Directed Metrics and Directed Graph Partitioning Problems

Moses Charikar* Konstantin Makarychev † Yury Makarychev ‡

Abstract

The theory of embeddings of finite metrics has provided a powerful toolkit for graph partitioning problems in undirected graphs. The connection comes from the fact that the integrality gaps of mathematical programming relaxations for sparsest cut in undirected graphs is exactly equal to the minimum distortion required to embed certain metrics into ℓ_1. No analog of this metric embedding theory exists for directed (asymmetric) metrics, the natural distance functions that arise in considering mathematical relaxations for directed graph partioning problems. We initiate a study of metric embeddings for directed metrics, motivated by understanding directed variants of sparsest cut.

It turns out that there are two different ways to formulate sparsest cut in directed graphs (depending on whether one insists on partitioning the graph into two pieces or not). Different subclasses of directed metrics arise in the consideration of mathematical relaxations for these two formulations and the embedding questions that result are quite different. Unlike in the undirected case, where the natural host space is ℓ_1, the host space in the directed case is not obvious and depends on the problem formulation. Our work is a first step at understanding this space of directed metrics, the resulting embedding questions and their relationships to directed graph partitioning problems.

1 Introduction

Embedding techniques have proved to be quite useful for rounding LP and SDP relaxations for graph partitioning problems. Our techniques for partitioning directed graphs are still quite weak. While the rich theory of embedding finite metrics into normed spaces (see [13, 17]) gives a powerful toolkit for partitioning undirected graphs, no analog of this toolkit exists for directed metrics, the natural asymmetric variant of symmetric metrics. Directed metrics arise naturally in considering LP and SDP relaxations for partitioning directed graphs. In this work, we initiate a study of embeddings for such directed metrics.

The close connection between metric embeddings and partitioning undirected graphs comes from the fact that the integrality gap of LP and SDP relaxations for sparsest cut is exactly equal to the minimum distortion required for embedding certain undirected metrics into ℓ_1. The integrality gap of the natural LP relaxation for sparsest cut is the same as the minimum distortion required to embed arbitrary metrics into ℓ_1 (as pointed out by Linial, London and Rabinovich [16]), while the integrality gap of a certain stronger SDP relaxation is equal to the distortion required for embedding a class of metrics called ℓ_2^2 metrics into ℓ_1. The former embedding question was addressed by Bourgain [4]. This latter embedding question has been the focus of much research recently [2, 5, 14].

Our investigation of directed metrics and their embedding questions is motivated by the study of directed variants of sparsest cut[1]. It turns out that are two distinct ways to formulate the sparsest cut problem for directed graphs. In one formulation, we require that the graph be partitioned into two pieces while the other formulation does not have this constraint. The resulting problems are called Bipartite Directed Sparsest Cut and Directed Sparsest Cut respectively. (These will be formally defined later).

We study the classes of directed metrics that arise in studying natural LP and SDP relaxations for these directed variants of sparsest cut. The basic classes of such metrics[2] and their inclusions are depicted in Figure 1. We will define these classes of metrics later. The integrality gap of a natural LP relaxation for bipartite directed sparsest cut is equal to the distortion required for embedding weighted directed metrics into

*Computer Science Dept., Princeton University. E-mail: moses@cs.princeton.edu. Supported by NSF ITR grant CCR-0205594, DOE Early Career Principal Investigator award DE-FG02-02ER25540, NSF CAREER award CCR-0237113 and an Alfred P. Sloan Fellowship.

†Computer Science Dept., Princeton University. E-mail: kmakaryc@cs.princeton.edu. Supported by a Gordon Wu fellowship.

‡Computer Science Dept., Princeton University. E-mail: ymakaryc@cs.princeton.edu. Supported by a Gordon Wu fellowship.

[1]Unless otherwise stated, our formulations of sparsest cut will involve non-uniform demands.

[2]We will use the terms semimetric and metric interchangeably, although technically the correct terminology is a semimetric where 0 distances between distinct points are allowed.

directed ℓ_1. The corresponding integrality gap question for a natural SDP relaxation is equal to the distortion for embedding directed ℓ_2^2 metrics into directed ℓ_1. On the other hand, for directed sparsest cut, the integrality gap for a natural LP relaxation is equal to the distortion for embedding directed metrics into directed 0-1 metrics.

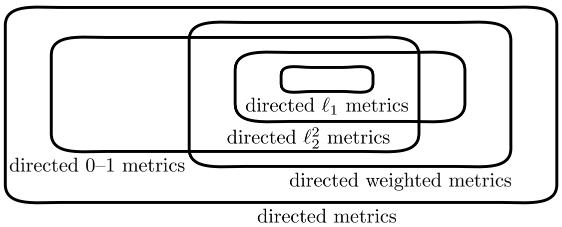

Figure 1: Classes of Directed Metric Spaces

Our contribution is identifying interesting classes of directed metrics and connecting embedding questions here to directed graph partioning problems. We clarify some of the relationships between these various classes of directed metrics and show distortion lower bounds for mappings between them. For Bipartite Directed Sparsest Cut, we show a strong hardness result of $\Omega(2^{(\log n)^\delta})$ assuming that 3-SAT cannot be solved in subexponential time (and $\Omega(n^\delta)$ assuming a certain conjecture about refuting random 3-SAT formulae). The hardness result for the problem with non-uniform demands is in contrast to the $O(\sqrt{\log n})$ approximation factor that can be achieved for the version with uniform demands [1]. We also show an $\Omega(n^\delta)$ lower bound for the related question of embedding directed ℓ_2^2 metrics into directed ℓ_1. This is based on an explicit combinatorial construction.

We hope that our work will initiate an investigation of directed metrics and their associated embedding questions. The most interesting open problem here seems to be to understand how well directed metrics embed into (convex combinations of) directed 0-1 metrics. As mentioned before, the minimum distortion is exactly equal to the integrality gap of a natural LP relaxation for directed sparsest cut. This is known to be $O(\sqrt{n})$ by the results of Hajiaghayi and Räcke [12], who extended work of Gupta [10] and Cheriyan, Karloff and Rabani [6] on Directed Multicut. However the best lower bound is only $\Omega(\log n)$ that follows from the undirected case. The approximability of Directed Sparsest Cut is related within logarithmic factors to that of Directed Multicut and both are very interesting open problems.

Organization: In Section 2 we introduce the notion of directed semimetric, define directed counterparts of metric spaces ℓ_p, as well as prove several theorems on the directed semimetrics. In Section 3 we define the directed partitioning problems that we study. Then in Section 4 we show that integrality gaps of an SDP relaxation for the Bipartite Directed Sparsest Cut Problem, and for an LP relaxation for the Directed Sparsest Cut Problem have natural embedding formulations. In Section 5 we show hardness results for Bipartite Directed Sparsest Cut by reduction from Maximum Edge Bipartite Clique. In Section 6 we show a lower bound of $\Omega(n^\delta)$ for embedding directed ℓ_2^2 semimetric spaces into directed ℓ_1. Finally, in Section 7 we introduce a notion of weak embeddability, and prove a directed counterpart of Bourgain's theorem for weak embeddings of directed semimetrics.

2 Definitions and Basic Facts

Definition 2.1. *A directed semimetric*[3] *is a set X with a distance function $d : X \times X \to \mathbb{R}^+ \cup \{0\}$ such that*

1. $\forall\, x, y \in X$ $d(x, y) \geq 0$.
2. $\forall\, x \in X$ $d(x, x) = 0$.
3. *The triangle inequality holds:* $\forall\, x, y, z \in V$ $d(x, y) + d(y, z) \geq d(x, z)$.

Definition 2.2. *Given a metric space (X, d), and a point $s \in X$, we define a directed distance $d_{X,s}$ by $d_{X,s}(x, y) = d(x, y) + d(x, s) - d(y, s)$.*

Observe that the set X with the distance $d_{X,s}$ is a directed semimetric. Indeed,

1. $d_{X,s}(x, y) = d(x, y) + d(x, s) - d(y, s) \geq 0$ by the triangle inequality for the metric d.
2. $d_{X,s}(x, x) = d(x, x) + d(x, x) - d(x, x) = 0$.
3. $d_{X,s}(x, y) + d_{X,s}(y, z) = d(x, y) + d(x, s) - d(y, s) + d(y, z) + d(y, s) - d(z, s) = [d(x, y) + d(y, z)] + d(x, s) - d(z, s) \geq d(x, z) + d(x, s) - d(z, s) = d_{X,s}(x, z)$.

Note that for a homogeneous metric space[4] X, directed semimetric spaces $(X, d_{X,s})$ are isometric for different choices of s. Therefore we can arbitrarily pick the point s. Now we can define directed counterparts of standard ("undirected") metric spaces.

[3] Directed semimetrics are sometimes called quasi-semimetrics.

[4] i.e. a space s.t. for every two points x and y there exists an isometry that maps x to y

Definition 2.3. *The directed semimetric ℓ_p is the space ℓ_p with distance function $d_{\ell_p,0}$. I.e. the directed distance between two points x and y equals to*

$$d_p(x,y) \equiv d_{\ell_p,0}(x,y) = (\sum_i |x_i - y_i|^p)^{1/p} + (\sum_i |x_i|^p)^{1/p} - (\sum_i |y_i|^p)^{1/p}.$$

In particular, for the directed ℓ_1, if all coordinates of x and y are positive, we have $d_1(x,y) = \sum_i |x_i - y_i| + \sum_i |x_i| - \sum_i |y_i| = \sum_i x_i \dot{-} y_i$. (where $a \dot{-} b \equiv \max(a-b,0)$).

Similarly, if X is an ℓ_2^2 space we define a directed semimetric ℓ_2^2 space as $(X, d_{X,s})$ (where $s \in X$).

Remark. Our choice of the definition of directed ℓ_p spaces was motivated by the following goals: 1) Properties of directed spaces ℓ_p should resemble those of undirected spaces ℓ_p. In particular, ℓ_p should isometrically embed into ℓ_1 for $p \in [1,2]$ (see below for the definitions, see also Corollary 2.1). 2) The definition should have applications to combinatorial problems, as we will show our definition of ℓ_1 and ℓ_2^2 have.

Another important distance is a cut metric. Given a set X, and a subset $S \subset X$, the cut metric d_S defined as follows: the distance between a point in S and a point in $X \setminus S$ is 1; the distances between points in S, and the distances between points in $X \setminus S$ are 0. Let us pick an $s \in S$, then we get the following directed distance:

$$d_{S,X,s}(x,y) = d_S(x,y) + d_S(x,s) - d_S(y,s) = \begin{cases} 2, & \text{if } x \in S,\ y \notin S; \\ 0, & \text{otherwise}; \end{cases}$$

For convenience, we scale this distance by half.

Definition 2.4. *Given a set X and a subset $S \subset X$, the directed cut metric $\delta_S(x,y)$ is*

$$\delta_S(x,y) \equiv \frac{1}{2} d_{S,X,s} = \begin{cases} 1, & \textit{if } x \in S,\ y \notin S; \\ 0, & \textit{otherwise}; \end{cases}$$

Definition 2.5. *The distortion of an embedding ϕ of a directed metric space X to a directed metric Y is equal to the infimum of D for which there exists a scale $c > 0$ s.t.*

$$c\, d_X(x_1,x_2) \le d_Y(\phi(x_1),\phi(x_2)) \le c\, D\, d_X(x_1,x_2).$$

If no such D exists we say that the distortion is infinite. We say that X embeds isometrically to Y (or (X is Y-embeddable) if there exists an embedding of X to Y with distortion 1.

Note that a convex combination of cut metrics embeds into ℓ_1. First we embed each cut metric into a one dimensional directed ℓ_1 space: we embed S into the point with the coordinate equal to the weight of the cut in the convex combination, we embed the complement of S into 0. Then we concatenate all these embeddings.

Lemma 2.1. *Any directed ℓ_1-embeddable semimetric d on a set X is a convex combination of directed cut metrics.*

Proof. We reduce the problem to the well known undirected counterpart of this lemma. Let ϕ be an isometric embedding of (X,d) to the directed ℓ_1 space. Then

$$d(x,y) = \|\phi(x) - \phi(y)\|_1 + \|\phi(x) - 0\|_1 - \|\phi(y) - 0\|_1.$$

We know that the ℓ_1 norm $\|\cdot\|_1$ restricted to the set $\phi(S) \cup \{0\}$ is a convex combination of cut metrics d_S: $\|x-y\|_1 = \sum_{S \subset \phi(X)} \alpha_S d_S(x,y)$ (since the cut metric for the set S is the same as for the set $X \setminus S$, here we assume that S always does not contain 0). Therefore, we get

$$\begin{aligned} d(x,y) &= \\ &= \sum_{S \subset \phi(X)} \alpha_S \Big(d_S(\phi(x),\phi(y)) + d_S(\phi(x),0) - d_S(\phi(y),0) \Big) \\ &= \sum_{S \subset \phi(X)} 2\alpha_S \delta_S(\phi(x),\phi(y)) = \sum_{S \subset X} 2\alpha_{\phi^{-1}(S)} \delta_S(x,y), \end{aligned}$$

which concludes the proof.

A natural question to ask is: If a metric space (X, d_X) is embeddable into a metric (Y, d_Y) with small distortion, is the corresponding directed semimetric space $(X, d_{X,s})$ embeddable into $(Y, d_{Y,t})$ (where $s \in X, t \in Y$) with small distortion?

Lemma 2.2. *Suppose a metric space (X, d_X) is isometrically embeddable into a metric space (Y, d_Y). Let $\phi : X \hookrightarrow Y$ be an isometric embedding. Then ϕ is an isometric embedding of the corresponding directed semimetric space $(X, d_{X,s})$ into $(Y, d_{Y,\phi(s)})$. Hence $(X, d_{X,s})$ is isometrically embeddable into $(Y, d_{Y,\phi(s)})$.*

Proof. We prove that ϕ is an isometric embedding of $(X, d_{X,s})$ into $(Y, d_{Y,\phi(s)})$. Indeed, $d_{Y,\phi(s)}(\phi(x),\phi(y)) = d_Y(\phi(x),\phi(y)) + d_Y(\phi(x),\phi(s)) - d_Y(\phi(y),\phi(s)) = d_X(x,y) + d_X(x,s) - d_X(y,s) = d_{X,s}(x,y)$.

This lemma yields the following result:

Corollary 2.1. *For every $p \in [1,2]$, the directed semimetric space ℓ_p is directed ℓ_1-embeddable. Directed ℓ_1 is directed ℓ_2^2-embeddable.*

However, it turns out that the result of Lemma 2.2 is very fragile: In section 6 we will construct a directed ℓ_2^2 metric that embeds into directed ℓ_1 with distortion at least n^δ (where n is the size of the directed ℓ_2^2 space, and $\delta > 0$ is a constant), but the underlying (undirected) ℓ_2^2 metric space embeds into (undirected) ℓ_1 with constant distortion.

Finally, another class of directed semimetrics, which we consider, is the class of convex combinations of directed 0–1 semimetrics, i.e. such semimetrics where every distance is either 0 or 1. In particular, since the directed cut metric is a directed 0–1 semimetric, every directed ℓ_1-embeddable semimetric is a convex combination of directed 0–1 semimetrics. Note, that in the undirected case the corresponding classes of metrics, ℓ_1-embeddable metrics and convex combinations of 0–1 metrics, coincide. However, in the directed case these two classes are very different.

THEOREM 2.1. *The least distortion with which every convex combination of directed* 0–1 *semimetrics on* n *points can be embedded into directed* ℓ_1 *is* $\Omega(n)$.

Proof. We construct a directed 0–1 semimetric space (X, d) not embeddable into directed ℓ_1 with distortion less then $\Omega(n)$. Put $n = 2k+2$. Let $X = \{s, t, a_i, b_i | 1 \leq i \leq k\}$. The distance function d is defined by:

- for every $x \in X$, $d(x, s) = d(t, x) = 0$;
- for every $i \neq j \in \{1, \ldots, k\}$, $d(b_i, a_j) = 0$;
- all other distances are equal to 1.

It is easy to verify that (X, d) is a directed semimetric. Assume that (X, d) embeds into directed ℓ_1 with distortion D. Consider the distance d' induced by the embedding ϕ of X into directed ℓ_1: $d'(x_1, x_2) = d_1(\phi(x_1), \phi(x_2))$, then

$$c\, d(x_1, x_2) \leq d'(x_1, x_2) \leq c\, D\, d(x_1, x_2).$$

Let us represent d' a convex combination of directed cut metrics $d' = \sum \alpha_S \delta_S$. Consider a cut[5] S that appears with a non-zero coefficient in this convex combination. Note that if $d(x_1, x_2) = 0$, then $d'(x_1, x_2) = 0$, hence $\delta_S(x_1, x_2) = 0$. From this fact and the definition of d, we get

- $s \in S$, $t \notin S$ (since for every x, $d(x, s) = d(t, x) = 0$);
- if $a_i \notin S$ then $b_j \notin S$ for every $j \neq i$ (since for every $i \neq j$, $d(b_i, a_j) = 0$);

[5] S is a proper subset of X

From the first item, we conclude that each cut S contributes α_S to the distance between s and t: $d'(s, t) = \sum_S \alpha_S$. From the second item, we conclude that each cut S contributes to at most one distance between b_i and a_i. Indeed, otherwise we would have that for some $i \neq j$, $a_i, a_j \notin S$, $b_i, b_j \in S$, which would contradict to the second item. Hence, $\sum_{i=1}^k d'(a_i, b_i) \leq \sum_S \alpha_S$.

We obtain

$$\begin{aligned} c\,k = c \sum_{i=1}^{k} d(b_i, a_i) \leq \sum_{i=1}^{k} d'(b_i, a_i) \leq \sum_S \alpha_S \\ = d'(s, t) \leq c\, D\, d(s, t) = c\, D. \end{aligned}$$

We conclude that $D \geq k = \frac{n-2}{2} = \Omega(n)$.

3 Directed Partitioning Problems

In this section we introduce several graph partitioning problems. In the next section we will show a relation between these problems and embedding questions.

Let us start with the Directed Sparsest Cut Problem. There are several non-equivalent ways to extend the definition of the Directed Sparsest Cut Problem from the undirected to the directed case. We call the first variant Bipartite Directed Sparsest Cut.

DEFINITION 3.1. (BIPARTITE DIRECTED SPARSEST CUT) *Let* $G = (V, E)$ *be a directed graph. We have* m *source terminal pairs* (s_k, t_k) $(1 \leq k \leq m)$. *Each edge* $e \in E$ *has capacity* cap_e, *each source terminal pair has demand* dem_i. *Our goal is to divide the graph into two parts* S *and* $T = V \setminus S$ *so as to minimize the ratio of the total capacity of cut edges to the separated demand:*

$$\sum_{\substack{(i,j) \in E \\ i \in S,\ j \in T}} \text{cap}_{(i,j)} \Bigg/ \sum_{k: s_k \in S,\ t_k \in T} \text{dem}_k.$$

A version of this problem with uniform demands was studied by Leighton and Rao [15], who presented an $O(\log n)$ approximation algorithm, and then by Agarwal, Charikar, Makarychev, and Makarychev [1], who found an $O(\sqrt{\log n})$ approximation algorithm. In Section 5 we will show that the version of this problem with non-uniform demands cannot be approximated within $O(2^{(\log n)^\delta})$ or $O(n^\delta)$ (depending on the complexity assumptions) in polynomial time.

In another variant of Directed Sparsest Cut Problem we have the same graph G, capacities, and demands, but our objective function is different.

DEFINITION 3.2. (DIRECTED SPARSEST CUT) *In this variant of the problem, our goal is to remove a set of edges* A *so as to minimize the ratio of the total capacity* *of cut edges to the separated demand:*

$$\sum_{(i,j)\in A} \text{cap}_{(i,j)} \Bigg/ \sum_{\substack{k\,:\,\text{there is no path}\\ \text{from } s_k \text{ to } t_k \text{ in } G-A}} \text{dem}_k.$$

Hajiaghayi and Räcke gave an $O(\sqrt{n})$ approximation algorithm for this problem [12]. Now we mention the Directed Multicut Problem, which is closely related to the Directed Sparsest Cut Problem.

Definition 3.3. (Directed Multicut Problem) *Let $G = (V, E)$ be a directed graph. We have m source terminal pairs (s_k, t_k) $(1 \le k \le m)$. Each edge $e \in E$ has capacity cap_e. Our goal is to remove a set of edges of minimal capacity so that every source terminal pair is separated.*

Cheriyan, Karloff, and Rabani gave an $O(\sqrt{n}\log m)$ approximation algorithm for this problem [6]. Then Gupta found an $O(\sqrt{n})$ approximation [10]. It turns out that given an α approximation for the Directed Multicut Problem, one can get an $\alpha \log D$ approximation for the Directed Sparsest Cut Problem [11] (where D is the total demand; without loss of generality we may assume that D is polynomial in n, hence $\log D = O(\log n)$); given an α approximation for the Directed Sparsest Cut Problem, one can get an $O(\alpha \log n)$ approximation for the Directed Multicut Problem using a divide-and-conquer approach.

4 Embeddings and Integrality Gaps

4.1 Bipartite Directed Sparsest Cut

Recall that in the undirected case the integrality gap of an SDP relaxation for the Sparsest Cut with non-uniform demands is equal to the minimal distortion with which every ℓ_2^2 space embeds into ℓ_1. In this section we establish a counterpart of this fact for the directed case.

First we state the Directed Sparsest Cut Problem as follows: *Find a directed cut metric $d = \delta_S$ that minimizes the objective function*

$$\sum_{(i,j)\in E} \text{cap}_{(i,j)}\, d(i,j) \Bigg/ \sum_k \text{dem}_k\, d(s_k, t_k). \tag{4.1}$$

By Lemma 2.1 any directed ℓ_1-embeddable metric is a convex combination of directed cut metrics. Therefore if we assume that d is a directed ℓ_1 embeddable metric, the objective function (4.1) will not change.

Now to approximate the directed cut metric it seems natural to replace the constraint that d is directed ℓ_1 embeddable with the constraint that it is a directed ℓ_2^2 semimetric. Then we obtain the following SDP relaxation:

$$\min \frac{1}{8} \sum_{(i,j)\in E} \text{cap}_{(i,j)}\, d(i,j)$$

where $d(i,j) = |v_i - v_j|^2 + |v_i - v_0|^2 - |v_j - v_0|^2$

$$\sum_k \text{dem}_k\, d(s_k, t_k) = 1$$

$$|v_i - v_k|^2 \le |v_i - v_j|^2 + |v_j - v_k|^2 \quad \forall\, i,j,k \in V \cup \{0\}$$

here we assign a vector v_i to each vertex i, and we introduce one additional vector v_0. In the intended solution the vector v_0 corresponds to the part T; and the vector $-v_0$ corresponds to the part S. In fact in [1] the same SDP relaxation was used to find an $O(\sqrt{\log n})$ approximation for the Bipartite Directed Sparsest Cut Problem with uniform demands.

The integrality gap of this SDP is closely connected to the the minimum distortion achievable for the embedding of a directed ℓ_2^2 semimetric into directed ℓ_1.

Theorem 4.1. *Denote the integrality gap of the SDP for the Bipartite Directed Sparsest Cut Problem on the complete graph on n vertices (for the worst choice of the demands) by $Gap_{BDSC}(n)$. Denote the minimal distortion with which every directed ℓ_2^2 semimetric embeds into directed ℓ_1 by $Distort_{\ell_2^2\to\ell_1}(n)$. Then*

$$Distort_{\ell_2^2\to\ell_1}(n) = Gap_{BDSC}(n).$$

Proof. The proof is exactly the same as in the undirected case. See Appendix A for the details.

4.2 Directed Sparsest Cut and Directed 0–1 Metrics

Consider the following natural LP relaxation for the Directed Sparsest Cut problem.

$$\min \sum_{e\in E} \text{cap}_e\, x_e$$

$$\sum_k \text{dem}_k\, dist(s_k, t_k) \ge 1$$

$$dist(i,j) + x(j,k) \ge dist(i,k) \quad \forall i,j,k \in V, (j,k) \in E$$

$$x(i,j) \ge 0 \quad \forall (i,j) \in E$$

$$dist(i,j) \ge 0 \quad \forall i,j \in V$$

We show that the integrality gap of this LP is closely linked to the minimal distortion.

Theorem 4.2. *Denote the integrality gap of the LP for the Directed Sparsest Cut Problem with non-uniform Demands on the complete graph on n vertices (for the worst choice of the demands) by $Gap_{DSC}(n)$. Denote the minimal distortion with which every directed semimetric embeds into a convex combination of 0–1 semimetrics by $Distort_{any\to\{0,1\}}(n)$. Then*

$$Gap_{DSC}(n) = Distort_{any\to\{0,1\}}(n).$$

Proof. See Appendix A for the proof.

Hajiaghayi and Räcke showed that the LP integrality gap is $O(\sqrt{n})$ [12]. This yields the following corollary.

Corollary 4.1. *Every directed semimetric can be embedded into a convex combination of 0–1 semimetrics with distortion $O(\sqrt{n})$.*

5 Hardness of Bipartite Directed Sparsest Cut

In this section we reduce the Maximum Edge Bipartite Clique Problem to the Directed Sparsest Cut Problem with non-uniform demands. Recently Feige and Kogan [8] showed that this problem cannot be approximated within $O(2^{(\log n)^\delta})$ for some $\delta > 0$ in polynomial time unless 3-SAT can be solved in time $2^{n^{3/4+\varepsilon}}$ for every $\varepsilon > 0$. They also conjectured that the problem cannot be solved for $O(n^\delta)$ for some $\delta > 0$. The conjecture was proved in [7] under a stronger complexity assumption. Therefore, the result of this section implies that the Directed Sparsest Cut Problem with non-uniform demands also cannot be approximated within $O(2^{(\log n)^\delta})$ or $O(n^\delta)$ (depending on the complexity assumptions) in polynomial time.

Definition 5.1. (Maximum Edge Bipartite Clique Problem) *Given a bipartite graph on the set of vertices (L, R), find a bipartite clique $(A, B) : A \subset L, B \subset R$ that maximizes the value of $|A||B|$.*

Theorem 5.1. *Let G be a bipartite graph on the set of vertices (L, R). We construct a directed graph H as follows. The set of vertices of H is $L \cup R \cup \{s, t\}$ (where s and t are two new vertices). Every pair of vertices $l \in L$ and $r \in R$ non-adjacent in G are connected with an edge of infinite (or sufficiently large) capacity in H. All vertices are connected to s with an edge of infinite capacity; t is connected to all vertices with an edge of infinite capacity; s is connected to t with an edge of capacity 1. There are no other edges in H. The demand between every $l \in L$ and every $r \in R$ is 1, all other demands are 0.*

Denote the optimum value of the Bipartite Directed Sparsest Cut Problem on the graph H by OPT_{BDSC}, and the optimum value of the Maximum Edge Bipartite Clique Problem by OPT_{MBC}. Then $OPT_{MBC} = OPT_{BDSC}^{-1}$.

Proof. First, let us prove that $OPT_{MBC} \le OPT_{BDSC}^{-1}$. Consider an optimum bipartite clique $(A, B) : A \subset L, B \subset R$. Let

$$S = \{s\} \cup A \cup (R \setminus B)$$
$$T = \{t\} \cup B \cup (L \setminus A)$$

Compute the ratio of the cut capacity to the separated demand. The only cut edge is (s, t). The cut separates all pairs of vertices $(a, b) : a \in A, b \in B$. Therefore the ratio is $\frac{1}{|A||B|} = 1/OPT_{MBC}$. We conclude that $OPT_{MBC} \le OPT_{BDSC}^{-1}$.

Now we prove the other direction: $OPT_{MBC} \ge OPT_{BDSC}^{-1}$. Let $(S, T \equiv \bar{S})$ be an optimum directed balanced cut. Since the capacity of the cut is finite no edge of infinite capacity is cut, this means that

- $s \in S$, and $t \in T$;
- for every vertices $a \in S \cap L$ and $b \in T \cap R$, (a, b) is not an edge of H, i.e. (a, b) is an edge in the graph G.

The first item implies that the capacity of the cut is 1. Therefore the separated demand equals $1/OPT_{BDSC} = |S \cap L| \times |T \cap R|$. The second item implies that $(S \cap L, T \cap R)$ is a bipartite clique. We have found a bipartite clique with $1/OPT_{BDSC}$ edges. This concludes the proof.

6 Lower bound for embedding directed ℓ_2^2 metrics into directed ℓ_1

In this section we show that the integrality gap of an SDP relaxation for the Maximum Edge Bipartite Clique Problem is $\Omega(n^\delta)$. What is a natural SDP relaxation for the problem? We can apply the reduction of Theorem 5.1 to the SDP relaxation for the Bipartite Directed Sparsest Cut, and obtain an SDP relaxation for the Maximum Edge Bipartite Clique:

$$\min \frac{8}{d(s,t)}$$

$$d(i, s) = 0 \;\text{ since edge } (i, s) \text{ of } H \text{ has infinite capacity;}$$
$$d(t, i) = 0 \;\text{ since edge } (t, i) \text{ of } H \text{ has infinite capacity;}$$
$$d(l, r) = 0 \;\; \forall\, l \in L, r \in R,\ l \text{ is not connected to } r \text{ in } G;$$
$$\sum_{l \in L, r \in R} d(l, r) = 1$$

Now we throw in additional constraints (note that the SDP becomes stronger): we require that $v_s = -v_0$, $v_t = v_0$, $|v_i| = |v_0|$. We also require that not only vectors v_i but all vectors $\pm v_i$ satisfy triangle inequalities w.r.t. the squared Euclidean distance. Finally we rescale the vectors so that all the vectors are unit vectors, and put $a_l = v_l$ for $l \in L$, and $b_r = -v_r$ for $r \in R$.

We get the following nice SDP relaxation for the Maximum Edge Bipartite Clique Problem:

$$\text{maximize } \frac{1}{4}\sum_{ij}\langle v_0 - a_i, v_0 - b_j\rangle$$

$\langle v_0 - a_i, v_0 - b_j\rangle = 0$ if i is not connected to j

vectors $\pm v_0$, $\pm a_i$, and $\pm b_j$ are unit vectors satisfying triangle inequalities w.r.t the squared Euclidean distance.

In the intended solution $a_i = -v_0$ if i belongs to the bipartite clique, and $a_i = v_0$, otherwise. Similarly, $b_j = -v_0$ if j belongs to the bipartite clique, and $b_j = v_0$, otherwise. We prove that the integrality gap of this relaxation is n^δ for some $\delta > 0$.

Construction 6.1. *Consider the following instance of the Maximum Edge Bipartite Clique Problem. Let C be the set of vertices of the d-dimensional hypercube $\{\pm\frac{1}{\sqrt{d}}\}^d$ that have an even number of entries equal to $\frac{1}{\sqrt{d}}$. Let G be a bipartite graph on $(L, R) = (C, C)$; two vertices $l \in L$ and $r \in R$ are connected if $\langle l, r\rangle \neq -\gamma$, where $\gamma \in (0, 1)$ is a fixed constant.*

Lemma 6.1. *Let $\gamma = \gamma(d)$ be a number in the interval $(0, \gamma_0)$, $\gamma_0 < 1$, s.t. $-\gamma$ is the inner product of some vectors from C (i.e. γ is of the form $\frac{d-4r}{d}$). Then the number of edges in the maximum bipartite clique (L_*, R_*) of G is at most $(2 - \varepsilon)^{2d}$ (for some constant $\epsilon > 0$ that depends only on γ_0).*

Proof. This is an easy consequence of one of the results of Frankl and Rödl [9, Theorem 1.4]. See Appendix B for the proof.

Theorem 6.1. *The value of the SDP solution of the problem is $\Omega(4^d)$.*

Proof. Let $\alpha = \frac{2\gamma^2}{(1+\gamma)^2}$ (it will become clear later why we chose α this way; for now you may think of α as an arbitrary constant between 0 and 2/3).

Let

$$G(u) = \sqrt{\frac{2\alpha - \alpha^2}{1 + 2\gamma}}\left[\sqrt{2\gamma}u \oplus u \otimes u\right].$$

Define a feasible SDP solution by

$$a_i = b_i = F(u) \equiv (1 - \alpha)v_0 \oplus G(u).$$

where a_i and b_i are vectors corresponding to the vertex $u \in C$. Now we verify that vectors from the set $S \equiv \{a_i, b_i\}$ form a feasible solution. They are unit vectors: $\|(1 - \alpha)v_0 \oplus \sqrt{\frac{2\alpha-\alpha^2}{1+2\gamma}}\left[\sqrt{2\gamma}u \oplus u \otimes u\right]\|_2^2 = (1 - \alpha)^2 + \frac{2\alpha-\alpha^2}{1+2\gamma}(1 + 2\gamma) = 1$.

We will need the following inequality below

$$\langle v_0 - F(u), v_0 - F(v)\rangle \geq 0. \tag{6.2}$$

Indeed, we have $\langle v_0 - F(u), v_0 - F(v)\rangle = \alpha^2 + \frac{2\alpha-\alpha^2}{1+2\gamma}\left[2\gamma\langle u, v\rangle + \langle u, v\rangle^2\right] \geq \alpha^2 + \frac{2\alpha-\alpha^2}{1+2\gamma}\left[-\gamma^2\right] = \alpha\left[\frac{2\gamma^2}{(1+\gamma)^2} - \frac{2-\frac{2\gamma^2}{(1+\gamma)^2}}{1+2\gamma}\gamma^2\right] = 0$.

Moreover, if $\langle u, v\rangle = -\gamma$, then the inequality above becomes an equality. The latter fact actually implies that the first SDP constraint is satisfied.

Now we verify that vectors satisfy triangle inequalities w.r.t. the squared Euclidean distance. Clearly, vertices $\{\pm G(u)\}$ satisfy triangle inequalities, since they are vertices of a multidimensional parallelepiped. So all the vertices from S satisfy triangle inequalities. Therefore, we have to consider the following cases (we omit the cases that can be obtained from the cases below by reflection in the origin):

1) one of the vertices is v_0; two others are in S;
2) one of the vertices is v_0; two others are in $-S$;
3) two of the vertices are v_0 and $-v_0$, and the third is in S;
4) one of the vertices is v_0; the second is in S, and the third is in $-S$;
5) two of the vertices are in S, and the third is in $-S$.

Lemma 6.2. *Triangle inequality is satisfied in all the cases above.*

Proof.

1) Obviously the triangle with these three vertices is isosceles. From inequality (6.2) we know that the angle at the vertex v_0 is not obtuse.

2) Similarly, the triangle with these three vertices is isosceles. The angle at the vertex v_0 is clearly acute.

3) The triangle with these vertices is a right triangle. Therefore all its angles are not obtuse.

4) Let the vertices be v_0, $F(u)$, and $-F(v)$ correspondingly. In the triangle with vertices v_0, $F(u)$, and $-F(v)$, the side from v_0 to $F(u)$ is shorter than the side from v_0 to $-F(v)$. Hence the angle at the vertex $-F(v)$ is smaller that one at the vertex $F(u)$. It suffices to check that angles at the vertices v_0 and $F(u)$ are not obtuse. We have,

$$\langle v_0 - F(u), v_0 + F(v)\rangle = \alpha(2 - \alpha) - \underbrace{\langle G(u), G(v)\rangle}_{\text{is maximal when } u = v}$$
$$\geq \alpha(2 - \alpha) - (1 - (1 - \alpha)^2) = 0$$

The other angle is not obtuse as well: $\langle F(u)-v_0, F(u)+F(v)\rangle = \langle F(u)-v_0, F(u)+v_0\rangle + \langle F(u)-v_0, F(v)-v_0\rangle$. The first term in the RHS is equal to zero, the second term is non-negative by inequality (6.2).

5) Let the vertices be $F(u)$, $F(v)$, and $-F(w)$ correspondingly. We have

$$\langle F(u)-F(v), F(u)+F(w)\rangle = \langle G(u)-G(v), G(u)+G(w)\rangle$$

We know that vectors $\{\pm G(x)\}$ satisfy triangle inequalities. So the RHS is non-negative. Now consider the last remaining case:

$$\begin{aligned}&\langle F(u)+F(w), F(v)+F(w)\rangle = 1 + \langle F(u), F(w)\rangle + \\ &\langle F(v), F(w)\rangle + \langle F(u), F(v)\rangle \geq 1 + 3\min_{x,y\in C}\langle F(x), F(y)\rangle \\ &= 1 + 3\min_{x,y\in C}(\langle v_0 - F(x), v_0 - F(y)\rangle - \alpha^2 + (1-\alpha)^2) \\ &= 4 - 6\alpha\end{aligned}$$

Here we use that $\min_{x,y\in C}\langle v_0 - F(x), v_0 - F(y)\rangle \geq 0$ by inequality (6.2). Since $\alpha < 2/3$, the quantity $4-6\alpha$ is positive.

We have shown that our solution is a feasible solution of the SDP.

Finally let us estimate the value of the objective function. For two random vertices u, and v from C we have: $E[\langle v_0 - F(u), v_0 - F(v)\rangle] = \alpha^2 + E[\langle G(u), G(v)\rangle] > \alpha^2$. Therefore the value of the objective function is at least $\alpha^2|C|^2 = \alpha^2 2^{2d-2}$

We have shown that the integrality gap of the SDP relaxation for the Maximum Edge Bipartite Clique Problem is $\Omega(n^\delta)$. This implies that the integrality gap of the SDP relaxation for the Bipartite Directed Sparsest Cut Problem is at least $\Omega(n^\delta)$, and therefore the least distortion $Distort_{\ell_2^2\to\ell_1}(n)$ for the embedding of a directed ℓ_2^2-semimetric into directed ℓ_1 is at least $\Omega(n^\delta)$.

7 Weak Embeddings

In this section we introduce a new notion of weak embeddability, and prove a directed counterpart of Bourgain's theorem for directed semimetrics.

Definition 7.1. *We say that a mapping ϕ from a directed semimetric space (X, d_X) to a directed semimetric space (Y, d_Y) is a weak embedding with distortion D if there is a scale factor $c > 0$ s.t. for every $x, y \in X$ we have*

- *if $d_X(x,y) \geq d_X(y,x)$ then $\frac{c}{D}\, d_X(x,y) \leq d_Y(\phi(x),\phi(y)) \leq c\, d_X(x,y)$;*
- *if $d_X(x,y) \leq d_X(y,x)$ then $d_Y(\phi(x),\phi(y)) \leq c\, d_X(x,y)$.*

We say that a directed semimetric space (X, d_X) is weakly embeddable to a directed semimetric space (Y, d_Y) with distortion D, if there exists a weak embedding with distortion D.

Definition 7.2. *A directed semimetric space (X, d) is a weighted directed semimetric space if there exists a weight function $w: X \to \mathbb{R}$ s.t. $d(x,y) = d_*(x,y) + w(x) - w(y)$, where $d_*(x,y)$ is a (undirected) semimetric.*

This definition is a slight generalization of Definition 2.2. In Definition 2.2 we only consider weight functions $w(x)$ of the form $d(x,s)$. The following theorem by Vitolo [18] gives an alternative characterization.

Theorem 7.1. (Vitolo) *Let (X, d) be a directed semimetric space. The following two conditions are equivalent: 1) (X, d) is a weighted directed semimetric space; 2) for every x, y, and z we have $d(x,y)+d(y,z)+d(z,x) = d(x,z)+d(z,y)+d(y,x)$.*

Now we are ready to state our embedding result.

Theorem 7.2. *Every weighted directed semimetric (X, d) on n points weakly embeds into directed ℓ_1 with distortion at most $O(\log n)$.*

The proof of this theorem is given in Appendix C.

Remark. The condition that (X, d) is weighted is important. Indeed, any weak embedding of a directed 0–1 semimetric with distortion D is also a (regular) embedding with the same distortion. Therefore, the directed 0–1 semimetric presented in the proof of Theorem 2.1 does not embed into directed ℓ_1 with distortion less than $\Omega(n)$.

Acknowledgements

We would like to thank Uri Feige for pointing us to [8].

References

[1] Amit Agarwal, Moses Charikar, Konstantin Makarychev, and Yury Makarychev, $O(\sqrt{\log n})$ approximation algorithms for Min UnCut, Min 2CNF Deletion, and directed cut problems. In Proc. 37th Annual ACM Symposium on Theory of Computing (STOC), pages 573-581, 2005.

[2] S. Arora, J. R. Lee, and A. Naor, Euclidean distortion and the sparsest cut. In Proc. 37th STOC, pages 553-562, 2005,

[3] S. Arora, S. Rao, and U. Vazirani, Expander flows,geometric embeddings and graph partitioning. In Proc. 36th STOC, pages 222-231, 2004.

[4] J. Bourgain, On Lipschitz embedding of finite metric spaces in Hilbert space. Israel J. Math., 52:46–52, 1985.
[5] S. Chawla, A. Gupta, and H. Räcke, Embeddings of negative-type metrics and an improved approximation to generalized sparsest cut. In Proc. 16th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), pages 102-111, 2005.
[6] Joseph Cheriyan, Howard Karloff, and Yuval Rabani, Approximating Directed Multicuts. In Proc. 42nd Annual Symposium on Foundations of Computing (FOCS), pages 320-328, 2001.
[7] Uriel Feige, Relations between average case complexity and approximation complexity. In Proc. 34th STOC, pages 534-543, 2002.
[8] Uriel Feige and Shimon Kogan, Hardness of approximation of the Balanced Complete Bipartite Subgraph problem. Technical Report MCS04-04, Department of Computer Science and Applied Math., Weizmann Institute.
[9] P. Frankl and V. Rödl, Forbidden intersections. *Trans. Amer. Math. Soc.*, 300:259–286, 1987.
[10] Anupam Gupta, Improved results for directed multicut. In Proc. 14th SODA, pages 454-455, 2003.
[11] M. Hajiaghayi and F. Leighton, On the max-flow min-cut ratio for directed multicommodity flows. Tech Report MIT-LCS-TR-910, 2003.
[12] Mohammad T. Hajiaghayi and Harald Räcke, An $O(\sqrt{n})$-Approximation Algorithm For Directed Sparsest Cut. To appear in IPL.
[13] P. Indyk and J. Matousek, Low Distortion Embeddings of Finite Metric Spaces, to appear in Handbook of Discrete and Computational Geometry (2nd edition) J. E. Goodman and J. O'Rourke, editors. CRC Press LLC.
[14] S. Khot and N. Vishnoi, The Unique Games Conjecture, Integrality Gap for Cut Problems and Embeddability of Negative Type Metrics into ℓ_1, To appear in Proc. 46th FOCS, 2005.
[15] F. T. Leighton and S. Rao, An Approximate Max-Flow Min-Cut Theorem for Uniform Multicommodity Flow Problems with Applications to Approximation Algorithms. In Proc. 29th FOCS, pages 422-431, 1988.
[16] N. Linial, E. London and Yu. Rabinovich, The geometry of graphs and some of its algorithmic applications Combinatorica (1995) 15, pp. 215-245
[17] J. Matousek. Embedding finite metric spaces into Euclidean spaces, chapter in *Lectures on discrete geometry*. Graduate Texts in Mathematics, Vol. 212, Springer Verlag, 2002.
[18] Paolo Vitolo, The representation of weighted quasimetric spaces. Rend. Istit. Mat. Univ. Trieste 31 (1999), no. 1-2, 95-100.

A Proof of Theorems 4.1 and 4.2

We will prove a general theorem that implies Theorems 4.1 and 4.2.

Let $\mathcal{A}$ and $\mathcal{B}$ be two closed convex cones of semimetrics on a set X. Denote the minimum distortion D with which every semimetric $d_A \in \mathcal{A}$ can be approximated by a semimetric $d_B \in \mathcal{B}$ by $Distort_{\mathcal{A}\to\mathcal{B}}$, i.e. the least D s.t. $\forall\, d_A \in \mathcal{A}\ \exists d_B \in \mathcal{B}\ d_B(x,y) \le d_A(x,y) \le D\, d_B(x,y)$.

Given an instance $\mathcal{I}$ of the Sparsest Cut Problem: graph $G = (X, E)$ with edge capacities cap_e, and demands dem_i, we define the optimum value $OPT_C(\mathcal{I})$ of the problem in the class of semimetrics $\mathcal{C}$, as

$$OPT_{\mathcal{C}}(\mathcal{I}) = \min_{d\in\mathcal{C}} \frac{\sum_{(x,y)\in E} \text{cap}_{(x,y)}\, d(x,y)}{\sum_i \text{dem}_i\, d(s_i,t_i)}.$$

Then the gap between two classes $\mathcal{A}$ and $\mathcal{B}$ is defined as

$$Gap_{\mathcal{A},\mathcal{B}} = \sup_{\mathcal{I}\text{ is an instance on }X} \frac{OPT_{\mathcal{B}}(\mathcal{I})}{OPT_{\mathcal{A}}(\mathcal{I})}.$$

THEOREM A.1. *For two closed cones of semimetrics $\mathcal{A}$, $\mathcal{B}$ on a set X, we have the identity $Distort_{\mathcal{A}\to\mathcal{B}} = Gap_{\mathcal{A},\mathcal{B}}$.*

Remark. This theorem applied to the cone of directed ℓ_2^2 semimetrics and the cone of directed ℓ_1 embeddable semimetrics yields Theorem 4.1; applied to the cone of all semimetrics and the cone generated by directed 0–1 semimetric yields Theorem 4.2.

Proof. [Proof of Theorem A.1]

The distortion $Distort_{\mathcal{A}\to\mathcal{B}}$ is equal to the maximum over all $d_A \in \mathcal{A}$ of the value of the following LP:

$$\begin{aligned}
&\min D\\
&\quad d_B \in \mathcal{B}\\
&\quad D d_A(x,y) - d_B(x,y) \ge 0 \qquad \forall\, x,y \in X\\
&\quad d_B(x,y) \ge d_A(x,y) \qquad \forall\, x,y \in X\\
&\text{minimization is over variables } d_B, \text{ and } D
\end{aligned}$$

The value of this LP equals the value of the dual LP:

$$\begin{aligned}
&\max \sum_{x,y} \beta_{x,y} d_A(x,y)\\
&\quad -\sum_{x,y} \alpha_{x,y} d(x,y) + \beta_{x,y} d(x,y) \le 0 \qquad \forall\, d \in \mathcal{B}\\
&\quad \sum_{x,y} \alpha_{x,y} d_A(x,y) \le 1
\end{aligned}$$

(minimization is over variables $\alpha_{x,y}$, and $\beta_{x,y}$)

We get, that $Distort_{\mathcal{A}\to\mathcal{B}}$ equals

$$\max_{d_A\in\mathcal{A}} \max_{\substack{\alpha_{x,y}\\ \beta_{x,y}}} \left(\frac{\sum \beta_{x,y} d_A(x,y)}{\sum \alpha_{x,y} d_A(x,y)} \Big/ \max_{d_B\in\mathcal{B}} \frac{\sum \beta_{x,y} d_B(x,y)}{\sum \alpha_{x,y} d_B(x,y)} \right)$$

$$= \max_{\substack{\alpha_{x,y}\\ \beta_{x,y}}} \left(\max_{d_A\in\mathcal{A}} \frac{\sum \beta_{x,y} d_A(x,y)}{\sum \alpha_{x,y} d_A(x,y)} \Big/ \max_{d_B\in\mathcal{B}} \frac{\sum \beta_{x,y} d_B(x,y)}{\sum \alpha_{x,y} d_B(x,y)} \right)$$

$$= \max_{\substack{\alpha_{x,y}\\ \beta_{x,y}}} \left(\min_{d_B\in\mathcal{B}} \frac{\sum \alpha_{x,y} d_B(x,y)}{\sum \beta_{x,y} d_B(x,y)} \Big/ \min_{d_A\in\mathcal{A}} \frac{\sum \alpha_{x,y} d_A(x,y)}{\sum \beta_{x,y} d_A(x,y)} \right)$$

(all the summations are over x and y)
Now we interpret values of $\alpha_{x,y}$ as capacities, and values of $\beta_{x,y}$ as demands, and get the statement of the theorem.

B Proof of Lemma 6.1

Lemma 6.1 *Let $\gamma = \gamma(d)$ be a number in the interval $(0, \gamma_0)$, $\gamma_0 < 1$, s.t. $-\gamma$ is the inner product of some vectors from C (i.e. γ is of the form $\frac{d-4r}{d}$). Then the number of edges in the maximum bipartite clique (L_*, R_*) of G is at most $(2-\varepsilon)^{2d}$ (for some constant $\epsilon > 0$ that depends only on γ_0).*

Proof. This is an easy consequence of one of the results of Frankl and Rödl [9, Theorem 1.4]:

Theorem B.1. (Frankl and Rödl) *Suppose $0 < \eta < \frac{1}{4}$ and two families $\mathcal{F}, \mathcal{G} \subset 2^X$ are given which satisfy $|F \cap G| \neq l$ for $F \in \mathcal{F}$, $G \in \mathcal{G}$. If $\eta d \leq l \leq (\frac{1}{2} - \eta)d$, then*

$$|\mathcal{F}|\,|\mathcal{G}| \leq (4 - \varepsilon(\eta))^d$$

where $\varepsilon(\eta)$ is a positive constant depending only on η.

Now we proceed along the lines of the proof of [9, Theorem 1.11]. Let $X = \{1, \ldots, d\}$; $\eta = \frac{1-\gamma_0}{8}$. We prove that $|L_*||R_*| \leq d^2(4-\varepsilon(\eta))^d$ (obviously this easily implies the statement of the theorem). Choose $1 < \alpha \leq d$ for which the set

$$L_\alpha = \{u \in L_* | u \text{ has exactly } \alpha \text{ entries equal to } \frac{1}{\sqrt{d}}\}$$

has maximal size. Likewise choose $1 < \beta \leq d$ for which the set

$$R_\beta = \{v \in R_* | v \text{ has exactly } \beta \text{ entries equal to } \frac{1}{\sqrt{d}}\}$$

has maximal size. Then $|L_\alpha| \geq |L_*|/d$, $|R_\alpha| \geq |R_*|/d$. Now, if $|L_\alpha| < (2 - \varepsilon(\eta)/2)^d$, then $|L_*||R_*| < d^2|L_\alpha||R_\beta| < d^2(2-\varepsilon(\eta)/2)^d \cdot 2^d = d^2(4-\varepsilon(\eta))^d$, and we are done. So below we assume that $|L_\alpha| \geq (2-\varepsilon(\eta)/2)^d$, and similarly $|R_\beta| \geq (2-\varepsilon(\eta)/2)^d$. Therefore, $\alpha = \frac{d}{2} + o(d)$, $\beta = \frac{d}{2} + o(d)$.

Now we are ready to apply Theorem B.1. Pick $l = \frac{\alpha+\beta}{2} - \frac{d(\gamma+1)}{4} = \frac{1-\gamma}{4}d + o(d)$. Clearly the condition of Theorem B.1 holds (for sufficiently large d):

$$(\frac{1}{2} - \eta)d > \frac{3}{8}d > l = \frac{1-\gamma}{4}d + o(d) > \eta d.$$

Define sets $\mathcal{F}$ and $\mathcal{G}$ as follows: $\mathcal{F} = \{F = \{i : u_i = \frac{1}{\sqrt{d}}\} \,|u \in L_\alpha\}$; $\mathcal{G} = \{G = \{i : v_i = \frac{1}{\sqrt{d}}\} \,|v \in R_\beta\}$. Pick $u \in L_\alpha$, $v \in R_\beta$. Now consider the corresponding sets F and G. We have

$$\langle u, v\rangle = \frac{d - 2|F \triangle G|}{d} = \frac{d - 2(\alpha + \beta - 2|F \cap G|)}{d}.$$

Since (L_*, R_*) is a bipartite clique, u and v are adjacent in G. Hence $\langle u, v\rangle \neq -\gamma$. Therefore,

$$|F \cap G| = \frac{\alpha+\beta}{2} + \frac{d(\langle u,v\rangle - 1)}{4} \neq \frac{\alpha+\beta}{2} - \frac{d(\gamma+1)}{4} = l.$$

By Theorem B.1, $|\mathcal{F}||\mathcal{G}| < (4 - \varepsilon(\eta))^d$. Therefore,

$$|L_*|\,|R_*| \leq d^2|L_\alpha|\,|R_\beta| = d^2|\mathcal{F}|\,|\mathcal{G}| < d^2(4 - \varepsilon(\eta))^d,$$

which concludes the proof.

C Proof of Theorem 7.2

Theorem 7.2. *Every weighted directed semimetric (X, d) on n points weakly embeds into directed ℓ_1 with distortion at most $O(\log n)$.*

Proof. We use the following form of Bourgain's theorem.

Theorem C.1. (Bourgain [4]) *For any semimetric space (Y, d_Y) on n points there exists an embedding $\phi : y \mapsto (\phi^1(y), \ldots, \phi^K(y))$ of Y into ℓ_1^K with distortion $D = O(\log n)$, where $K = \Theta((\log n)^2)$ s.t.*

- *$|\phi^i(y_1) - \phi^i(y_2)| \leq d_Y(y_1, y_2)$;*
- *$\|\phi(y_1) - \phi(y_2)\|_1 \geq c \log n \; d_Y(y_1, y_2)$ for some fixed constant $c > 0$.*

We represent the weighted directed semimetric d as $d(x, y) = d_*(x, y) + w(x) - w(y)$. And then we apply Bourgain's theorem to the semimetric space (X, d_*), and get the embedding ϕ. Define a new embedding $\psi : X \hookrightarrow \ell_1^{2K}$ by $\psi^i(x) = \phi^i(x) + w(x)$ for $1 \leq i \leq K$; $\psi^i(x) = -\phi^{i-K}(x) + w(x)$ for $K+1 \leq i \leq 2K$ (here for each dimension i of the original embedding we introduce two dimensions i and $i+K$ to the new embedding). Note that adding a constant to the weight function $w(x)$ does not change the semimetric d, so we assume that $w(x)$ is chosen s.t. all coordinates $\psi^i(x)$ are positive. Let us estimate the distortion of ψ. Observe, that $\psi^i(x_1) \dot{-} \psi^i(x_2) \leq \max(|\phi^i(x_1) - \phi^i(x_2)| + w(x_1) - w(x_2), 0) \leq$

$d_*(x_1, x_2) + w(x_1) - w(x_2) = d(x_1, x_2)$, for $1 \leq i \leq K$; the same bound holds for $i > K$. Therefore, the directed ℓ_1 distance between $\psi(x_1)$ and $\psi(x_2)$ is bounded by

$$d_1(\psi(x_1), \psi(x_2)) = 2\sum_{i=1}^{2K} \psi^i(x_1) \dot{-} \psi^i(x_2) \leq 4K \cdot d(x_1, x_2).$$

Now if $d(x_1, x_2) \geq d(x_2, x_1)$, that is $w(x_1) \geq w(x_2)$, we have

$$\begin{aligned}
&(\psi^i(x_1) \dot{-} \psi^i(x_2)) + (\psi^{i+K}(x_1) \dot{-} \psi^{i+K}(x_2)) \\
&\quad \geq \big((\phi^i(x_1) + w(x_1)) \dot{-} (\phi^i(x_2) + w(x_2))\big) \\
&\qquad + \big((-\phi^i(x_1) + w(x_1)) \dot{-} (-\phi^i(x_2) + w(x_2))\big) \\
&\quad \text{assuming } \phi^i(x_1) \geq \phi^i(x_2) \text{ (the other case is similar)} \\
&\quad \geq (\phi^i(x_1) + w(x_1)) - (\phi^i(x_2) + w(x_2)) \\
&\quad = |\phi^i(x_1) - \phi^i(x_2)| + w(x_1) - w(x_2).
\end{aligned}$$

Finally, we have $d_1(\psi(x_1), \psi(x_2)) = 2\sum_{i=1}^{K}\Big((\psi^i(x_1) \dot{-} \psi^i(x_2)) + (\psi^{i+K}(x_1) \dot{-} \psi^{i+K}(x_2))\Big) \geq 2\sum_{i=1}^{K}\Big(|\phi^i(x_1) - \phi^i(x_2)| + w(x_1) - w(x_2)\Big) \geq 2c\log n \ d_*(x_1, x_2) + 2K(w(x_1) - w(x_2)) \geq 2c\log n \ d(x_1, x_2)$.

We have showed that if $d(x_1, x_2) \geq d(x_2, x_1)$, then $2c\log n \ d(x_1, x_2) \leq d_1(\psi(x_1), \psi(x_2)) \leq 4K \cdot d(x, y)$; if $d(x_1, x_2) \geq d(x_2, x_1)$, then $d_1(\psi(x_1), \psi(x_2)) \leq 4K \cdot d(x, y)$. Therefore, ψ is a weak embedding with distortion at most $\frac{4K}{2c\log n} = O(\log n)$.