Cuts, Trees and ℓ_1-Embeddings of Graphs
(Preliminary Version)

Anupam Gupta* Ilan Newman† Yuri Rabinovich† Alistair Sinclair*

Abstract

Motivated by many recent algorithmic applications, this paper aims to promote a systematic study of the relationship between the topology of a graph and the metric distortion incurred when the graph is embedded into ℓ_1 space. The main results are:

1. *Explicit constant-distortion embeddings of all series-parallel graphs, and all graphs with bounded Euler number. These are thus the first natural families known to have constant distortion (strictly greater than 1). Using the above embeddings, we obtain algorithms to approximate the sparsest cut in such graphs to within a constant factor.*

2. *A constant-distortion embedding of outerplanar graphs into the restricted class of ℓ_1-metrics known as "dominating tree metrics". We also show a lower bound of $\Omega(\log n)$ on the distortion for embeddings of series-parallel graphs into (distributions over) dominating tree metrics. This shows, surprisingly, that such metrics approximate distances very poorly even for families of graphs with low treewidth, and excludes the possibility of using them to explore the finer structure of ℓ_1-embeddability.*

1 Introduction

Let $G = (V, E)$ be an undirected graph. Each assignment of non-negative weights to the edges of G naturally defines a metric space (V, μ),[1] where for each pair of vertices $x, y \in V$, $\mu(x, y) = d_G(x, y)$ is the shortest-path distance between them. We say that the metric μ is *supported on* (or *generated by*) G. Let (S, ρ) be another metric space. An *embedding* of G into (S, ρ) is a mapping $\phi : V \to S$. The *distortion* of ϕ is the smallest value $c \geq 1$ such that

$$d_G(x, y) \;\leq\; \rho(\phi(x), \phi(y)) \;\leq\; c\, d_G(x, y) \qquad \forall x, y \in V.$$

Thus the distortion measures the maximum factor by which any distance is stretched in the embedding.

In recent years, the idea of embedding a graph into a "nice" metric space with low distortion has emerged as a useful ingredient in the design and analysis of algorithms in a variety of domains. "Nice" metric spaces are those with well-studied structural properties, such as Euclidean or ℓ_1 space, or those which are particularly simple, such as trees and distributions over them. A very incomplete list of applications includes approximation algorithms for graph and network problems, such as sparsest cut [23, 2], minimum bandwidth [16, 9], low-diameter decompositions [23], optimal group Steiner trees [18, 12], and online algorithms for metrical task systems and file migration problems [4, 6]. These applications, together with its intrinsic mathematical interest, have made the study of low-distortion embeddings a significant field in its own right.

Most of the embeddings considered in the literature have been for metrics supported on general graphs, and give results that bound the worst-case distortion over all graphs. However, if it is guaranteed that the input graph has a certain structure, it is plausible that better embeddings can be found. This is quite intuitive: it is clear that *any* metric is supported on the complete graph on its points, while only a very limited set of metrics can be generated by weighting the edges of, say, a tree. Thus the complexity (and hence embeddability) of a metric generated by a graph G intrinsically depends on the topology of G. At present, very little is known about this interplay between the topological and metrical properties of the graph; the search for connections between the two is emerging as an intriguing and challenging area.

This paper focuses on this problem, and in particular on the relationship between the topology of graphs and their optimal (or near-optimal) embeddings into ℓ_1 (i.e., real space of arbitrary dimension, endowed with the ℓ_1 met-

*Computer Science Division, UC Berkeley CA 94720-1776. Email: {angup, sinclair}@cs.berkeley.edu. Supported in part by NSF grants CCR-9505448 and CCR-9820951.

†Computer Science Department, University of Haifa, Haifa 31905, Israel. Email: {ilan, yuri}@cs.haifa.ac.il.

[1] More correctly, a *semi-metric* space, since we allow $\mu(x, y) = 0$ even when $x \neq y$.

ric). Embeddings into ℓ_1 have been widely studied, and are of special importance due their connection with finding the *sparsest cut* in multicommodity flow networks, which in turn is a key ingredient in approximate solutions of many other problems in such areas as VLSI layout, network routing and efficient simulations of one network by another [8, 22, 21]. Though finding the exact sparsest cut is a computationally hard problem, efficient approximation algorithms for it can be obtained by embedding a natural metric associated with the optimal multicommodity flow into ℓ_1; the approximation ratio depends essentially on the distortion.

One driving force behind this paper is the intriguing conjecture that any metric supported on a planar graph (henceforth called a *planar metric*) can be embedded into ℓ_1 with constant distortion. Even more generally, we conjecture that this holds for any family of graphs closed under taking minors. There is some evidence to suggest that planar metrics are better behaved than general metrics with respect to ℓ_1-embeddability. In an interesting recent development, Rao [27] has given $O(\sqrt{\log n})$-distortion embeddings of n-point planar metrics into ℓ_1, while the lower bound for general metrics is $\Omega(\log n)$. This result, and the decomposition lemma of [20] on which it is based, attest to the special structure of planar metrics. Further evidence for this is provided by Bartal *et al.*, who have shown that any planar metric can be approximated to within a distortion of $O(\log n)$ by a distribution over dominating tree metrics [7], while the best known upper bound for general metrics is still $O(\log n \log\log n)$ [5].

Despite this promise, current techniques are apparently inadequate to resolve the above conjecture. For embeddings into ℓ_1, a celebrated result of Bourgain [10] tells us that *any* metric supported on an n-vertex graph (i.e., any metric on n points) can be embedded into ℓ_1 with distortion $O(\log n)$; unfortunately, the embedding technique is not sensitive to the topology and incurs a $O(\log n)$-distortion even for the metric generated by the unit-weighted path P_n. Similarly, the method of Bartal *et al.* of finding distributions over dominating trees is limited by a lower bound of $\Omega(\log n)$ for embedding the $n \times n$ grid [1, 7]. Lastly, Rao gives embeddings into ℓ_1 by first embedding into ℓ_2, an approach that is limited by a lower bound of $\Omega(\sqrt{\log\log n})$ for embedding trees into ℓ_2 [11, 24].

In this paper, we systematically explore how the topology of a graph affects the distortion incurred by ℓ_1-embeddings of metrics supported on it. Using the intimate connection between ℓ_1-embeddability of metrics supported on a graph and multicommodity flow problems defined on it, it is possible to show that graphs all of whose metrics are *isometrically* embeddable into ℓ_1 (i.e., embeddable with distortion 1) are exactly the graphs which exclude $K_{2,3}$ as a minor, which essentially corresponds to the class of *outerplanar* graphs. As a natural next step, we consider the family of graphs which have K_4 as an excluded minor. These are graphs with treewidth 2, and essentially correspond to the familiar class of *series-parallel* graphs. Our first main result is an explicit ℓ_1-embedding of these graphs with small constant distortion. This is the first natural family known to have a constant distortion strictly bigger than 1. In addition, our construction implies a simple polynomial time algorithm for finding a sparsest cut within a constant factor of optimal in series-parallel graphs. In a similar vein, we also show that any family of graphs with bounded Euler characteristic can be embedded into ℓ_1 with constant distortion. The technique we use for these results is to explicitly construct a set of *cut metrics* whose sum approximates the original graph metric very closely; cut metrics are trivially embeddable into ℓ_1.

We then go on to study the complementary approach of approximating a metric by a probability distribution over *(dominating) tree metrics*. Since trees are isometrically embeddable into ℓ_1, this gives us another way of embedding graphs into ℓ_1. Furthermore, embeddings based on such metrics have proved easy to work with and possess additional properties that have been exploited in devising approximation algorithms and online algorithms for many problems (see, e.g., [4, 6, 3, 18, 29, 12, 14]). It is natural to ask if we can obtain the above embeddability results for outerplanar and series-parallel graphs using these more restricted metrics. The answers are mixed. On the one hand, we show that this is possible for outerplanar graphs, at a small cost: we give an explicit embedding for such graphs into a distribution over dominating tree metrics with distortion 8 (compared to distortion 1 obtained using cuts). On the other hand, we prove a complementary negative result by exhibiting a family of series-parallel graphs for which *any* distribution over dominating tree metrics must necessarily incur a distortion of $\Omega(\log n)$. This indicates that this approach breaks down at a very early stage (even for graphs of treewidth 2), and suggests that such embeddings offer little hope for exploring the finer structure of ℓ_1-embeddings.

The organization of the paper closely follows the above outline. After a short section containing some definitions and notation, we briefly illuminate the connection between flows and ℓ_1-embeddings in section 3, and then give the embeddings of series-parallel graphs and graphs with small Euler number in section 4. Finally, in section 5, we present our positive and negative results on embeddings into tree distributions. The details of some of the proofs are deferred to the full version of the paper.

2 Definitions and notation

Metrics: In this paper, we shall only consider finite metrics; unless otherwise stated, the number of points will be denoted by n. For more details on many of the concepts used here, see the book of Deza and Laurent [15]. Given two metric spaces, (V, ν) and (W, μ), and a map $f : V \to W$, define the following quantities:

$$\|f\| = \max_{x,y \in V} \frac{\mu(f(x), f(y))}{\nu(x,y)};$$
$$\|f^{-1}\| = \max_{x,y \in V} \frac{\nu(x,y)}{\mu(f(x), f(y))}.$$

Then we say that f has *contraction* $\|f^{-1}\|$, *expansion* $\|f\|$ and *distortion* $D(f) = \|f\| \cdot \|f^{-1}\|$. We say that (V, ν) r-*approximates* (W, μ) (or that the *distortion* between μ and ν is at most r) if there exists a map $f : V \to W$ with $D(f) \leq r$. Often we shall consider two distance functions μ and ν over the same vertex set V. In such cases, we shall assume that f is the identity map. Also, μ will be said to *dominate* ν if for all $x, y \in V$, $\mu(x,y) \geq \nu(x,y)$.

Let $G = (V, E)$ be an undirected graph. A metric (V, μ) is *supported on* G if it is the shortest path metric of G w.r.t. some nonnegative weighting of the edges E. Unless specified otherwise, we shall assume that the edge-weights $w(\cdot)$ satisfy $w(e) = \mu(e)$, where μ is the shortest-path metric of G with weights w.

For a set $S \subseteq V$, the *cut metric* δ_S on V is defined by $\delta_S(x,y) = 1$ if $|S \cap \{x,y\}| = 1$, and 0 otherwise. An important observation is that the ℓ_1-embeddable metrics on V are precisely those which can be written as sums of cut metrics on V with nonnegative coefficients [15].

Multicommodity flows: A *multicommodity flow network* (V, E, P) is specified by an undirected graph $G = (V, E)$, where E is the set of edges along which flow can be routed, and a set $P \subseteq V \times V$ of pairs of vertices of G between which demands can be placed. Assigning nonnegative *capacities* C to the graph edges E and *demands* D to the demand pairs P gives us a particular instance (V, C, D) of the *multicommodity flow* problem. For background, see the survey by Shmoys [28].

The optimal solution to this problem is the maximum value λ such that there is a multicommodity flow f respecting the edge capacities that satisfies a λ-fraction of each demand. We shall refer to λ as *MaxFlow*(V, C, D). Its value (as well as an actual flow f which realizes it) can be found in polynomial time using linear programming.

A closely related problem is the *sparsest cut* problem, which entails finding a partition $(A, \bar{A})$ of V that minimizes the ratio

$$\kappa(A) = \frac{Capacity(A, \bar{A})}{Demand(A, \bar{A})}.$$

We shall refer to $\kappa = \min_A \kappa(A)$ as *MinCut*(V, C, D). In contrast with the case when there is just one commodity, the MinCut may not be equal to the MaxFlow. The ratio between the MinCut and the MaxFlow, γ, is called the *gap* of the instance (V, C, D). It is easy to see that $\kappa \geq \lambda$, and thus $\gamma \geq 1$. From the computational point of view, computing the value of the MinCut is an NP-hard problem.

Graph Minors: The graph $G = (V, E)$ has an H-*minor* if we can perform a series of edge-deletion and edge-contraction operations on G to obtain a graph G' that is isomorphic to H. Note that each vertex of G' corresponds to a (connected) set of vertices of G which were contracted to it. For $U \subseteq V$, we say that G has an H-*minor w.r.t.* U if it has an H-minor G' such that for every vertex of G', the corresponding set of vertices of G contains a vertex from U. Finally, we say that G is H-*free* (w.r.t. U) if it has no H-minor (w.r.t. U).

It can be verified that K_4-free graphs are those whose biconnected components (or blocks) are series-parallel graphs, and $K_{2,3}$-free graphs are those whose blocks are either outerplanar or isomorphic to K_4.

3 Multicommodity Flows and Metrics

Multicommodity flows have long been an object of study in combinatorial optimization (see [17] for a historical survey). The classical theory was mainly concerned with the following question: Under what conditions on the flow network (V, E, P) is the MaxFlow equal to the MinCut for every setting of capacities C and demands D? As we shall see below, this question is equivalent to the following question concerning the ℓ_1-embeddability of metrics: What are the conditions on (V, E, P) such that, for every metric μ generated by $G = (V, E)$, there is an ℓ_1-embeddable metric ν such that μ dominates ν, and $\mu = \nu$ on P?

In light of this equivalence, the classical results about flows (in cases where the gap is 1) have consequences for ℓ_1-embeddability and *vice versa*. For instance, a well-known theorem due to Okamura and Seymour [25] says that if $G = (V, E)$ is a planar graph with outer face F, and $P \subseteq F \times F$, then the MaxFlow and MinCut are equal for all instantiations of C and D. Taking $G = (V, E)$ to be an outerplanar graph, setting $P = V$ and using the above equivalence, we can infer that all metrics supported on outerplanar graphs can be isometrically embedded into ℓ_1. To state this and other such results succinctly, let us introduce some notation. For a metric μ, let $c_1(\mu)$ be the minimum distortion between μ and ρ, where ρ ranges over all ℓ_1 metrics, and let $c_1(G)$ be the maximum value of $c_1(\mu)$ for all metrics μ supported on G. Hence, we have just seen that $c_1(G) = 1$ for every outerplanar graph G.

In fact, this turns out to be almost a characterization of graphs G with $c_1(G) = 1$. Note that each block of a $K_{2,3}$-free graph is outerplanar or isomorphic to K_4, and a graph is ℓ_1-embeddable iff each of its blocks is. We have already seen that outerplanar graphs are ℓ_1-embeddable, and it can be readily shown that the same holds for all metrics on four points. Thus every $K_{2,3}$-free graph G has $c_1(G) = 1$. Conversely, it is well known that the metric of the unit-weighted $K_{2,3}$ is not ℓ_1-embeddable, and therefore if a graph has a $K_{2,3}$-minor, there is a simple way to weight its edges so that the resulting metric is not ℓ_1-embeddable. Hence we have the following characterization:

Proposition 3.1 *The class of graphs for which $c_1(G) = 1$ is exactly the class of $K_{2,3}$-free graphs.*

Much recent research on multicommodity flows has been directed towards the case where equality does not hold, and to finding good bounds on the ratio γ between MinCut and MaxFlow. This study was pioneered in the paper of Leighton and Rao [22], and the results presented there were extended in a long sequence of papers by several authors (see [28] for a detailed account). The best results known [23, 2] show that for any flow network (V, E, P), the gap between the MaxFlow and the MinCut can never be more than $O(\log |P|)$, and hence $O(\log n)$, a bound which is tight for constant degree expanders. Better results have been obtained for planar graphs, showing that in such graphs the gap γ never exceeds $O(\sqrt{\log n}\,)$ [27] and, moreover, is a constant in the important case of uniform demands [20].

It turns out that there is an intimate relationship between the gap γ and $c_1(G)$ even in the case where the MaxFlow is not equal to the MinCut:

Theorem 3.2 *For any graph $G = (V, E)$, the worst possible gap γ attained by a multicommodity flow problem on G is exactly $c_1(G)$.*

One direction of this theorem can be easily seen. Indeed, it is known (see, e.g., [23] or [15, page 135]) that $\mathrm{MaxFlow}(V, C, D)$ is equal to the minimum of $\frac{C\cdot\delta}{D\cdot\delta}$, where δ ranges over all metrics on G, while $\mathrm{MinCut}(V, C, D)$ is the minimum of the same ratio where δ ranges over all ℓ_1-*embeddable* metrics on G. Now, given a flow of value λ and a metric δ which achieves $\lambda = \frac{C\cdot\delta}{D\cdot\delta}$, we can embed δ into an ℓ_1-metric δ' with distortion at most $c_1(G)$, and hence $\frac{C\cdot\delta'}{D\cdot\delta'} \le \lambda\, c_1(G)$. This metric δ' can then be used to find a cut of value at most $\lambda\, c_1(G)$. The other direction is more involved, and the proof is left for the final version of the paper.

In view of Theorem 3.2, the study of the value of $c_1(G)$ assumes prime importance. Recall that by Proposition 3.1, the graphs for which $c_1(G) = 1$ are exactly the $K_{2,3}$-free graphs. It is not a coincidence that this characterization is in terms of excluded minors: let $\mathcal{F}_c$ denote the family of graphs G for which all metrics generated by G are embeddable into ℓ_1 with distortion at most c. One can show that, for any c, $\mathcal{F}_c$ is minor-closed. Indeed, edge-deletion corresponds to assigning an edge the value ∞, and edge-contraction corresponds to assigning it the value 0. It therefore seems natural to consider families of graphs defined in terms of excluded minors. An intriguing conjecture is that the converse assertion also holds: i.e., for any minor-closed family $\mathcal{F}$ of graphs, there is a constant $c_{\mathcal{F}} \ge 1$ such that for all $G \in \mathcal{F}$, $c_1(G) \le c_{\mathcal{F}}$.

The results in the next section provide further evidence in favor of this conjecture. We consider the next natural minor-closed class of graphs containing $K_{2,3}$, namely series-parallel graphs, and show that they are ℓ_1-embeddable with constant distortion. In addition, we bound the distortion $c_1(G)$ of a graph in terms of its Euler characteristic alone, and thus establish an infinite sequence of natural minor-closed families with constant distortion, namely those with bounded Euler characteristic.

4 Constant-distortion embeddings for some graph families

In this section, we shall present explicit ℓ_1-embeddings which show that the natural minor-closed families of series-parallel graphs and graphs with bounded Euler characteristic can be embedded into ℓ_1 with constant distortion. These are the first non-trivial results exhibiting such non-isometric embeddings with constant distortion.

4.1 Series-parallel graphs

We show that any metric supported on a series-parallel (SP) graph is embeddable in ℓ_1 with constant distortion. In fact, our argument works for the slightly more general class of treewidth 2 graphs, i.e., graphs whose blocks are SP graphs. (Recall that this is a minor closed family with K_4 as the excluded minor.)

Theorem 4.1 *Let $G = (V, E)$ be a weighted graph with treewidth 2, and let $\mu = \mu_G$ be the metric induced by the edge weights of G. Then there exists an ℓ_1-embeddable metric $\widetilde{\mu}$ and a constant $c < 14$ such that for every $u, v \in V$,*

$$\frac{1}{c}\,\mu(u,v) \le \widetilde{\mu}(u,v) \le \mu(u,v).$$

Moreover, this embedding preserves the length of edges, i.e., for every $(u, v) \in E$, $\widetilde{\mu}(u,v) = \mu(u,v)$.

Treewidth 2 graphs can also be defined using the following composition procedure: start with a single edge and repeatedly attach a single new vertex to the endpoints of an already existing edge (which we call its *parent* edge). After all the vertices have been attached, remove an arbitrary subset of the edges. W.l.o.g., we assume that the weight of any edge e is $\mu(e)$, and that no edges are removed at the end.

Proof of Theorem 4.1: Let $G^2, G^3, ..., G^n = G$ be the sequence of intermediate graphs in the above composition of G. The approximating metric $\widetilde{\mu}$ will be constructed inductively. For the base case, G^2 is a single edge $e = (a, b)$, and we set $\widetilde{\mu}(a, b) = \mu(a, b)$.

For the inductive step, assume that G^i is obtained from G^{i-1} by attaching a new vertex x to the endpoints of the edge (s, t). Let $\delta = (\mu(x, s) + \mu(x, t) - \mu(s, t))/2$, and define "weights" $w'(x, s) = \mu(x, s) - \delta$ and $w'(x, t) = \mu(x, t) - \delta$. Note that $w'(x, s) + w'(x, t) = \mu(s, t)$. Now, the value of $\widetilde{\mu}(x, \cdot)$, where $\cdot$ stands for any vertex of $V(G^i) - \{x\}$, is defined as

$$\begin{aligned}\widetilde{\mu}(x, \cdot) = \delta + & \frac{w'(x, t)}{w'(x, s) + w'(x, t)}\,\widetilde{\mu}(s, \cdot) \\ & + \frac{w'(x, s)}{w'(x, s) + w'(x, t)}\,\widetilde{\mu}(t, \cdot).\end{aligned} \tag{4.1}$$

It can be easily verified that $\widetilde{\mu}$ is a metric, and further that it may be represented as a non-negative sum of cut metrics on V and thus is ℓ_1-embeddable.

For our analysis, it will actually be useful to have an equivalent "backwards" description of this construction. If the edge (s, t) is the parent of x, we take out a cut metric corresponding to x (having weight δ), and then we collapse the vertex x to either s or t, with probabilities P_s and P_t inversely proportional to the weights $w'(x, t)$ and $w'(x, s)$ respectively. If we envision this process as starting from the final vertex and collapsing the current "last" vertex onto one of the endpoints of its parent edge, then the distance from vertex x to vertex y is the expected value of the separating cuts taken out before x and y collapse to the same point. This view of the embedding as a random process will enable us to bound the distortion.

We first show that the embedding is a contraction, i.e., for every pair of vertices $u, v \in V$, $\widetilde{\mu}(u, v) \le \mu(u, v)$. Note that it suffices to show the stronger statement that every edge of G maintains its length, i.e., for every $e = (u, v)$, $\widetilde{\mu}(u, v) = \mu(u, v)$, since the expansion of a graph embedding is always largest for the edges of the graph. The equality clearly holds for the base case. For the inductive step, let us attach the vertex x to the edge (s, t). We assume that the equality holds for all edges in G^{i-1}, and in particular for the edge (s, t). We just have to show the equality for the two new edges (x, s) and (x, t). But from (4.1) we have $\widetilde{\mu}(x, s) = \delta + P_t\, \widetilde{\mu}(s, t) = \delta + P_t\, \mu(s, t) = \mu(x, s)$. A similar argument shows this for the edge (x, t) as well.

Before we bound the contraction of the embedding, let us introduce some terminology. An *ancestor* edge of a vertex x is either its parent edge (s, t), or an ancestor edge of s or t. For any two vertices x and y, we can define the *last common ancestor* edge as the common ancestor edge that was added last during the composition process for G. It is easy to see that, if (s, t) is the last common ancestor of x and y, then $\mu(x, y) = \min\{\, \mu(x, s) + \mu(y, s),\ \mu(x, t) + \mu(y, t)\, \}$.

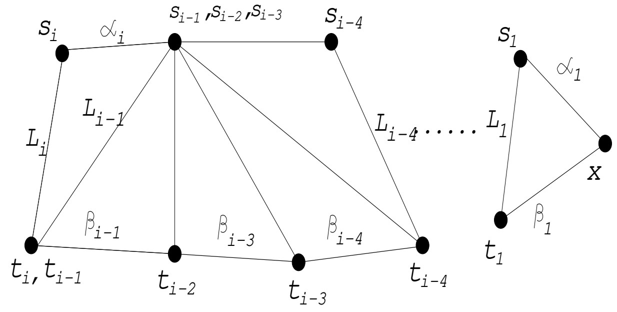

Figure 4.1. Proof of Claim 4.2

Claim 4.2 *If s lies on an ancestor edge of x, then for any $\frac{1}{2} < \xi < 1$ we have*

$$\widetilde{\mu}(x, s) \;\ge\; \frac{(1 - \xi)(2\xi - 1)}{1 + \xi}\,\mu(x, s).$$

Proof of Claim 4.2: Let $(s_1, t_1), (s_2, t_2), \ldots, (s_k, t_k) = (s, t)$ be the sequence of ancestor edges of x up to the edge on which s lies, as in Figure 4.1. Let $s_0 = t_0 = x$, and for all $i \in \{1, \ldots, k\}$ define

$$L_i = \mu(s_i, t_i); \qquad \alpha_i = \mu(s_{i-1}, s_i); \qquad \beta_i = \mu(t_{i-1}, t_i).$$

Note that for each $i \ge 2$, either $t_{i-1} = t_i$ with $\beta_i = 0$, or $s_{i-1} = s_i$ with $\alpha_i = 0$. An important observation is that all the values $\widetilde{\mu}(x, s_i)$ and $\widetilde{\mu}(x, t_i)$ are defined by the graph induced by the vertices $\{s_j, t_j \mid 0 \le j \le i\}$ (again, as in Figure 4.1), and thus it suffices to bound the value of $\widetilde{\mu}$ for this subgraph.

Using the random process definition of $\widetilde{\mu}$, we define P_s^i (resp., P_t^i) to be the probability that at the time x has collapsed to (s_i, t_i), x is at s_i (resp., at t_i). Let Δ_i be the expected value of the sum of the weights of the cuts removed before this happens. Note that all these cuts separate x from s. From the definition of the random process, we have

$$\begin{aligned}\widetilde{\mu}(x, s_{i-1}) &= \Delta_{i-1} \;+\; P_t^{i-1}\, L_{i-1}; \\ \widetilde{\mu}(x, t_{i-1}) &= \Delta_{i-1} \;+\; P_s^{i-1}\, L_{i-1}.\end{aligned} \tag{4.2}$$

Now assume that s_i, t_i are situated as in Figure 4.1; the other case (in which s_i coincides with s_{i-1}) is handled symmetrically. In this case, $\widetilde{\mu}(x, t_i) = \widetilde{\mu}(x, t_{i-1})$. Let us relate the value of $\widetilde{\mu}(x, s_i)$ to the values of $\widetilde{\mu}(x, s_{i-1})$ and $\widetilde{\mu}(x, t_{i-1})$. As defined earlier, the weight of the cut δ_i is $(\mu(s_{i-1}, s_i) + \mu(s_{i-1}, t_i) - L_i)/2$. From the definition of the random process, we get

$$\begin{aligned}\widetilde{\mu}(x,s_i) &= P_t^{i-1}\,(\Delta_{i-1}+L_i)\\ &\quad + P_s^{i-1}\left(\Delta_{i-1}+\delta_i+\frac{\alpha_i-\delta_i}{L_i}\,L_i\right)\\ &= \Delta_{i-1} \;+\; P_t^{i-1}\,L_i \;+\; P_s^{i-1}\,\alpha_i.\end{aligned} \tag{4.3}$$

From (4.2) and (4.3), we get the following bound:

$$\begin{aligned}\widetilde{\mu}(x,s_i) &= \Delta_{i-1}+P_t^{i-1}\,(L_i+\alpha_i)+(P_s^{i-1}-P_t^{i-1})\alpha_i\\ &\geq \Delta_{i-1}+P_t^{i-1}\,L_{i-1}+(P_s^{i-1}-P_t^{i-1})\alpha_i\\ &= \widetilde{\mu}(x,s_{i-1})+(P_s^{i-1}-P_t^{i-1})\alpha_i,\end{aligned} \tag{4.4}$$

where the second line follows from the triangle inequality. By a similar argument, we get that

$$\widetilde{\mu}(x,s_i) \geq \widetilde{\mu}(x,t_{i-1})+(P_t^{i-1}-P_s^{i-1})L_i. \tag{4.5}$$

Finally, we consider the change in Δ:

$$\Delta_i-\Delta_{i-1} \;=\; P_s^{i-1}\,\delta_i \;=\; P_s^{i-1}\cdot\frac{\alpha_i+L_{i-1}-L_i}{2}.$$

Substituting this last expression for the value of $(\Delta_i - \Delta_{i-1})$, and using (4.3) and (4.2), we get:

$$\begin{aligned}&\widetilde{\mu}(x,s_i)+\frac{2P_t^{i-1}}{P_s^{i-1}}(\Delta_i-\Delta_{i-1})\\ &= \Delta_{i-1}+P_t^{i-1}L_i+P_s^{i-1}\alpha_i+P_t^{i-1}(\alpha_i+L_{i-1}-L_i)\\ &= \widetilde{\mu}(x,s_{i-1})+\alpha_i.\end{aligned} \tag{4.6}$$

Let ξ be the constant in the statement of Claim 4.2. The calculations above ((4.4), (4.5) and (4.6) respectively) imply that:

1. If $P_s^{i-1}>\xi$, then $\widetilde{\mu}(x,s_i) \;\geq\; \widetilde{\mu}(x,s_{i-1})+(2\xi-1)\alpha_i$.
2. If $P_t^{i-1}>\xi$, then $\widetilde{\mu}(x,s_i) \;\geq\; \widetilde{\mu}(x,t_{i-1})+(2\xi-1)L_i$.
3. If $1-\xi \;\leq\; P_s^{i-1} \;\leq\; \xi$, then $\widetilde{\mu}(x,s_i)+\frac{2\xi}{1-\xi}(\Delta_i-\Delta_{i-1}) \;\geq\; \widetilde{\mu}(x,s_{i-1})+\alpha_i$.

We are now in position to bound $\widetilde{\mu}(x,s_k)$ from below in terms of $\mu(x,s_k)$. For this purpose, we will construct a path between x and s_k in G, and show that for every edge on this path, we get a substantial contribution to $\widetilde{\mu}(x,s_k)$. Since the length of the path cannot be less than $\mu(x,s_k)$, this will yeild the desired lower bound.

The path P from s_k to x in G will be defined as follows. Let us assume we have already constructed some initial segment of P, and have reached one endpoint of the edge (s_i,t_i), with s_i,t_i again situated as in Figure 4.1 (the other case is symmetrical). If we are at t_i, we do nothing and observe that in fact we have reached (s_{i-1},t_{i-1}). If we are at s_i, we consider the value of P_t^{i-1} defined above. If, on the other hand, $P_t^{i-1} > \xi$, we add to P the edge (s_i,t_{i-1}) of length L_i and proceed to (s_{i-1},t_{i-1}); otherwise, we add to P the edge (s_i,s_{i-1}) of length α_i and proceed to (s_{i-1},t_{i-1}).

Clearly, P is a well defined path from s_k to x in G. Moreover, by our choice of P and the preceding analysis, if P is $\langle s_k=p_0\to p_1\to p_2\to\ldots\to p_m=x\rangle$, then for every edge $(p_{j-1},p_j)\in P$,

$$\begin{aligned}&\widetilde{\mu}(x,p_j)-\widetilde{\mu}(x,p_{j-1})+\frac{2\xi}{1-\xi}\,(\Delta_{p_j}-\Delta_{p_{j-1}})\\ &\qquad\geq (2\xi-1)\,\mu(p_{j-1},p_j),\end{aligned}$$

where (with slight abuse of notation) Δ_{p_j} denotes the Δ_i corresponding to point p_j along the path. Summing up these expressions, we get that

$$\begin{aligned}\widetilde{\mu}(x,s_k)+\frac{2\xi}{1-\xi}\,\Delta_k &\geq (2\xi-1)\cdot(\text{the } \mu\text{-length of } P)\\ &\geq (2\xi-1)\,\mu(x,s_k).\end{aligned} \tag{4.7}$$

Since $\widetilde{\mu}(x,s_k)\geq\Delta_k$, this completes the proof of Claim 4.2. ■

Claim 4.3 *Let x and x^* be two vertices of G such that neither lies on an ancestor edge of the other. Then, for any $\frac{1}{2}<\xi<1$, we have*

$$\widetilde{\mu}(x,x^*) \;\geq\; \frac{(1-\xi)(2\xi-1)}{1+\xi}\,\mu(x,x^*).$$

Proof of Claim 4.3: Let (s,t) be the last common ancestor edge of x and x^*. Let P_s (resp., P_s^*) denote the probability that x (resp., x^*) collapses to s, and P_t and P_t^* be the corresponding values for t. Also let Δ (resp., Δ^*) be the expected value of the sum of the weights of cut metrics removed in the process of collapsing x (resp., x^*) to the edge (s,t). Clearly,

$$\widetilde{\mu}(x,x^*) = \Delta+\Delta^*+(P_sP_t^*+P_tP_s^*)\,\mu(s,t). \tag{4.8}$$

Moreover, it can be easily verified that $P_sP_t^*+P_tP_s^* \geq \frac{1}{2}\min\{P_s+P_s^*;\; P_t+P_t^*\}$. Substituting this into (4.8), assuming (w.l.o.g.) that the minimum is attained at t, and then using (4.2), we get that

$$\widetilde{\mu}(x,x^*) \geq \tfrac{1}{2}\,(\widetilde{\mu}(x,s)+\Delta)+\tfrac{1}{2}\,(\widetilde{\mu}(x^*,s)+\Delta^*)\,. \tag{4.9}$$

Multiplying inequality (4.7) by $\zeta=\frac{1-\xi}{1+\xi}$, and adding it to $(\frac{1}{2}-\zeta)$ times the inequality $\widetilde{\mu}(x,s)\geq\Delta$, we get

$$\frac{1}{2}\,(\widetilde{\mu}(x,s)+\Delta) \;\geq\; \frac{(1-\xi)(2\xi-1)}{1+\xi}\,\mu(x,s)\,.$$

An analogous bound holds for $\widetilde{\mu}(x^*,s)$. These bounds, together with equation (4.9), imply the claim. ■

The claims 4.2 and 4.3 show that, for any two vertices, the metric $\widetilde{\mu}$ closely approximates the original distance μ. Choosing ξ optimally to be $\sqrt{3}-1$, we get that the distortion is at most 13.92, thus proving the theorem. ■

With the aid of a little graph-theoretic machinery, we can generalize Theorem 4.1 as follows. The proof is deferred to the full version of the paper.

Theorem 4.4 *Let $G = (V, E)$ be a graph, and let the set of demand pairs be $P = U \times U$ for some $U \subseteq V$. If G contains no K_4-minor w.r.t. U, then for every assignment of edge capacities C and demands D, the MaxFlow-MinCut gap γ is less than 14.*

4.1.1 Approximating the sparsest cut for series-parallel graphs

The iterative procedure used in the above proof can be exploited to find a near-optimal cut in SP graphs in polynomial time. We briefly sketch the algorithm here. To approximate the MinCut in an SP graph, we first solve the corresponding multicommodity flow problem, and find the metric μ minimizing $\frac{C\cdot\mu}{D\cdot\mu}$ (see the discussion following Theorem 3.2). By Theorem 4.1, we can find an ℓ_1-metric $\widetilde{\mu}$ that 14-approximates μ. Recall the manner in which $\widetilde{\mu}$ is built (see equation (4.1) and the alternative description following it): at each step, it is a weighted sum of three ℓ_1-metrics $\widetilde{\mu}_1, \widetilde{\mu}_2$ and $\widetilde{\mu}_3$. Consequently, at least one of these metrics must yield a value $\frac{C\cdot\widetilde{\mu}_i}{D\cdot\widetilde{\mu}_i}$ which is at most $\frac{C\cdot\widetilde{\mu}}{D\cdot\widetilde{\mu}}$. Choosing the minimizing metric and continuing with the corresponding subgraph, we will eventually reach a point where the remaining metric is a cut metric. This cut achieves the desired approximation ratio. ■

4.2 Embedding graphs with few edges

Recall that, for any graph $G = (V, E)$, the *Euler characteristic* $\chi(G)$ is defined as $|E| - |V| + 1$. It is easy to see that, for each $c \in \mathbb{Z}^+$, the family of graphs $\mathcal{F}_c = \{G \,|\, \chi(G) \le c\}$ is minor-closed. The following theorem shows that graphs with low $\chi(G)$ can be embedded well into ℓ_1.

Theorem 4.5 *Any graph G can be embedded into ℓ_1 with distortion $O(\log\chi(G))$.*

Before proving the theorem, let us introduce the important notion of a *geodesic cut*. Given a weighted graph G, a geodesic cut C of G is an edge cut such that any geodesic path in G intersects C at most once. For each such C, there exists a (maximal) $\alpha_C > 0$ such that

$$d_G = \alpha_C\delta_C + d_{G^-},$$

where δ_C is the cut metric corresponding to C, and G^- is obtained from G by shortening all edges in C by α_C.

Proof of Theorem 4.5 (Sketch): W.l.o.g., we may assume that G is 2-connected and, furthermore, not a cycle, since any cycle metric is isometrically l_1-embeddable. Define a *branch* to be a maximal path in G, each of whose internal vertices has degree two. Thus, each of its endpoints has degree at least three. Removing from each branch B the cut metrics corresponding to the geodesic cuts it contains, we necessarily arrive at a situation where the length of each branch is equal to the distance between its endpoints, $d(B)$. Now for each B, choose uniformly and independently a value $r_B \in [0, d(B)]$, and "collapse" all vertices at current distance at most r from one of its endpoints to this endpoint, and all the other vertices to the other endpoint. Note that the distance between the endpoints remains unchanged. An analysis similar to (but simpler than) that for Theorem 4.1 shows that this procedure introduces a distortion of at most four.

Now we are left with a distribution over graphs, in which each graph G' has all vertices of degree at least three and Euler characteristic $\chi(G') = \chi(G)$. This implies that each G' has at most $O(\chi(G))$ vertices, and it can be embedded into ℓ_1 with a further distortion of $O(\log\chi(G))$ using, e.g., the embedding of Bourgain [10]. Since each G' embeds well, so does any distribution over them, which completes the proof. ■

5 Embedding graphs into tree metrics

The algorithms for ℓ_1-embeddings described in the previous section were based on constructing an approximating set of cut metrics. A complementary approach for embedding a metric (V, μ) into ℓ_1 is to specify a probability distribution over *trees* containing V, such that the *expected* distance between any two vertices x and y in V approximates $\mu(x, y)$ well. Since trees can be embedded isometrically into ℓ_1, this also gives an ℓ_1-embedding, which can be made explicit by sampling trees from the distribution and embedding them into orthogonal subspaces. Of particular interest and importance are embeddings into distributions over *dominating* trees, in which the distance function in each tree dominates μ. Finding low-distortion embeddings of this kind has consequences for the design of many approximation algorithms [4, 3, 18, 29, 12, 14] and online algorithms [4, 6]. Formally:

Definition 5.1 *A graph G is α-*probabilistically approximated *by a distribution $\mathcal{D}$ over trees if*
(1) *each tree T in the distribution $\mathcal{D}$ has $V(G) \subseteq V(T)$;*
(2) *for all $x, y \in V$ and T in the distribution, d_T* dominates *d_G, i.e., $d_G(x, y) \le d_T(x, y)$; and*

(3) *for all $x, y \in V$, the expected distance* $\mathrm{E}_{\mathcal{D}}[d_T(x,y)] \leq \alpha \cdot d_G(x,y)$.

We shall also refer to this as an "embedding of G with distortion α into a tree distribution $\mathcal{D}$".

Distributions over trees were first studied by Karp, who showed that distances in the unweighted cycle C_n can be $2(1-\frac{1}{n})$-probabilistically approximated by a distribution over subtrees [19]. The distribution is very simple: each possible spanning tree of G is output with probability $1/n$. This is in sharp contrast to the deterministic case, where any tree (not necessarily a subtree) approximating the cycle has $\Omega(n)$ distortion [26]. This line of enquiry was further developed in several papers [1, 4, 5, 7, 13], where distributions over arbitrary dominating trees were considered. The state-of-the-art results show that any graph with n vertices can be embedded into tree distributions with distortion $O(\log n \log\log n)$ [5]. In the special case where the graph excludes a $K_{s,s}$-minor, a distortion of $O(s^3 \log n)$ can be achieved [7].

In line with our general approach, we now study the embeddability of minor-closed families into tree distributions. The first result of this section shows that all $K_{2,3}$-free graphs can be embedded into tree distributions with distortion at most eight.

Theorem 5.2 *Given any metric d_G supported on a $K_{2,3}$-free graph G, there is a distribution $\mathcal{D}$ over subtrees of G such that $\mathrm{E}_{\mathcal{D}}[d_T(x,y)] \leq 8 \cdot d_G(x,y)$ for all $x, y \in G$.*

Proof: As usual, it suffices to embed only the biconnected components of the $K_{2,3}$-free graph, each of which must be either K_4 or outerplanar. Since any metric on four points can be embedded into a distribution over dominating trees with distortion less than three, it suffices to bound the distortion for 2-connected outerplanar graphs.

Before we give the embedding, let us recall a few facts about outerplanar graphs. Any 2-connected outerplanar graph G is either a simple cycle or, inductively, it is formed by taking an outerplanar graph H and attaching a path between the endpoints of some edge $e = (u,v)$ on its outer face. We shall refer to this as a *natural composition* of G. Assuming w.l.o.g. that the length of any edge is equal to the distance between its endpoints, we note that at any stage of the composition the metric of a partial graph H coincides with the metric induced on H by G.

The embedding is inductive and follows the natural composition. At each stage, we maintain a random spanning tree of the current outerplanar graph. To embed the initial cycle, we just pick an edge from the cycle with probability proportional to its length, and delete it to get a spanning tree. For the inductive step, let us assume that the graph G is formed by attaching a path P to the endpoints of an edge in H. Assuming inductively that H is embedded into a (random) tree T_H, we pick an edge e' from P with probability proportional to its length, and set $T_G = T_H \cup (P - \{e'\})$. It is easy to see that this is a spanning tree of G.

To bound the expected distortion, note that it suffices to give an upper bound on the expansion of any edge. We will show that the expected expansion is (under a certain condition) at most four. In the base case, an edge is picked at random from a cycle, and the expected distance between the endpoints of any edge is at most twice its length. For the inductive step, assume G was obtained by attaching P to vertices u and v in H. Further, let $d = d_H(u,v) = d_G(u,v)$, and let l be the length of P. Let us inductively assume that $\mathrm{E}\left[d_{T_H}(x,y)\right] \leq \alpha d_H(x,y)$ for all $x, y \in H$. Note that the distance between any two points in T_H does not increase in T_G, and hence we just have to bound the distortion for edges on the path P. For any edge e' on P with length x, the expected distance between its endpoints is at most

$$\left(\frac{x}{l}\right) \cdot (\alpha d + l - x) + \left(\frac{l-x}{l}\right) \cdot x \;\leq\; x\left(2 + \alpha\left(\frac{d}{l}\right)\right).$$

Thus, if we knew in advance that $d/l \leq \frac{1}{2}$, then setting $\alpha = 4$ would ensure that the expected distance is at most αx, completing the proof.

While it is possible that d/l is very close to 1 in general weighted outerplanar graphs, we will show that, given any outerplanar graph G, one can construct (in polynomial time) another outerplanar graph G' on the same vertex set, and a natural composition of G', in which the length of every path P added in this composition is at least twice the distance between the points to which it is added. In fact, G' will be a subgraph of G, and the weight of an edge e in G' will be at least half its weight in G, and at most its weight in G. This, in turn, will imply that $d_{G'} \leq d_G \leq 2d_{G'}$, i.e., G' 2-approximates the distances of G.

Here is a sketch of a recursive procedure to construct G' from G, which also defines a natural composition for G'. If G is a cycle, we set $G' = G$. Otherwise, let us assume that G is obtained by adding a path P to the endpoints of the edge e in H, and H' has been obtained by applying the preprocessing recursively to H. If the length of P is at least twice the length of the edge e in H', then we simply set $G' = H' \cup P$. (Note that the length of an edge $e = (u,v)$ may be greater than the distance between u and v in H' due to previous changes in edge lengths.) Else, we multiply the lengths of the edges of P by a factor $\theta = \mathrm{length}_{H'}(e)/\mathrm{length}_G(P)$, so that the length of P becomes exactly the length of e. Then we remove the edge e from H' and add the path P (with the modified edge-lengths) to get G'. Note that this recursive procedure alters the length of any edge at most once, and while doing so, it shrinks it by a factor of at most two. Furthermore, the length of any edge is at least the G'-distance between its endpoints. Also, the G'-length of any path that is added between the endpoints of e is either exactly equal, or at least twice as long as the length of e. To complete the preprocessing, we remove all edges which are not essential, i.e., whose lengths are strictly greater than the distance between their endpoints. It can be easily checked that the remaining graph still has all the desired properties with respect to the new composition.

Now the final embedding procedure is the following. Given an arbitrary outerplanar graph G, we perform the preprocessing step to construct G'. This incurs a distortion of at most two. We then embed G' into a distribution of subtrees of G' incurring a further expected distortion of at most four, and hence a combined expected distortion of at most eight. It is now easy to check that restoring the original edge lengths (as in G) to the subtrees of G' that appear in this distribution can only improve the distortion, thus proving the theorem. ■

Remark 5.3 Using a similar (but simpler) method combined with the construction of [5], one can show that every graph with Euler characteristic $\chi(G)$ can be embedded into a dominating tree distribution with distortion $O(\log \chi(G) \log\log \chi(G))$, a result analogous to Theorem 4.5 for embeddings into tree distributions.

In view of the results of the previous sections, Theorem 5.2 may inspire hope that embeddings into tree distributions with constant distortion exist for other minor-closed families, such as series-parallel graphs. Our final result shows that this is not so; we prove a lower bound of $\Omega(\log n)$ on the distortion for embedding series-parallel graphs into dominating tree distributions. This extends results of Alon *et al.* [1] and Bartal *et al.* [7], who gave a more involved lower bound for the n-vertex grid, and shows that approximating graph metrics by distributions over tree metrics breaks down even for relatively simple families of graphs.

Theorem 5.4 *There exists an infinite family of series-parallel graphs $\{G_k\}$, such that any α-approximation of the metric of G_k by a distribution over dominating trees has $\alpha = \Omega(\log |G_k|)$.*

The proof uses the following result:

Theorem 5.5 (Rabinovich & Raz [26]) *The distortion of any embedding of the unit-weighted cycle C_n into an (arbitrary) tree is at least $n/3 - 1$.*

Proof of Theorem 5.4: The graphs G_k are defined recursively as follows. Let G_0 be a single unit-weighted edge between s_0 and t_0. Inductively, let H_{i+1} consist of two copies of G_i in series, and G_{i+1} consist of two copies of H_{i+1} in parallel between s_{i+1} and t_{i+1}, which we call its terminals (see Figure 5.2). The graph G_k has $n = 4^k$ edges and $\Theta(n)$ vertices. Observe that for any G_i with terminals s_i and t_i, both the s_i-t_i distance and the size of every s_i-t_i minimal cut are 2^i.

We use Yao's lemma [30] to prove the lower bound. As in the lower bound proofs of [1, 7], it suffices to specify a distribution $\mathcal{D}$ over the edges of $G = G_k$, such that any tree T with $V(G) \subseteq V(T)$ and $d_T \geq d_G$ has a large expected distortion, i.e., there is a constant c such that $\mathrm{E}_{\,e=(x,y)\in\mathcal{D}}[d_T(x,y)] \geq ck$.

Let the distribution $\mathcal{D}$ be uniform over the edges of G: this implies that the average distortion is simply the sum of the tree-distances between the endpoints of the edges divided by 4^k. Since all edges are unit-weighted, the tree-distance is just the expansion of the edge.

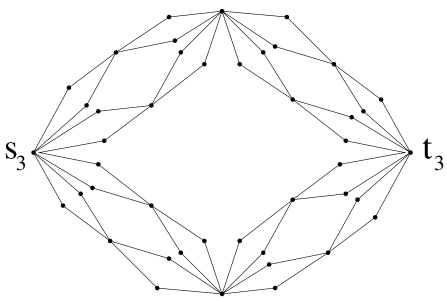

Figure 5.2. The graph G_3

Fix T to be any tree which contains the vertices of $G = G_k$ and dominates distances in it. Let us pick a cycle which goes around the graph G (i.e., a simple cycle which includes the terminals s_k and t_k); its length in G is 2^{k+1}. By Theorem 5.5, at least one of the edges on this cycle has expansion $\geq 2^{k+1}/3 - 1$. Let us color all such edges on this cycle with the color k. We now pick another cycle of the same length from among the uncolored edges and perform a similar coloring operation. We continue this process until each cycle going around G_k contains a colored edge. This implies that at least one of the four copies of G_{k-1} comprising G_k has its terminals separated by edges colored k. We choose one such minimal cut of size 2^{k-1} and uncolor all the other edges which have color k. Now each of the edges still colored k suffers an expansion of at least $2^{k+1}/3 - 1$.

We now look at all copies of G_{k-1} whose terminals have not been separated by colored edges (there are three such copies), and repeating the process we get $3 \times 2^{k-2}$ edges having expansion at least $2^k/3 - 1$. Continuing this way, at the i-th recursive step there are $4^i - (4^{i-1} + 4^{i-2} + \cdots + 1) \geq 4^{i-1}$ copies of G_{k-i} whose terminals are not separated by previously colored edges. (Note that previously colored copies of, say, G_{k-i+2} may contain uncolored copies of

G_{k-i}.) From each such copy we get a contribution of $2^{k-i-1} \times (2^{k-i+1}/3 - 1) \geq 4^{k-i}/3 - 2^{k-i-1}$ towards the aggregated expansion. Summing over all i, the sum of the edge expansions in $G = G_k$ is at least $\Omega(k4^{k-1})$, which in turn implies that the expected expansion is at least $\Omega(k) = \Omega(\log |G_k|)$. ■

Acknowledgments

We would like to thank Gil Kalai, Alexander Karzanov and Mike Saks for insightful discussions.

References

[1] N. Alon, R. M. Karp, D. Peleg, and D. B. West. A graph-theoretic game and its applications to the k-server problem. *SIAM Journal on Computing*, 24(1):78–100, 1995.

[2] Y. Aumann and Y. Rabani. An $O(\log k)$ approximate min-cut max-flow theorem and approximation algorithm. *SIAM Journal on Computing*, 27(1):291–301, Feb. 1998.

[3] B. Awerbuch and Y. Azar. Buy-at-bulk network design. In *38th Annual Symposium on Foundations of Computer Science*, pages 542–547, 1997.

[4] Y. Bartal. Probabilistic approximations of metric spaces and its algorithmic applications. In *37th Symposium on Foundations of Computer Science*, pages 184–193, 1996.

[5] Y. Bartal. On approximating arbitrary metrics by tree metrics. In *30th Annual ACM Symposium on Theory of Computing*, pages 161–168, 1998.

[6] Y. Bartal, A. Blum, C. Burch, and A. Tomkins. A polylog(n)-competitive algorithm for metrical task systems. In *30th ACM Symposium on Theory of Computing*, pages 711–719, 1997.

[7] Y. Bartal, G. Konjevod, R. Ravi, and F. S. Salman. On approximating planar metrics by tree metrics. Manuscript.

[8] S. N. Bhatt and F. T. Leighton. A framework for solving VLSI graph layout problems. *J. of Comp. and System Sciences*, 28(2):300–343, Apr. 1984.

[9] A. Blum, G. Konjevod, R. Ravi, and S. Vempala. Semi-definite relaxations for minimum bandwidth and other vertex-ordering problems. In *30th Annual ACM Symposium on Theory of Computing*, pages 100–105, 1998.

[10] J. Bourgain. On Lipschitz embeddings of finite metric spaces in Hilbert space. *Israel J. of Math.*, 52:46–52, 1985.

[11] J. Bourgain. The metrical interpretation of superreflexivity in Banach spaces. *Israel J. of Math.*, 56(2):222–230, 1986.

[12] M. Charikar, C. Chekuri, A. Goel, and S. Guha. Rounding via trees: deterministic approximation algorithms for group Steiner trees and k median. In *30th Annual ACM Symposium on Theory of Computing*, pages 114–123, 1998.

[13] M. Charikar, C. Chekuri, A. Goel, S. Guha, and S. A. Plotkin. Approximating a finite metric by a small number of tree metrics. In *39th Symposium on the Foundations of Computer Science*, pages 379–388, 1998.

[14] M. Charikar and B. Raghavachari. The finite capacity dial-a-ride problem. In *39th Symposium on the Foundations of Computer Science*, pages 458–467, 1998.

[15] M. M. Deza and M. Laurent. *Geometry of Cuts and Metrics*. Springer Verlag, 1997.

[16] U. Feige. Approximating the bandwidth via volume respecting embeddings. In *30th Annual ACM Symposium on Theory of Computing*, pages 90–99, 1998.

[17] A. Frank. Packing paths, circuits, and cuts – a survey. In B. Korte, L. Lovász, H. J. Prömel, and A. Schrijver, editors, *Paths, Flows and VLSI-Layout*, pages 47–100. Springer-Verlag, 1990.

[18] N. Garg, G. Konjevod, and R. Ravi. A polylogarithmic approximation algorithm for the group Steiner tree problem. In *Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 253–259, 1998.

[19] R. M. Karp. A $2k$-competitive algorithm for the circle. Manuscript, 1989.

[20] P. Klein, S. A. Plotkin, and S. B. Rao. Excluded minors, network decomposition, and multicommodity flow. In *25th Annual ACM Symposium on the Theory of Computing*, pages 682–690, 1993.

[21] P. Klein, S. Rao, A. Agrawal, and R. Ravi. An approximate max-flow min-cut relation for undirected multicommodity flow, with applications. *Combinatorica*, 15(2):187–202, 1995.

[22] F. T. Leighton and S. B. Rao. An approximate max-flow min-cut theorem for uniform multicommodity flow problems with applications to approximation algorithms. In *29th Annual Symposium on Foundations of Computer Science*, pages 422–431, 1988.

[23] N. Linial, E. London, and Y. Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995. (Preliminary version in: *35th Annual Symposium on Foundations of Computer Science*, pages 577–591, 1994.).

[24] J. Matoušek. On embedding trees into uniformly convex Banach spaces. *Israel J. of Math.*, To appear, 1999. (Czech version in : *Lipschitz distance of metric spaces*, C.Sc. degree thesis, Charles University, 1990).

[25] H. Okamura and P. D. Seymour. Multicommodity flows in planar graphs. *Journal of Comb. Theory Series B*, 31:75–81, 1981.

[26] Y. Rabinovich and R. Raz. Lower bounds on the distortion of embedding finite metric spaces in graphs. *Discrete & Computational Geometry*, 19(1):79–94, 1998.

[27] S. B. Rao. Small distortion and volume preserving embeddings for planar and Euclidean metrics. In *15th Annual ACM Symposium on Computational Geometry*, pages 300–306, 1999.

[28] D. B. Shmoys. Cut problems and their application to divide-and-conquer. In D. S. Hochbaum, editor, *Approximation Algorithms for NP-hard Problems*, pages 192–235. PWS Publishing, 1997.

[29] B. Y. Wu, G. Lancia, V. Bafna, K.-M. Chao, R. Ravi, and C. Y. Tan. A polynomial time approximation scheme for minimum routing cost spanning trees. In *Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 21–32, 1998.

[30] A. C. Yao. Probabilistic computations: Toward a unified measure of complexity. In *18th Annual Symposium on Foundations of Computer Science*, pages 222–227, 1977.