Approximation Algorithms for Low-Distortion Embeddings Into Low-Dimensional Spaces

Mihai Bădoiu* Kedar Dhamdhere†‡ Anupam Gupta† Yuri Rabinovich§
Harald Räcke†‡ R. Ravi¶‖ Anastasios Sidiropoulos**

Abstract

We present several approximation algorithms for the problem of embedding metric spaces into a line, and into the two-dimensional plane. Among other results, we give an $O(\sqrt{n})$-approximation algorithm for the problem of finding a line embedding of a metric induced by a given unweighted graph, that minimizes the (standard) multiplicative distortion. We give an improved $\tilde{O}(n^{1/3})$ approximation for the case of metrics generated by unweighted trees. This is the first result of this type.

1 Introduction

Embedding distance matrices into geometric spaces (most notably, into low-dimensional spaces) is a fundamental problem occurring in many applications. In the context of data visualization, this approach allows the user to observe the structure of the data set and discover its interesting properties. In computational chemistry, this approach is used to recreate the geometric structure of the data from the distance information. The problem is of interest in many other areas, see [Wor] for a discussion.

The methods for computing such embeddings have their roots in work going back to the first half of the 20th century, and in the more recent work of Shepard [She62a, She62b], Kruskal [Kru64a, Kru64b], and others. The area is usually called *Multi-dimensional Scaling* (MDS) and is a subject of extensive research [Wor]. However, despite significant practical interest, few theoretical results exist in this area (see Related Work). The most commonly used algorithms are heuristic (e.g., gradient-based method, simulated annealing, etc) and are often not satisfactory in terms of the running time and/or quality of the embeddings.

In this paper we present algorithms for the following fundamental embedding problem: given a graph $G = (V, E)$ inducing a shortest path metric $M = M(G) = (V, D)$, find a mapping f of V into a *line* that is non-contracting (i.e., $|f(u) - f(v)| \geq D(u, v)$ for all $u, v \in V$) which minimizes the distortion $c_{line}(M, f) = \max_{u,v \in V} \frac{|f(u)-f(v)|}{D(u,v)}$. That is, our goal is to find $c_{line}(M) = \min_f c_{line}(M, f)$. For the case when G is an *unweighted* graph, we show the following algorithms for this problem (denote $n = |V|$):

- A polynomial (in fact, $O(n^3 c)$-time) c-approximation algorithm for metrics M for which $c_{line}(M) \leq c$. This also implies an $O(\sqrt{n})$-approximation algorithm for any M (Section 2).
- A polynomial-time $\tilde{O}(\sqrt{c})$ approximation algorithm for metrics generated by unweighted trees. This also implies an $\tilde{O}(n^{1/3})$-approximation algorithm for these metrics (Section 3).
- An exact algorithm, with running time $n^{O(c_{line}(M))}$ (Section 4).

For the case when G is a weighted graph, we obtain the following result. For induced metrics M such that $c_{line}(M) = 1 + \epsilon < 1.5$, we give an algorithm that finds a line embedding f such that $c_{line}(M, f) = 1 + O(\epsilon)$. In other words, the algorithm constructs a good embedding for metrics that are very well embeddable into a line. The algorithm proceeds by computing an MST T of M, and then ordering the nodes according to T. Thus, its running time is $O(n^2)$ in the worst case, and is even more efficient for metric spaces that support faster MST computation. We also note that ordering the metric nodes using MST is a popular heuristic (e.g., see [BJDG+03]). To our knowledge, our result provide

*MIT Computer Science and Artificial Intelligence Laboratory; Cambridge, MA 02139; {mihai,tasos}@theory.lcs.mit.edu

†School of Computer Science; Carnegie Mellon University; Pittsburgh, PA 15213; {kedar,anupamg,harry}@cs.cmu.edu

‡Supported by NSF ITR grants CCR-0085982 and CCR-0122581.

§Department of Computer Science; University of Haifa; 31905 Haifa, Israel; yuri@cslx.haifa.ac.il

¶Tepper School of Business; Carnegie Mellon University; Pittsburgh, PA 15213; ravi@andrew.cmu.edu

‖Supported by NSF ITR grants CCR-0085982, CCR-0122581 and NSF CCF 04-30751.

**Supported by the Paris Kanellakis Fellowship Fund, and by the Alexandros S. Onassis Public Benefit Foundation.

the first known provable guarantee for this heuristic. The details have been omitted from this extended abstract.

We complement our algorithmic results by showing that a-approximating the value of $c_{line}(M)$ is NP-hard for certain $a > 1$ in Section 5. In particular, this justifies the exponential dependence on $c_{line}(M)$ in the running time bound for the exact algorithm.

We also study the problem of embedding metrics into the *plane* in 6. In particular, we focus on embedding metrics $M = (X, D)$ which are induced by a set of points in a unit sphere S^2. Embedding such metrics is important, e.g., for the purpose of visualizing point-sets representing places on Earth or other planets, on a (planar) computer screen.[1] In general, we show that an n-point spherical metric can be embedded with distortion $O(\sqrt{n})$, and this bound is optimal in the worst case. (The lower bound is shown by resorting to the Borsuk-Ulam theorem [Bor33], which roughly states that any continuous mapping from S^2 into the plane maps two antipodes of S^2 into the same point.) For the algorithmic problem of embedding M into the plane, we give a 3.512-approximation algorithm, when D is the Euclidean distance in $\mathbb{R}^3$. For the case where D corresponds to the geodesic distance in S^2, our algorithm can be re-analyzed to give a an approximation guarantee of 3.

To our knowledge, our results provide the first non-trivial approximation guarantees for the standard (multiplicative) notion of distortion for embeddings into low-dimensional spaces.

1.1 Related work

Combinatorial vs Algorithmic Problem. The problem of finding low-distortion embeddings of metrics into geometric spaces has been long a subject of extensive mathematical studies. During the last few years, such embeddings found multiple and diverse uses in computer science as well; many such applications have been surveyed in [Ind01]. However, the problems addressed in this paper are fundamentally different from those investigated in the aforementioned literature. In a nutshell, our problems are *algorithmic*, as opposed to *combinatorial*. More specifically, we are interested in finding the best distortion embedding of a *given* metric (which is an algorithmic problem) as opposed to the best distortion embedding for a *class* of metrics (which is a combinatorial problem). Thus, we define the quality of an embedding algorithm as the worst-case *ratio* of the distortion obtained by the algorithm to the best achievable distortion. In contrast, the combinatorial approach focuses on providing the worst-case upper bounds for the distortion itself. Thus, the problems are fundamentally different, which raises new interesting issues.

Despite the differences, we mention two combinatorial results that are relevant in our context. The first one is the [LLR94] adaptation of Bourgain's construction [Bou85] that enables embedding of an arbitrary metric into $l_2^{O(\log^2 n)}$ with maximum multiplicative distortion $O(\log n)$. It should be noted, however, that for the applications mentioned earlier, the most interesting spaces happen to be low-dimensional. Similarly, any metric can be embedded into d-dimensional Euclidean space with multiplicative distortion $O(\min[n^{\frac{2}{d}} \log^{3/2} n, n])$ and no better than $\Omega(n^{1/\lfloor (d+1)/2 \rfloor})$ [Mat96]. However, the worst-case guarantees are rather large for small d, especially for the case $d = 1$ that we consider here.

Previous Work on the Algorithmic Problem. To our knowledge there have been few *algorithmic* embedding results. Hastad et al. gave a 2-approximation algorithm for embedding an arbitrary metric into a line $\Re$, when the *maximum additive two-sided error* was considered; that is, the goal was to optimize the quantity $\max_{u,v} ||f(u) - f(v)| - D(u, v)|$. They also showed that the same problem cannot be approximated within 4/3 unless $P = NP$ [HIL98, Iva00]. Bădoiu extended the algorithm to the 2-dimensional plane with maximum two-sided additive error when the distances in the target plane are computed using the l_1 norm [Bŏ03]. Bădoiu, Indyk and Rabinovich [BIR03] gave a weakly-quasi-polynomial time algorithm for the same problem in the l_2 norm.

Very recently, Kenyon, Rabani and Sinclair [KRS04] gave *exact* algorithms for minimum (multiplicative) distortion embeddings of metrics *onto* simpler metrics (e.g., line metrics). Their algorithms work as long as the minimum distortion is small, e.g., constant. We note that constraining the embeddings to be *onto* (not *into*, as in our case) is crucial for the correctness of their algorithms.

In general, one can choose non-geometric metric spaces to serve as the host space. For example, in computational biology, approximating a matrix of distances between different genetic sequences by an ultrametric or a tree metric allows one to retrace the evolution path that led to formation of the genetic sequences. Motivated by these applications M. Farach-Colton and S. Kannan show how to find an *ultrametric* T with minimum possible maximum additive distortion [FCKW93]. There is also a 3-approximation algorithm for the case of embedding arbitrary metrics into weighted tree metrics to minimize the maximum additive two-sided er-

[1] Indeed, the whole field of cartography is devoted to low-distortion representations of spherical maps in the plane.

ror [ABFC+96]. [Dha04] recently gave an $O(\log^{1/p} n)$-approximation for embedding arbitrary n-point metrics into the line to minimize the ℓ_p norm of the two-sided error vector $||f(u) - f(v)| - D(u,v)|$.

Distortion vs Bandwidth. In the context of unweighted graphs, the notion of minimum distortion of an embedding into a line is closely related to the notion of a graph *bandwidth*. Specifically, if the non-contraction constraint $|f(u) - f(v)| \geq D(u,v)$ is replaced by a constraint $|f(u) - f(v)| \geq 1$ for $u \neq v$, then $c_1(M(G))$ becomes precisely the same as the bandwidth of the graph G.

There are several algorithms that approximate the bandwidth of a graph [Fei00, Gup00]. Unfortunately, they do not seem applicable in our setting, since they do not enforce the non-contraction constraint for all node pairs. However, in the case of *exact* algorithms the situation is quite different. In particular, our exact algorithm for computing the distortion is based on the analogous algorithm for the bandwidth problem by Saxe [Sax80].

2 A c-approximation algorithm

We start by stating an algorithmic version of a fact proved in [Mat90].

LEMMA 2.1. *Any shortest path metric over an unweighted graph $G = (V, E)$ can be embedded into a line with distortion at most $2n - 1$ in time $O(|V| + |E|)$.*

Proof. Let T be a spanning tree of the graph. We replace every (undirected) edge of T with a pair of opposite directed edges. Since the resulting graph is Eulerian, we can consider an Euler tour C in T. Starting from an arbitrary node, we embed the nodes in T according to the order that they appear in C, ignoring multiple appearances of a node, and preserving the distances in C. Clearly, the resulting embedding is non-contracting, and since C has length $2n$, the distortion is at most $2n - 1$. □

Note that the $O(n)$ bound is tight, e.g. when G is a star.

Let $G = (V, E)$ be a graph, such that there exists an embedding of G of distortion c. The algorithm for computing an embedding of distortion at most $O(c^2)$ is the following:

1. Let f_{OPT} be an optimal embedding of G (note that we just assume the existence of such an embedding, without computing it). Guess nodes $t_1, t_2 \in V$, such that $f_{OPT}(t_1) = \min_{v \in V} f_{OPT}(v)$, and $f_{OPT}(t_2) = \max_{v \in V} f_{OPT}(v)$.
2. Compute the shortest path $p = v_1, v_2, \ldots, v_L$ from t_1 to t_2.
3. Partition V into disjoint sets $V_1, V_2, \ldots V_L$, such that for each $u \in V_i$, $D(u, v_i) = \min_{1 \leq j \leq L} D(u, v_j)$. Break ties so that each V_i is connected.
4. For $i = 1 \ldots L$, compute a spanning tree T_i of the subgraph induced by V_i, rooted at v_i. Embed the nodes of V_i as in the proof of Lemma 2.1, leaving a space of length $|V_i|$ between the nodes of V_i and V_{i+1}.

LEMMA 2.2. *For every $i, 1 \leq i \leq L$, and for every $x \in V_i$, we have $D(v_i, x) \leq c/2$.*

Proof. Assume that the assertion is not true. That is, there exists v_i, and $x \in V_i$, such that $D(x, v_i) > c/2$. Consider the optimal embedding f_{OPT}. By the fact that v_1 and v_L are the left-most and right-most embedded nodes in the embedding f_{OPT}, it follows that there exists $j, 1 \leq j < L$, such that $f_{OPT}(x)$ lies between $f_{OPT}(v_j)$, and $f_{OPT}(v_{j+1})$. W.l.o.g., assume that $f_{OPT}(v_j) < f_{OPT}(x) < f_{OPT}(v_{j+1})$. Since $x \in V_i$, we have $|f_{OPT}(v_{j+1}) - f_{OPT}(v_j)| = f_{OPT}(v_{j+1}) - f_{OPT}(x) + f_{OPT}(x) - f_{OPT}(v_j) \geq D(v_{j+1}, x) + D(x, v_j) \geq 2D(x, v_i) > c$. This is a contradiction, since the expansion of f_{OPT} is at most c. □

LEMMA 2.3. *For every i, $1 \leq i \leq L - c + 1$, we have $\sum_{j=i}^{i+c-1} |V_j| \leq 2c^2$.*

Proof. Assume that there exists i such that $\sum_{j=i}^{i+c-1} |V_j| > 2c^2$. Note that

$$\max_{i \leq j_1 < j_2 \leq i+c-1} |f_{OPT}(v_{j_1}) - f_{OPT}(v_{j_2})| \leq c(c-1).$$

Moreover, since $\sum_{j=i}^{i+c-1} |V_j| > 2c^2$, we have $\max_{u,w \in \bigcup_{j=i}^{i+c-1} V_j} |f_{OPT}(u) - f_{OPT}(w)| \geq 2c^2$. It follows that there exists $u \in V_l$, for some l, with $i \leq l \leq i+c-1$, such that $|f_{OPT}(v_l) - f_{OPT}(u)| \geq \frac{2c^2 - c(c-1)}{2} > c^2/2$. Since the expansion is at most c, we have $D(v_l, u) > c/2$, contradicting Lemma 2.2. □

LEMMA 2.4. *The embedding computed by the algorithm is non-contracting.*

Proof. Let $x, y \in V$. If x and y are in the same set V_i, for some i, then clearly $|f(x) - f(y)| \geq D(x, y)$, since the distance between x and y produced by an traversal of the spanning tree of the graph induced by V_i is at least the distance of x and y on T_i, which is at least $D(x, y)$.

Assume now that $x \in V_i$ and $y \in V_j$, for some $i < j$. We have $|f(y) - f(x)| \geq |V_i| + 2\sum_{l=i+1}^{j-1} |V_l| + |V_j| \geq |V_i| + |V_j| + j - i > D(x, v_i) + D(y, v_j) + D(v_i, v_j) \geq D(x, y)$. □

LEMMA 2.5. *The distortion of the embedding computed by the algorithm is at most $4c^2$.*

Proof. It suffices to show that for each $\{x, y\} \in E$, $|f(x) - f(y)| \leq 4c^2$. Let $x \in V_i$, and $y \in V_j$. If $|i - j| \leq 2c$, then by Lemma 2.3 we obtain that $|f(x) - f(y)| \leq 4c^2$.

Assume now that there exist nodes $x \in V_i$ and $y \in V_j$, with $\{x, y\} \in E$, and $|i - j| > 2c$. By Lemma 2.2, we obtain that $D(v_i, x) \leq c/2$, and $D(y, v_j) \leq c/2$, and thus $|i - j| = D(v_i, v_j) \leq c + 1$, a contradiction. □

THEOREM 2.1. *The described algorithm computes a non-contracting embedding of maximum distortion $O(c^2)$, in time $O(n^3 c)$.*

Proof. By Lemmata 2.4 and 2.5, it follows that the computed embedding is non-contracting and has distortion at most $O(c^2)$. In the beginning of the algorithm, we compute all-pairs shortest paths for the graph. Next, for each possible pair of nodes t_1 and t_2, the described embedding can be computed in linear time. Thus, the total running time is $O(n^2|E|) = O(n^3 c)$. □

THEOREM 2.2. *There exists a $O(\sqrt{n})$-approximation algorithm for the minimum distortion embedding problem.*

Proof. If the optimal distortion c is at most $\sqrt{n}$, then the described algorithm computes an embedding of distortion at most $O(c\sqrt{n})$. Otherwise, the algorithm described in Lemma 2.1, computes an embedding of distortion $O(n)$. Thus, by taking the best of the above two embeddings, we obtain an $O(\sqrt{n})$-approximation. □

3 Better embeddings for unweighted trees

For the case of trees, we use a similar framework as for general graphs: we divide the tree along the path from t_1 to t_2 and obtain connected components $V_1, \ldots, V_L$ each with $\text{diam}(V_i) \leq c$ and $\sum_{j=i}^{i+c-1} |V_j| \leq 2c^2$. Instead of a spanning tree on each V_i, we give a more sophisticated embedding. We consider all the vertices in $X_i = \cup_{j=i}^{i+c} V_j$ together. Lemma 2.2 gives the following bound on the diameter of the set X_i.

LEMMA 3.1. *The diameter of the set X_j (for $j = 1, 2, \ldots$) is at most $2c$.*

We use the following straightforward lower bound on the distortion for embedding X_j.

The *local density* Δ of G is defined as

$$\Delta = \max_{v \in V, r \in \mathbb{R}_{>0}} \left\{ \frac{|B(v, r)| - 1}{2r} \right\} ,$$

where $|B(v, r)| = \{u \in V \mid d(u, v) \leq r\}$ denotes the ball of nodes within distance r from v. Intuitively, a high local density tells us that there are dense clusters in the graph, which will cause a large distortion. The following lemma formalizes this intuition.

LEMMA 3.2. **[Local Density]** *Let G denote a graph with local density Δ. Then any map of G into the line has distortion at least Δ.*

3.1 Prefix Embeddings.

We first prove that it suffices to consider embeddings where each prefix of the associated tour forms a connected component of the tree; this will allow us to considerably simplify all our later arguments.

LEMMA 3.3. **[Prefix Embeddings]** *Given any graph G, there exists an embedding of G into the real line with the following two properties:*

1. *Walk from left to right on the line, the set of points encountered up to a certain point forms a connected component of G.*
2. *The distortion of this map is at most twice the optimal distortion.*

Proof. Consider the optimal embedding f^*, and let $v_1, v_2, \ldots, v_n$ be the order of the points in this embedding. (We will blur the distinction between a vertex v and its image $f^*(v)$ on the line.) Without loss of generality, we can assume that the distance between any two adjacent points v_i and v_{i+1} in this embedding is their shortest path distance $D(v_i, v_{i+1})$.

Let i be the smallest index such that $\{v_1, v_2, \ldots, v_i\}$ does not form a connected subgraph; hence there exists some vertex on every v_{i-1}-v_i path that has not yet been laid out. We pick a shortest path P, take the vertex w in $P \setminus \{v_1, v_2, \ldots, v_{i-1}\}$ closest to v_{i-1}, and place it at distance $D(v_{i-1}, w)$ to the right of v_{i-1} in the embedding. We repeat this process until Property 1 is satisfied; it remains to bound the distortion we have introduced.

Note that the above process moves each vertex at most once, and only moves vertices to the left. We claim that each vertex is moved by at most distance c, where c is the optimal distortion. Indeed, consider a vertex w that is moved when addressing the v_{i-1}-v_i path, and let v_k be a neighbor of w among $v_1, \ldots, v_{i-1}$. Note that the distance $|f^*(v_k) - f^*(w)|$ between these two vertices is at most c in the optimal embedding. Since w stays to the right of v_k, the distance by which w is moved is at most c.

In short, though the above alterations move vertices to the left, whilst keeping others at their original

locations in f^*, the distance between the endpoints of an edge increases by at most c. Since the distance $|f^*(v) - f^*(u)|$ was at most c to begin with, we end up with an embedding with (multiplicative) distortion at most $2c$, proving the lemma. □

Henceforth, we will only consider embeddings that satisfy the properties stated in Lemma 3.3. The bound on the increase in distortion is asymptotically best possible: for the case of the n-vertex star $K_{1,n-1}$, the optimal distortion is $\approx n/2$, but any prefix embedding has distortion at least $n-2$.

3.2 The Embedding Algorithm. In this section, we give an algorithm which embeds trees with distortion $g(c) = 2\Delta\sqrt{c \log c} + c$, where Δ is the local density and c the optimal distortion. The algorithm proceeds in rounds: in round i, we lay down a set Z_i with about $g(c)$ vertices. To ensure that the neighbors of vertices are not placed too far away from them, we enforce the condition that the vertices in Z_i include all the neighbors of vertices in $\cup_{j<i} Z_j$ that have not already been laid out.

It is this very tension between needing to lay out a lot of vertices and needing to ensure their neighbors can be laid out later on, that leads to the following algorithm. In fact, we will mentally separate the action of laying out the neighbors of previously embedded vertices (which we call the *BFS part* of the round) from that of laying out new vertices (which we call the *DFS part*).

We assume that we know the left-most vertex r in the prefix embedding; we can just run over all the possible values of r to handle this assumption. Let $N(X)$ denote the set of neighbors of vertices in a set $X \subseteq V$.

We define a *light path ordering* on the vertices of the tree T. The light path ordering is a DFS ordering which starts at root r and at each point enters the subtree with smallest number of vertices in it.

Algorithm Tree-Embed:

1. let $C \leftarrow \{r\}$ denote the set of vertices already visited. Set $i \leftarrow 1$.
2. while $C \neq V(T)$ do
   (Round i BFS)
3. Visit all vertices in $N(C) \setminus C$; let $C \leftarrow C \cup N(C)$
   (Round i DFS)
5. set B to be a set of $g(c)$ vertices of $V(T) \setminus C$ in the *light path ordering*. Visit all vertices in B; let $C \leftarrow C \cup B$.
6. endwhile

LEMMA 3.4. [**Number of rounds**] *The algorithm Tree-Embed requires at most* $\sqrt{c \log^{-1} c}$ *iterations.*

Proof. By the very definition of the algorithm, the set C grows by at least $g(c)$ in every iteration. Note that the diameter of the tree is bounded by $2c$ and its local density is Δ. Therefore, the number of nodes in the tree is at most $2\Delta c$. Hence, within $(2\Delta c)/g(c) \leq \sqrt{c \log^{-1} c}$ iterations, all the vertices of the tree will be visited. □

The heart of the proof is to show that visiting the vertices in Steps 3 and 5 does not incur too much distortion; it may be the case that the size of $N(C) \setminus C$ may be too large, or even that these vertices may be separated very far from each other.

LEMMA 3.5. [**Span of boundary**] *The size of the induced spanning tree on the boundary* $N(C) \setminus C$ *is bounded by* $g(c)$.

Proof. Consider the set C_i of vertices that have been visited by round i. Consider a vertex x visited in round j of the DFS for some $j \leq i$. Note that the children of the vertex x will be visited *after* x. We say that x is a *branching point* if not all the children of x were visited in the same round as x. The branching point x is *active* after round i if at least one of the vertices below it has not been visited by round i; otherwise it is *inactive*. We claim that all the active branching points in C_i lie on some root-leaf path. This follows because the light path ordering is a DFS ordering. Therefore, if some vertices below a branching point x have not been visited, then the DFS part of the algorithm will not visit a different subtree.

Note that each active branching point (except possibly the lowest one) has at least two children and the algorithm visits the child which has a smaller number of vertices in its subtree. Recall that the size of the tree is bounded by $2c^2$ by Lemma 2.3. Therefore, the number of active branching points on a root to leaf path is at most $2 \log c + 1$.

We claim that every node in $N(C_i) \setminus C_i$ is within a distance of $i+1$ of some active branching point. We prove this by induction on i. Before the first round, this property is true, since $C_0 = \{r\}$. Now assume the property for $i-1$ and consider a vertex $v \in N(C_i) \setminus C_i$. Let u be the neighbor of v such that $u \in C_i$. If u was visited in the round i of the DFS, then u is an active branching point, since its child v has not been visited in the same round. Otherwise, if u was visited in round i of the BFS, then u is within distance i of some branching point x. Since v is below x and has not been visited after round i, the branching point x must be active. Therefore, v is within distance $i+1$ from some active branching point.

Consider an active branching point x and let N_x contain the points from $N(C_i) \setminus C_i$ that are within distance $i+1$ from x. Then, we can bound the span of the induced tree on N_x using the local density bound. The number of vertices in the induced tree on N_x is bounded by $(i+1)\Delta$. Thus, for each active branching point, the number of vertices in the induced tree is bounded by $\Delta\sqrt{c\log^{-1} c}$. Since there are $2\log c + 1$ branching points overall, the sum of spans over all the active branching points is at most $2\Delta\sqrt{c\log c}$. Note that, all the active branching points are on a single root-leaf path. Therefore, connecting all the branching points in $N(C_i) \setminus C_i$ requires only a path of length c. Hence, the total span of vertices in $N(C_i)\setminus C_i$ is bounded by $g(c)$. □

LEMMA 3.6. *The span of the tree induced on the vertices visited in any iteration is bounded by* $2g(c)$.

Proof. From Lemma 3.5, the span of the vertices visited in Step 3 of the algorithm is bounded by $g(c)$. The number of new vertices visited in Step 5 of the algorithm is bounded by $g(c)$. Since, we visit a set of connected components, their span is bounded by $g(c)+\text{span}(N(C)\setminus C)$. Therefore, the span of the vertices visited in each iteration is bounded by $2g(c)$. □

LEMMA 3.7. *The distortion of the embedding produced by Algorithm Tree-Embed is* $4g(c)$.

Proof. For a pair of vertices that are visited during the same iteration, the distance in the embedding is bounded by $2g(c)$ (from Lemma 3.6). Therefore, the distortion of such a pair is bounded by $4g(c)$. So, consider an edge (x, y) such that x and y were visited in different iterations. Note that, Step 1 of the algorithm ensures that if x is visited in iteration i, then y is visited in iteration $i+1$. Therefore, the distance between x and y in the embedding is bounded by $4g(c)$. Hence, the distortion is bounded by $4g(c)$. □

Concatenating the embeddings. In order to concatenate the embeddings of $X_1, X_2, \ldots$, it is enough to observe that since the input graph is a tree, there is only one edge connecting components X_i and X_{i+1} for all i. Consider the last vertex in X_i, viz. v_{ic}. To produce an embedding of the component X_i using Algorithm Tree-Embed, we use a light path ordering of X_i assuming that the subtree containing v_{ic} is the heaviest subtree. Hence v_{ic} is last in the light path ordering of X_i and is visited in the last iteration of the Algorithm Tree-Embed. This makes sure that the distortion of the edge (v_{ic}, v_{ic+1}) is smaller than $2g(c)$. Changing the light path ordering in this way does not affect the bound on the distortion proved in Lemma 3.7. Thus we get the following result.

THEOREM 3.1. *There is a polynomial time algorithm that finds an embedding of an unweighted tree with distortion* $8\Delta\sqrt{c\log c} + 4c$.

COROLLARY 3.1. *There is a polynomial time algorithm that finds an embedding of an unweighted tree with distortion within a factor* $O((n\log n)^{1/3})$ *of the optimal distortion.*

4 A dynamic programming algorithm for graphs of small distortion

Given a connected simple graph $G = (V, E)$ and an integer c, we consider the problem of deciding whether there exists a non-contracting embedding of G into the integer line with maximum distortion at most c.

Note that the maximum distance between any two points in an optimal embedding can be at most $c(n-1)$, and there always exists an optimal embedding with all the nodes embedded into integer coordinates. W.l.o.g., in the rest of this section, we will only consider embeddings of the form $f : V \to \{0, 1, \ldots, c(n-1)\}$. Furthermore, if G admits an embedding of distortion c, then the maximum degree of G is at most $2c$. Thus, we may also assume that G has maximum degree $2c$.

DEFINITION 1. (PARTIAL EMBEDDING) *Let* $V' \subseteq V$. *A* partial embedding *on* V' *is a function* $g : V' \to \{0, 1, \ldots, c(n-1)\}$.

DEFINITION 2. (FEASIBLE PARTIAL EMBEDDING) *Let* f *be a partial embedding on* V'. f *is called* feasible *if there exists an embedding* g *of distortion at most* c, *such that for each* $v \in V'$, *we have* $g(v) = f(v)$, *and for each* $u \notin V'$, *it is* $g(u) > \max_{w\in V'} f(w)$.

DEFINITION 3. (PLAUSIBLE PARTIAL EMBEDDING) *Let* f *be a partial embedding on* V'. f *is called* plausible *if*

- *For each* $u, v \in V'$, *we have* $|f(u) - f(v)| \geq D(u, v)$.
- *For each* $u, v \in V'$, *if* $\{u, v\} \in E$, *then* $|f(u) - f(v)| \leq c$.
- *Let* $L = \max_{v\in V'} f(v)$. *For each* $u \in V'$, *if* $f(u) \leq L - c$, *then for each* $w \in V$ *such that* $\{u, w\} \in E$, *we have* $w \in V'$.

LEMMA 4.1. *If a partial embedding is feasible, then it is also plausible.*

Proof. Let f be a partial embedding over V', such that f is feasible, but not plausible, and let $L =$

$\max_{v \in V'} f(v)$. It follows that there exists $\{u, w\} \in E$, with $u \in V'$, such that $f(u) \leq L - c$, and $w \notin V'$. Since f is feasible, there exists an embedding g of distortion at most c, satisfying $g(u) = f(u) \leq L - c$, and $g(w) > L$. Thus, $|g(u) - g(w)| > c$, a contradiction. □

Definition 4. (Active Region) *Let f be a partial embedding over V'. The* active region *of f is a couple (X, Y), where $X = \{(u_1, f(u_1)), \ldots, (u_{|X|}, f(u_{|X|}))\}$ is a set of $\min\{2c + 1, |V'|\}$ couples, where $\{u_1, \ldots, u_{|X|}\}$ is a subset of V', such that $f(u_i) = \max_{u \in V' \setminus \{u_{i+1}, \ldots, u_{|X|}\}} f(u)$, and Y is the set of all edges in E having exactly one endpoint in V'.*

Lemma 4.2. *Let f_1 be a plausible partial embedding over V_1, and f_2 be a plausible partial embedding over V_2. If f_1 and f_2 have the same active region, then*

- *$V_1 = V_2$.*
- *f_1 is feasible if and only if f_2 is feasible.*

Proof. Let $L = \max_{v \in V'} f(v)$. To prove that $V_1 \subseteq V_2$, assume that there exists $v \in V_1 \setminus V_2$. Let p be a path starting at v, and terminating at some node in $V_1 \cap V_2$, and let v'' be the first node in $V_1 \cap V_2$ visited by p, and v' be the node visited exactly before v''. Clearly, $v' \in V_1 \setminus V_2$, and v' is not in the active region, thus $f_1(v') < L - 2c$. Furthermore, by the definition of a plausible partial embedding, since the edge $\{v'', v'\}$ has exactly one endpoint in V_2, it follows that $f_2(v'') > L - c$. Thus, $|f_1(v') - f_1(v'')| = |f_1(v') - f_2(v'')| > c$, contradicting the fact that f_1 is plausible. Similarly we can show that $V_2 \subseteq V_1$, and thus $V_1 = V_2$.

Assume now that f_1 is feasible, thus there exists an embedding g_1 of distortion at most s, such that for each $v \in V_1$, we have $f_1(v) = g_1(v)$, and for each $v \notin V_1$, we have $g_1(v) > L$. Consider the embedding g_2, where $g_2(u) = f_2(u)$, if $u \in V_2$, and $g_2(u) = g_1(u)$ otherwise. It suffices to show that g_2 is non-contracting and has distortion at most c.

If g_2 has distortion more than c, then since f_2 is a plausible partial embedding, and g_1 has distortion at most c, it follows that there exists an edge $\{u, w\}$, with $u \in V_2$ and $w \notin V_2$, such that $|g_2(u) - g_2(w)| > c$. Since the edge $\{u, w\}$ has exactly one endpoint in V_2, it follows that $f_2(u) > L - c$, and thus u is in the active region, and $f_2(u) = f_1(u)$. Thus, we obtain that $|g_1(u) - g_1(w)| = |g_2(u) - g_2(w)| > c$, a contradiction. Thus, g_2 has distortion at most c.

If g_2 is a contraction, then there exist nodes u and w such that $|g_2(u) - g_2(w)| < D(u, w)$. Since f_2 is plausible, and g_2 is non-contracting, we obtain that exactly one of the nodes u and w is in V_2. W.l.o.g., assume that $u \in V_2$ and $w \notin V_2$, and thus $f_2(u) > L - c$. Thus, u must be in the active region, and we obtain that $f_2(u) = f_1(u)$, and thus $|g_1(u) - g_1(w)| = |g_2(u) - g_2(w)| < D(u, w)$, a contradiction. We have shown that g_2 is non-contracting and has distortion at most c, thus f_2 is feasible. □

Lemma 4.3. *For fixed values of c, the number of all possible active regions for all the plausible partial embeddings is at most $O(n^{4c+2})$.*

Proof. Let f be a plausible partial embedding, with active region (X, Y), such that $|X| = i$. It is easy to see that every edge in Y has exactly one endpoint in X. Since the degree of every node is at most $2c$, after fixing X, the number of possible values for Y is at most 2^{2ic}. Also, the number of possible different values for X is at most $\binom{n}{i}(nc)^i$. Thus, the number of possible active regions for all plausible partial embeddings is at most $\sum_{i=1}^{2c+1} \binom{n}{i}(nc)^i 2^{2ic} = O(n^{4c+2})$. □

Definition 5. (Successor of a Partial Embedding) *Let f_1 and f_2 be plausible partial embeddings on V_1 and V_2 respectively. f_2 is a* successor *of f_1 if and only if*

- *$V_2 = V_1 \cup \{u\}$, for some $u \notin V_1$.*
- *For each $u \in V_1 \cap V_2$, we have $f_1(u) = f_2(u)$.*
- *If $u \in V_2$ and $u \notin V_1$, then $f_2(u) = \max_{v \in V_2} f_2(v)$.*

Let P be the set of all plausible partial embeddings, and let $\hat{P}$ be the set of all active regions of the embeddings in P. Consider a directed graph H with $V(H) = \hat{P}$. For each $\hat{x}, \hat{y} \in V(H)$, $(\hat{x}, \hat{y}) \in E(H)$ if and only if there exist plausible embeddings x, y, such that $\hat{x}$ and $\hat{y}$ are the active regions of x and y respectively, and y is a successor of x.

Lemma 4.4. *Let x_0 be the active region of the empty partial embedding. G admits a non-contracting embedding of distortion at most c, if and only if there exists a directed path from x_0 to some node x in H, such that $x = (X, Y)$, with $X \neq \emptyset$ and $Y = \emptyset$.*

Proof. If there exists a path from x_0 to some node $x = (X, Y)$, with $X \neq \emptyset$ and $Y = \emptyset$, then since $X \neq \emptyset$, it follows that x is not the active region of the empty partial embedding. Furthermore, since G is connected and $Y = \emptyset$, it follows that x is the active region of a plausible embedding f of all the nodes of G. By the definition of a plausible embedding, it follows that f is a non-contracting embedding of G with distortion at most c.

If there exists a non-contracting embedding f of G, with distortion at most c, then we can construct a path in H, visiting nodes $y_0, y_1, \ldots, y_{|V|}$, as follows: For each

i let f_i be the partial embedding obtained from f by considering only the i leftmost embedded nodes, and let y_i be the active region of f_i. Clearly, each f_i is a feasible embedding, and thus by Lemma 4.1, it is also plausible. Moreover, $y_0 = x_0$, and for each $0 < i \leq |V|$, it is easy to see that f_i is a successor of f_{i-1}, and thus $(y_{i-1}, y_i) \in E(H)$. Since, $f_{|V|}$ is an embedding of all the nodes of G, the active region $y_{|V|} = (X_{|V|}, Y_{|V|})$ satisfies $X_{|V|} \neq \emptyset$, and $Y_{|V|} = \emptyset$. □

Using Lemma 4.4, we can decide whether there exists an embedding of G as follows: We begin at node x_0, and we repeatedly traverse edges of H, without repeating nodes. Note that we do not compute the whole H from the beginning, but we instead compute only the neighbors of the current node. This is done as follows: At each step i, we maintain a plausible partial embedding g_i, such that each partial embedding induced by the j leftmost embedded nodes in g_i, has active region equal to the jth node in the path from x_0 to the current node. We consider all the plausible embeddings obtained by adding a rightmost node in g_i. The key property is that by Lemma 4.2, the active regions of these embeddings are exactly the neighbors of the current node. This is because an active region completely determines the subset of embedded nodes, as well as the feasibility of such a plausible embedding. By Lemma 4.3, the above procedure runs in polynomial time when s is fixed.

THEOREM 4.1. *For any fixed integer c, we can compute in polynomial time a non-contracting embedding of G, with distortion at most c, if one exists.*

5 Hardness of approximation

In this section we show that the problem of computing minimum distortion embedding of unweighted graphs is NP-hard to a-approximate for certain $a > 1$. This is done by a reduction from TSP over $(1, 2)$-metrics. Recall that the latter problem is NP-hard to approximate up to some constant $a > 1$.

Recall that a metric $M = (V, D)$ is a $(1, 2)$-metric, if for all $u, v \in V$, $u \neq v$, we have $D(u, v) \in \{1, 2\}$. Let $G(M)$ be a graph (V, E) where E contains all edges $\{u, v\}$ such that $D(u, v) = 1$.

The reduction F from the instances of TSP to the instances of the embedding problem is as follows. For a $(1, 2)$-metric M, we first compute $G = (V, E) = G(M)$. Then we construct a copy $G' = (V', E')$ of G, where V' is disjoint from V. Finally, we add a vertex o with an edge to all vertices in $V \cup V'$. In this way we obtain the graph $F(M)$.

The properties of the reduction are as follows.

LEMMA 5.1. *If there is a tour in M of length t, then $F(M)$ can be embedded into a line with distortion at most t.*

Proof. The embedding $f : F(M) \to \Re$ is constructed as follows. Let $v_1, \ldots, v_n, v_1$ be the sequence of vertices visited by a tour T of length t. The embedding f is obtained by placing the vertices V in the order induced by T, followed by the vertex o and then the vertices V'. Formally:

- $f(v_1) = 0$, $f(v_i) = f(v_{i-1}) + D(v_{i-1}, v_i)$ for $i > 1$
- $f(o) = f(v_n) + 1$
- $f(v'_1) = f(o) + 1$, $f(v'_i) = f(v'_{i-1}) + D(v'_{i-1}, v'_i)$ for $i > 1$

It is immediate that f is non-contracting. In addition, the maximum distortion (of at most t) is achieved by the edges $\{o, v_1\}$ and $\{o, v'_n\}$. □

LEMMA 5.2. *If there is an embedding f of $F(M)$ into a line that has distortion s, then there is a tour in M of length at most $s + 1$.*

Proof. Let $H = F(M)$. Let $U = u_1 \ldots u_{2n}$ be the sequence of the vertices of $V \cup V'$ in the order induced by f. Partition the range $\{1 \ldots 2n\}$ into maximal intervals $\{i_0 \ldots i_1 - 1\}, \{i_1 \ldots i_2 - 1\}, \ldots, \{i_{k-1} \ldots i_k - 1\}$, such that for each interval I, the set $\{u_i : i \in I\}$ is either entirely contained in V, or entirely contained in V'. Recall that H has diameter 2. Since f has distortion s, it follows that $|f(u_1) - f(u_{2n})| \leq 2s$. Moreover, from non-contraction of f it follows that $|f(u_{i_j - 1}) - f(u_{i_j})| = 2$ for all j. It follows that if we swap any two subsequences of U corresponding to different intervals I and I', then the resulting mapping of $V \cup V'$ into $\Re$ is still non-contracting (with respect to the metric induced by H). Therefore, there exists a mapping f' of $V \cup V'$ into $\Re$ which is non-contracting, in which all vertices of V precede all vertices of V', and such that the diameter of the set $f'(V \cup V')$ is at most $2s$. Without loss of generality, assume that the diameter Δ of $f'(v)$ is not greater than the diameter of $f'(V')$. This implies that $\Delta \leq (2s - 2)/2 = s - 1$. Therefore, the ordering of the vertices in V induced by f' corresponds to a tour in M of length at most $\Delta + 2 \leq s + 1$. □

COROLLARY 5.1. *There exists a constant $a > 1$ such that a-approximating the minimum distortion embedding of an unweighted graph is NP-hard.*

6 Embedding spheres into the plane

Let $M = (X, D)$ be a metric induced by a set X of n points on a unit sphere S^2, under the Euclidean distance

in $\mathbb{R}^3$. Let $c_p^d(M)$ denote the minimum distortion of any embedding of M into l_p^d.

THEOREM 6.1. *If $M = (X, D)$ is the metric induced by a set X of n points on a unit sphere S^2, under the Euclidean distance in $\mathbb{R}^3$, then $c_2^2(M) = O(\sqrt{n})$.*

Proof. Since the size of the surface of S^2 is constant, it follows that there exists a cap K in S^2, of size $\Omega(1/n)$, such that $X \cap K = \emptyset$. Let p_0 be the center of K on S^2, and p_0' be its antipode. By rotating S^2, we may assume that $p_0 = (0, 0, 1)$, and thus $p_0' = (0, 0, -1)$.

For points $p, p' \in S^2$, let $\rho_S(p, p')$ be the geodesic distance between p and p' in S^2. Consider the mapping $f : X \to \mathbb{R}^2$, such that for every point $p \in X$, with $p = (x, y, z)$, we have $f(p) = \left(\rho_S(p, p_0')\frac{x}{\sqrt{x^2+y^2}}, \rho_S(p, p_0')\frac{y}{\sqrt{x^2+y^2}}\right)$, if $p \neq p'$, and $f(p) = (0, 0)$, if $p = p'$. It is straightforward to verify that f is non-contracting.

CLAIM 1. *The expansion of f is maximized for points p, q, on the perimeter of K, which are antipodals with respect to K.*

Proof. Let $p, q \in S^2$. W.l.o.g., we assume that $p = (0, \sin\phi_p, 1 + \cos\phi_p)$, and $q = (\sin\phi_q \sin\theta_q, \sin\phi_q \cos\theta_q, 1 + \cos\phi_q)$, for some $0 \leq \phi_p, \phi_q \leq \phi$, and $0 \leq \theta_q \leq \pi$. The images of p and q are $f(p) = (0, \phi_p)$, and $f(q) = (\phi_q \sin\theta_q, \phi_q \cos\theta_q)$, respectively. Let $h = \frac{\|f(p)-f(q)\|}{\|p-q\|}$, be the expansion of f in the pair p, q. We obtain:

$$h^2 = \frac{\phi_q^2 + \phi_p^2 - 2\phi_q\phi_p\cos\theta_q}{2 - 2\cos\phi_p\cos\phi_q - 2\sin\phi_p\sin\phi_q\cos\theta_q}$$

Observe that since $\sin\phi_p \leq \phi_p$, and $\sin\phi_q \leq \phi_q$, it follows that h^2 is maximized when $\cos\theta_q$ is minimized. That is, the expansion is maximized for $\theta_q = \pi$.

Thus, we can assume that the expansion of f is maximized for points $p, q \in S^2$, with $p = (0, \sin\phi_p, 1 + \cos\phi_p)$, and $q = (0, -\sin\phi_q, 1 + \cos\phi_q)$. For such points, the expansion is $\frac{\phi_p+\phi_q}{2\sin\frac{\phi_p+\phi_q}{2}}$. It follows that the expansion is maximized when $\phi_p + \phi_q$ is maximized, which happens when p and q are on the perimeter of K. □

We pick p and q on the perimeter of K, such that p is the antipode of q w.r.to K. Let ϕ_K be the angle of K, and set $r_K = \phi_K/2$. We have $r_K = \Omega(1/\sqrt{n})$, and $\|f(p)-f(q)\| = 2\pi - 2r_K$, while $\|p-q\| = 2\sin r_K$. Thus, the expansion is at most $\frac{\pi-r_K}{\sin r_K}$. W.l.o.g., we can assume that $r_K \leq \pi/2$, since otherwise we can simply consider a smaller cap K. Thus, $\frac{\pi-r_K}{\sin r_K} \leq 2\frac{\pi-r_K}{\pi r_K} < \frac{2}{r_K} = O(\sqrt{n})$. Since the embedding is non-contracting, it follows that the expansion is $O(\sqrt{n})$. □

THEOREM 6.2. *There exists a metric $M = (X, D)$, induced by a set X of n points on a unit sphere S^2, under the Euclidean distance in $\mathbb{R}^3$, such that any mapping $f : X \to \mathbb{R}^2$ has distortion $\Omega(\sqrt{n})$.*

Proof. Let $X \subset S^2$ be a set of n points, such that X is a $O(1/\sqrt{n})$-net of S^2, and let $f : X \to \mathbb{R}^2$ be a non-expanding embedding. Since $S^2 \subset \mathbb{R}^3$, by Kirszbraun's Theorem ([Kir34], see also [LN]), we obtain that f can be extended to a non-expanding mapping $f' : S^2 \to \mathbb{R}^2$. Also, by the Borsuk-Ulam Theorem, it follows that there exist antipodals $p, q \in S^2$, such that $f'(p) = f'(q)$. Since X is an $O(1/\sqrt{n})$-net, there exist points $p', q' \in X$, such that $\|p - p'\| = O(1/\sqrt{n})$, and $\|q-q'\| = O(1/\sqrt{n})$. Since f is non-expanding, it follows that $\|f(p') - f(q')\| = O(1/\sqrt{n})$. On the other hand, we have $\|p - q\| = 2$, and thus $\|p' - q'\| = \Omega(1)$. Thus, f has distortion $\Omega(\sqrt{n})$. □

THEOREM 6.3. *There exists a polynomial-time, 3.512-approximation algorithm, for the problem of embedding a finite sub-metric of S^2 into $\mathbb{R}^2$.*

Proof. [Proof sketch] We apply the embedding of Theorem 6.1, by choosing K to be the largest empty cap in S^2. Let r_K be the radius of K. By using an analysis similar to the one of Theorem 6.1, we obtain that the distortion of the embedding is at most $\frac{\pi-r_K}{\sin r_K}$. Moreover, by using the analysis of Theorem 6.2, we can show that the distortion of an optimal embedding is at least $\max\{1, \frac{\cos r_K}{2\sin\frac{r_K}{2}}\}$. By simple calculations, we obtain that the distortion is maximized for $r_K = 2\tan^{-1}\frac{(\sqrt{3}-1)3^{3/4}\sqrt{2}}{6} \approx 0.749$, for which we obtain that the approximation ratio is less than 3.512. □

For the case where the metric $M = (X, D)$ corresponds to the geodesic distances between the points of the sphere, we can show using the same techniques that the algorithm of Theorem 6.3, is in fact a 3-approximation.

7 Acknowledgments

This work is a combined version of two earlier papers by Badoiu, Rabinovich & Sidiropoulos, and by Dhamdhere, Gupta, Räcke & Ravi which obtained nearly identical results. We would like to thank the SODA committee to facilitate the merger, and Piotr Indyk for numerous comments and insights on the problems discussed in this paper.

References

[ABFC+96] R. Agarwala, V. Bafna, M. Farach-Colton, B. Narayanan, M. Paterson, and M. Thorup. On the approximability of numerical taxonomy: (fitting distances by tree metrics). *7th Symposium on Discrete Algorithms*, 1996.

[BIR03] M. Badoiu, P. Indyk, and Y. Rabinovich. Approximate algorithms for embedding metrics into low-dimensional spaces. *Manuscript*, 2003.

[BJDG+03] Z. Bar-Joseph, E. D. Demaine, D. K. Gifford, A. M. Hamel, Tommi S. Jaakkola, and Nathan Srebro. K-ary clustering with optimal leaf ordering for gene expression data. *Bioinformatics*, 19(9):1070–8, 2003.

[Bor33] K. Borsuk. Drei Sätze über die n-dimensionale euklidische Sphäre. *Fund. Math.*, 20:177–190, 1933.

[Bou85] J. Bourgain. On lipschitz embedding of finite metric spaces into hilbert space. *Isreal Journal of Mathematics*, 52:46–52, 1985.

[Bŏŏ3] M Bădoiu. Approximation algorithm for embedding metrics into a two-dimensional space. *14th Annual ACM-SIAM Symposium on Discrete Algorithms*, 2003.

[Dha04] Kedar Dhamdhere. Approximating additive distortion of embeddings into line metrics. *7th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX)*, 2004.

[FCKW93] M. Farach-Colton, S. Kannan, and T. Warnow. A robust model for finding optimal evolutionary tree. *Proceedings of the Symposium on Theory of Computing*, 1993.

[Fei00] U. Feige. Approximating the bandwidth via volume respecting embeddings. *Journal of Computer and System Sciences*, 60(3):510–539, 2000.

[Gup00] Anupam Gupta. Improved bandwidth approximation for trees. *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms*, 2000.

[HIL98] J. Hastad, L. Ivansson, and J. Lagergren. Fitting points on the real line and its application to rh mapping. *Lecture Notes in Computer Science*, 1461:465–467, 1998.

[Ind01] P. Indyk. Tutorial: Algorithmic applications of low-distortion geometric embeddings. *Proceedings of the Symposium on Foundations of Computer Science*, 2001.

[Iva00] L. Ivansson. Computational aspects of radiation hybrid. *Doctoral Dissertation, Department of Numerical Analysis and Computer Science, Royal Institute of Technology*, 2000.

[Kir34] M. D. Kirszbraun. Über die zusammenziehenden und lipschitzschen Transformationen. *Fund. Math.*, 22:77–108, 1934.

[KRS04] C. Kenyon, Y. Rabani, and A. Sinclair. Low distortion maps between point sets. *Proceedings of the Symposium on Theory of Computing*, 2004. to appear.

[Kru64a] J.B. Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29:1–27, 1964.

[Kru64b] J.B. Kruskal. Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 29:115–129, 1964.

[LLR94] N. Linial, E. London, and Y. Rabinovich. The geometry of graphs and some of its algorithmic applications. *Proceedings of 35th Annual IEEE Symposium on Foundations of Computer Science*, pages 577–591, 1994.

[LN] J. R. Lee and A. Naor. Absolute lipschitz extendability. *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics*. to appear.

[Mat90] J. Matoušek. Bi-lipschitz embeddings into low-dimensional euclidean spaces. *Comment. Math. Univ. Carolinae*, 31:589–600, 1990.

[Mat96] J. Matoušek. On the distortion required for embedding finite metric spaces into normed spaces. *Israel Journal of Mathematics*, 93:333–344, 1996.

[Sax80] J. B. Saxe. Dynamic-programming algorithms for recognizing small-bandwidth graphs in polynomial time. *SIAM J. Algebraic Discrete Methods*, 1:363–369, 1980.

[She62a] R. N. Shepard. The analysis of proximities: Multidimensional scaling with an unknown distance function 1. *Psychometrika*, 27:125–140, 1962.

[She62b] R. N. Shepard. The analysis of proximities: Multidimensional scaling with an unknown distance function 2. *Psychometrika*, 27:216–246, 1962.

[Wor] Working Group on Algorithms for Multidimensional Scaling. Algorithms for multidimensional scaling. http://dimacs.rutgers.edu/Workshops/Algorithms/.