Comput. 24 (1995), 78–100.

R. F. Arens, J. Eells, Jr., On embedding uniform and topological spaces, *Pacific J. Math.*, **6** (1956), 397–403.

Y. Bartal, Probabilistic approximation of metric spaces and its algorithmic applications, In: *37th Annual Symposium on Foundations of Computer Science* (Burlington, VT, 1996), 184–193, IEEE Comput. Soc. Press, Los Alamitos, CA, 1996.

Y. Bartal, On approximating arbitrary metrics by tree metrics. STOC 98. Proceedings of the 30th annual ACM symposium on theory of computing, Dallas, TX, USA, May 23–26, 1998. New York, NY: ACM, Association for Computing Machinery. 161–168 (1998).

F. Baudier, C. Gartland, Th. Schlumprecht. L_1-distortion of Wasserstein metrics into L_1: A tale of two dimensions. *Trans. Amer. Math. Soc.* Ser. B **10** (2023), 1077–1118.

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

J. Bourgain, Remarks on the extension of Lipschitz maps defined on discrete sets and uniform homeomorphisms, in: *Geometrical aspects of functional analysis* (1985/86), 157–167, *Lecture Notes in Math.*, **1267**, Springer, Berlin, 1987.

M. 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.

S. J. Dilworth, D. Kutzarova, M. I. Ostrovskii, Lipschitz-free spaces on finite metric spaces, *Canad. J. Math.*, **72** (2020), 774–804

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.

- S. J. Dilworth, D. Kutzarova, M. I. Ostrovskii, Cycle spaces: invariant projections and applications to transportation cost, *Proc. Symp. Pure Math.*, **106** (2024), 97–131.
- P. Erdős, L. Pósa, On the maximal number of disjoint circuits of a graph. *Publ. Math. Debrecen* **9** (1962), 3–12.
- J. Fakcharoenphol, S. Rao, K. Talwar, A tight bound on approximating arbitrary metrics by tree metrics, *J. Comput. System Sci.*, **69** (2004), no. 3, 485–497.
- C. Gartland, M. Ostrovskii, Lower estimates for L_1-distortion of transportation cost spaces, *STOC'26—Proceedings of the 58th Annual ACM Symposium on Theory of Computing*, 698-709, ACM, New York, 2026. Full version: `arXiv:2602.14852v1`.
- C. Gartland, M. Ostrovskii, Y. Rabani, R. Young, L_1-distortion of Earth Mover Distances and Transportation Cost Spaces on High Dimensional Grids, `arXiv:2602.19434v1`, 2026.

A. Gupta, I. Newman, Y. Rabinovich, A. Sinclair, Cuts, trees and ℓ_1-embeddings of graphs, *Combinatorica*, **24** (2004) 233–269.

P. Indyk, N. Thaper, Fast image retrieval via embeddings, *Proc. of the 3rd Int'l Workshop on Statistical and Computational Theories of Vision*, 2003.

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*, 189-203, Adv. Anal. Geom., **2**, De Gruyter, Berlin, 2020.

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

G. Konjevod, R. Ravi, F. S. Salman, On approximating planar metrics by tree metrics. *Inform. Process. Lett.* **80** (2001), no. 4, 213–219.

- R. Medina, G. Tresch, Transportation cost spaces and stochastic trees, `arXiv:2501.08656v1`.
- A. Naor, G. Schechtman. Planar earthmover is not in L_1. *SIAM J. Computing*, **37** (2007), 804–826.
- S. Ostrovska, M. I. Ostrovskii, On relations between transportation cost spaces and ℓ_1, *J. Math. Anal. Appl.*, **491** (2020), no. 2, 124338.
- S. Ostrovska, M. I. Ostrovskii, Isometric structure of transportation cost spaces on finite metric spaces. *Rev. R. Acad. Cienc. Exactas Fís. Nat.* Ser. A Mat. RACSAM **116** (2022), no. 4, Paper No. 153.
- M. I. Ostrovskii. Sobolev spaces on graphs. *Quaestiones Mathematicae*, **28** (2005), 501–523.
- M. I. Ostrovskii, *Metric Embeddings: Bilipschitz and Coarse Embeddings into Banach Spaces*, de Gruyter Studies in Mathematics, **49**. Walter de Gruyter & Co., Berlin, 2013.

- Th. Schlumprecht, G. Tresch, Stochastic embeddings of graphs into trees. *Banach J. Math. Anal.* **18** (2024), no. 4, Paper No. 61, 25 pp.
- G. Skandalis, A. Valette On the transportation cost norm on finite metric graphs, `arXiv:2601.16859`
- A. N. Tolstoi, Methods of removing inefficient transportation in planning (Russian), Socialist Transport **9** (1939) 28–51.
- A. M. Vershik, Classification of finite metric spaces and combinatorics of convex polytopes. *Arnold Math. J.* **1** (2015), no. 1, 75–81.