APPROXIMATION ALGORITHMS FOR LOW-DISTORTION EMBEDDINGS INTO LOW-DIMENSIONAL SPACES[*]

ANASTASIOS SIDIROPOULOS[†], MIHAI BADOIU[‡], KEDAR DHAMDHERE[§],
ANUPAM GUPTA[¶], PIOTR INDYK[‖], YURI RABINOVICH[#], HARALD RACKE[†‡] AND
R. RAVI[‡‡]

Abstract. We present several approximation algorithms for the problem of embedding metric spaces into a line, and into the 2-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 induced by unweighted trees.

Key words. metric embeddings, distortion, line, sphere, approximation algorithms

AMS subject classifications. 68W25, 68W05, 68Q25

DOI. 10.1137/17M1113527

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* 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 $M = M(G) =$

[†]Department of Computer Science, University of Illinois at Chicago, Chicago, IL 60607 (sidiropo@gmail.com).

[‡]Edgestream Partners L. P., Princeton, NJ 08540 (mihai@mit.edu).

[§]Google Inc., Mountain View, CA 94043 (kedar@cs.cmu.edu).

[¶]Department of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213 (anupamg@cs.cmu.edu).

[‖]MIT, Cambridge, MA 02139 (indyk@mit.edu).

[#]Department of Computer Science, University of Haifa, Haifa, Israel (yuri@cslx.haifa.ac.il).

[††]Technische Universität München, München, Germany (harald.raecke@tum.de).

[‡‡]CMU, Pittsburgh, PA 15213 (ravi@andrew.cmu.edu).

(V, D), find a mapping f of V into a *line* that is noncontracting (i.e., $|f(u) - f(v)| \geq D(u, v)$ for all $u, v \in V$) and minimizes the distortion

$$c_{\text{line}}(M, f) = \max_{u,v \in V} \frac{|f(u) - f(v)|}{D(u, v)}.$$

That is, our goal is to find $c_{\text{line}}(M) = \min_f c_{\text{line}}(M, f)$. 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) $O(c)$-approximation algorithm for metrics M for which $c_{\text{line}}(M) \leq c$. 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 $n^{O(c_{\text{line}}(M))}$ (section 4).

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

We also study the problem of embedding metrics into the *plane* in section 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-approximation algorithm, when D is the geodesic distance in S^2. For the case where D corresponds to the Euclidean distance in $\mathbb{R}^3$, our algorithm can be reanalyzed to give an approximation guarantee of 3.512.

1.1. Related work. Combinatorial versus 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 construc-

---

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

tion [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 of low dimension. 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, Ivansson, and Lagergren, gave a 2-approximation algorithm for embedding an arbitrary metric into a line $\mathbb{R}$, 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 [Bǒ03]. Bădoiu, Indyk and Rabinovich [BIR03] gave a weakly-quasi-polynomial time algorithm for the same problem in the l_2 norm.

In general, one can choose nongeometric 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 evolutionary path that led to formation of the genetic sequences. Motivated by these applications Farach-Colton and 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 error [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 versus 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 noncontraction 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 $c_{\text{line}}(M(G))$ 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 noncontraction 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].

More recent work. Since the conference version of this paper appeared, several results on exact and approximation algorithms for minimum distortion embeddings have been obtained. Most notably, embeddings for general metrics into the real line were obtained in [BCIS05, NR15]. For the case where the spread is polynomial and the optimum distortion is constant, the results in [NR15] give a quasi-polynomial exact algorithm and a polynomial time $O(1)$-approximation. It has also been shown that for unweighted graphs the problem of minimum distortion embedding into the line is fixed-parameter tractable, parameterized by the distortion [FFL+13]. Specifically, Fellows et al. [FFL+13] have shown that deciding whether an n-vertex graph admits an

embedding with distortion at most c can be done in time $2^{O(c \log c)} n^{O(1)}$. Furthermore Lokshtanov, Marx, and Saurabh [LMS18] have shown that there is no algorithm with running time $2^{o(c \log c)} n^{O(1)}$, unless the exponential time hypothesis fails. An exact algorithm with running time $2^{O(n)}$ has also been obtained [FLS09]. Finally, structural properties of minimum-distortion embeddings for unweighted trees into the line have been obtained in [CK11].

There has also been a series of papers for the case of embedding into low-dimensional spaces. Approximation algorithms for embedding ultrametrics into constant-dimensional Euclidean space have been obtained [BCIS06, OS08]. The algorithm in [BCIS06] implies an $O(1)$-factor approximation when the optimum is constant, and the result in [OS08] is an $O(\log n)$-approximation for the case of polynomial spread. On the lower bound side, it has been shown by [MS10] that for any fixed $d \geq 3$, it is NP-hard to approximate the minimum distortion embedding into d-dimensional Euclidean space within a factor better than $\Omega(n^{\alpha/d})$ for some fixed $\alpha > 0$. Moreover, for any fixed $d \geq 2$, it is NP-hard to distinguish whether the minimum distortion is at most $O(1)$, or at least $\Omega(n^{\beta/d})$, for some fixed β. In [ESZ10] it has been shown that for embedding into $\mathbb{R}^2$, it is NP-hard to distinguish whether the minimum distortion is at most c, or at least c', for some constants $0 < c < c'$.

Approximation algorithms for embedding into trees have also been considered. For the case of embedding the shortest-path metric of an unweighted graph, an $O(1)$-approximation for the case of constant distortion has been obtained [BIS07]. For the case of embedding general metrics into trees, the algorithm in [BIS07] gives a $n^{o(1)}$-approximation for the case of constant distortion and polynomial spread.

We finally remark that many similar questions for computing a minimum-distortion bijection between two given finite metric spaces have also been considered [KRS09, HP05, PS05, KS07].

2. An $O(c)$-approximation algorithm. We start by stating an algorithmic version of a fact proved in [Mat90].

LEMMA 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 $O(|V|+|E|)$. Moreover the embedding is noncontracting, has expansion at most $2n-1$, and its image is contained inside an interval of length at most $2n-1$.*

Proof. Let T be a spanning tree of the graph. We replace every (undirected) edge of T with a pair of oppositely 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 noncontracting. Since C has length $2n$, it follows that the image of the embedding is contained inside an interval of length at most $2n-1$, and thus the expansion is at most $2n-1$. Therefore the distortion is at most $2n-1$. □

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

Let $G = (V, E)$ be a graph, such that there exists an embedding of G with 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 $f_{OPT}(t_2) = \max_{v \in V} f_{OPT}(v)$.
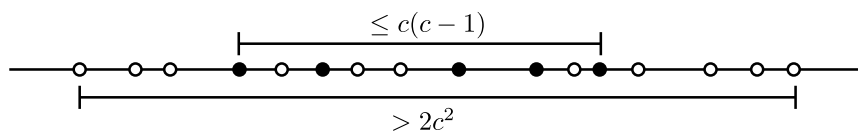2. Compute the shortest path $p = v_1, v_2, \ldots, v_L$ from t_1 to t_2.

FIG. 1. *Depiction of the proof of Lemma* 3. *The points in* $f_{OPT}(\{v_i, \ldots, v_{i+c-1}\})$ *and* $f_{OPT}(V_i \cup \cdots \cup V_{i+c-1})$ *are depicted as black and white disks, respectively.*

3. Partition V into disjoint sets $V_1, V_2, \ldots, V_L$, such that for each $u \in V_i$,

$$D(u, v_i) = \min_{1 \le j \le L} D(u, v_j).$$

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 1, leaving a space of length $|V_i| + |V_{i+1}| + 1$ between the nodes of V_i and V_{i+1}.

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

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 leftmost and rightmost embedded nodes in the embedding f_{OPT}, it follows that there exists $j, 1 \le j < L$, such that $f_{OPT}(x)$ lies between $f_{OPT}(v_j)$ and $f_{OPT}(v_{j+1})$. Without loss of generality (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 $|f_{OPT}(v_{j+1}) - f_{OPT}(v_j)| = f_{OPT}(v_{j+1}) - f_{OPT}(x) + f_{OPT}(x) - f_{OPT}(v_j) \ge D(v_{j+1}, x) + D(x, v_j) \ge 2D(x, v_i) > c$. This is a contradiction, since the distortion of f_{OPT} is at most c. □

LEMMA 3. *For every* i, $1 \le i \le L - c + 1$, *we have* $\sum_{j=i}^{i+c-1} |V_j| \le 2c^2$.

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

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

$$\le \max_{i \le j_1 < j_2 \le i+c-1} \left(\sum_{j=j_1}^{j_2-1} |f_{OPT}(v_j) - f_{OPT}(v_{j+1})| \right)$$

$$\le \max_{i \le j_1 < j_2 \le i+c-1} \left(\sum_{j=j_1}^{j_2-1} cD(v_j, v_{j+1}) \right)$$

$$\le c(c - 1),$$

where the first inequality follows by the triangle inequality, and the second line by the fact that f_{OPT} has expansion at most c. Moreover, since $\sum_{j=i}^{i+c-1} |V_j| > 2c^2$, we have $\max_{u,w \in \bigcup_{j=i}^{i+c-1} V_j} |f_{OPT}(u) - f_{OPT}(w)| \ge 2c^2$. In other words, $f_{OPT}(\{v_i, \ldots, v_{i+c-1}\})$ is contained in some interval of length at most $c(c - 1)$, while the minimum interval containing $f_{OPT}(V_i \cup \cdots \cup V_{i+c-1})$ has length greater than $2c^2$ (see Figure 1). It follows that there exists $u \in V_l$, for some l, with $i \le l \le i + c - 1$, such that $|f_{OPT}(v_l) - f_{OPT}(u)| \ge \frac{2c^2 - c(c-1)}{2} > c^2/2$. Since the distortion is at most c, we have $D(v_l, u) > c/2$, contradicting Lemma 2. □

LEMMA 4. *The embedding computed by the algorithm is noncontracting.*

Proof. Let $x, y \in V$. If x and y are in the same set V_i, for some i, then since the embedding computed by Lemma 1 is noncontracting, we get $|f(x) - f(y)| \ge D(x, y)$.

Assume now that $x \in V_i$ and $y \in V_j$ for some $i < j$. We have

$$
\begin{aligned}
|f(y) - f(x)| &\geq \sum_{t=i}^{j-1} (|V_t| + |V_{t+1}| + 1) \\
&\geq |V_i| + (j - i) + |V_j| \\
&\geq D(x, v_i) + D(v_i, v_j) + D(v_j, y) \\
&\geq D(x, y). \qquad \square
\end{aligned}
$$

LEMMA 5. *The expansion of the embedding computed by the algorithm is at most* $16c^2$.

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 $y \in V_j$. Consider first the case $|i - j| \leq 2c$. We may assume w.l.o.g. that $i \leq j$. Then by Lemma 3 we have $\sum_{\ell=i}^{j} |V_j| \leq 4c^2$. By Lemma 1 we have that for each $\ell \in \{i, \ldots, j\}$, $f(V_\ell)$ is contained in some interval of length at most $2|V_\ell| - 1$. Moreover, by the description of the algorithm, we have that for each $\ell' \in \{i, \ldots, j-1\}$, there is a gap of length $|V_{\ell'}| + |V_{\ell'+1}| + 1$ between $f(V_{\ell'})$ and $f(V_{\ell'+1})$. Therefore $f(V_i \cup \cdots \cup V_j)$ is contained in an interval of length at most $(\sum_{\ell=i}^{j} 2|V_\ell| - 1) + (\sum_{\ell'=i}^{j-1} |V_{\ell'}| + |V_{\ell'+1}| + 1) \leq 4 \sum_{\ell=i}^{j} |V_\ell| \leq 16c^2$. Since $\{x, y\} \subseteq V_i \cup \cdots \cup V_j$, it follows that $|f(x) - f(y)| \leq 16c^2$, and thus the expansion is at most $16c^2$ in this case.

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, we obtain that $D(v_i, x) \leq c/2$ and $D(y, v_j) \leq c/2$, and thus by the triangle inequality we get $|i - j| = D(v_i, v_j) \leq D(v_i, x) + D(x, y) + D(y, v_j) \leq c + 1$, a contradiction. $\qquad \square$

We use the following straightforward lower bound on the optimal distortion, which will also be used in subsequent sections. 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 6 (local density). *Let G denote a graph with local density Δ. Then any map of G into the line has distortion at least Δ.*

Proof. Let f_{OPT} be any optimal embedding of distortion c. Suppose for the sake of contradiction that $\Delta > c$. Thus there exists some vertex u in G, and some $r > 0$, such that $|B(u, r)| - 1 > 2rc$. Let $V^- = \{v \in B(u, r) : f_{OPT}(v) < f_{OPT}(u)\}$ and $V^+ = \{v \in B(u, r) : f_{OPT}(v) > f_{OPT}(u)\}$. We may assume w.l.o.g. that $|V^-| \geq |V^+|$, since otherwise we may consider the optimal embedding $-f_{OPT}$. Since $|V^+ \cup V^-| = B(u, r) \setminus \{r\}$, we get $|V^-| \geq (|B(u, r)| - 1)/2 > rc$. Let $V^- = \{v_1, \ldots, v_\ell\}$, where $f_{OPT}(v_1) < \cdots < f_{OPT}(v_\ell)$. By the triangle inequality and the fact that f_{OPT} is noncontracting, we get

$$
\begin{aligned}
|f_{OPT}(v_1) - f_{OPT}(u)| &= |f_{OPT}(v_\ell) - f_{OPT}(u)| + \sum_{i=1}^{\ell-1} |f_{OPT}(v_i) - f_{OPT}(v_{i+1})| \\
&\geq \ell = |V^-| > rc \geq cD(v_1, u),
\end{aligned}
$$

which contradicts the fact that f_{OPT} has expansion at most c, and concludes the proof. $\qquad \square$

COROLLARY 1. *The maximum degree of G is at most $2c$.*

Proof. The proof follows immediately from Lemma 6 by considering the local density for balls of radius $r = 1$. □

THEOREM 1. *The described algorithm computes a noncontracting embedding of maximum distortion $O(c^2)$ in time $O(n^3c)$.*

Proof. By Lemmas 4 and 5, it follows that the computed embedding is noncontracting 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. By Corollary 1 we have $|E| = O(cn)$. Thus, the total running time is $O(n^2|E|) = O(n^3c)$. □

THEOREM 2. *There exists an $O(\sqrt{n})$-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 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 $\operatorname{diam}(V_i) \le c$ and $\sum_{j=i}^{i+c-1} |V_j| \le 2c^2$. 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. We also denote by T_i the subtree of T induced on X_i. Lemma 2 gives the following bound on the diameter of the set X_i.

LEMMA 7. *The diameter of the set X_j (for $j = 1, 2, \ldots$) is at most $2c$.*

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 8 (prefix embeddings). *Given any graph G, there exists an embedding of G into the real line with the following two properties:*
1. *Walking 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.) W.l.o.g., 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, while 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 distortion at most $2c$, proving the lemma.   □

Henceforth, we will only consider embeddings that satisfy the properties stated in Lemma 8. 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 $O(g(c))$, where $g(c) = 2\Delta\sqrt{c \log c} + c$, Δ is the local density, and c the optimal distortion.

We first describe an embedding of X_j, for any integer j. At the end of this section we explain how these embeddings can be combined to obtain an embedding of T. In order to simplify notation, we fix some integer j and focus on the set X_j and the tree T_j.

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 breadth first search (*BFS) part* of the round) from that of laying out new vertices (which we call the depth first search (*DFS) part*).

We assume that we know the leftmost 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_j. If T_{j+1} is empty, then the light path ordering is a DFS ordering which starts at root r of T_j and at each point enters a subtree with the smallest number of vertices in it. Otherwise, if T_{j+1} is nonempty, let v^* be the vertex of T_j that has the root of T_{j+1} as a child; then the light path ordering is a DFS ordering which starts at root r of T_j and at each point considers the following three cases: (i) if v^* is not a descendant of the current vertex then it enters a subtree with the smallest number of vertices in it; (ii) if v^* is the only descendant then it enters the subtree rooted at v^*; and (iii) otherwise it enters the subtree with the smallest number of vertices, excluding the subtree rooted at v^*.

For some $Y \subset X_j$, we define the *induced spanning tree on Y* to be the minimum subtree T' of T that contains all the vertices in Y. We remark also that $T' \subseteq T_j$. We also define *the span of the induced spanning tree on Y*, or simply *the span of Y*, to be the number of vertices in T'.

Initially, we consider all vertices as being not *visited*. At certain points during the execution of the algorithm some vertices are marked as visited; we say that the algorithm *visits* those vertices at that point of the execution.

The embedding is computed inductively, from left to right, as follows. The algorithm proceeds in iterations. During each iteration, the algorithm computes the

embedding of two new subsets of vertices. For a subset Y, we say that the algorithm visits Y to denote the following procedure: We compute the induced spanning tree T' on Y; we then compute an embedding f_Y of T' using the algorithm in Lemma 1; finally, we extend the current embedding to Y by placing the image of f_Y to the right of the image of the current embedding, leaving the minimum possible gap to ensure that the resulting map is noncontracting.

Algorithm Tree-Embed.

1. let $C \leftarrow \{r\}$ denote the set of vertices
    already visited. Set $i \leftarrow 1$.
2. while $C \neq V(T_j)$ do
    (Round i BFS)
3.   Visit all vertices in $N(C) \setminus C$;
    let $C \leftarrow C \cup N(C)$
    (Round i DFS)
4.   set B to be a set of the first $g(c)$ vertices
    of $V(T_j) \setminus C$ in the light path ordering that have not yet been visited.
    Visit all vertices in B; let $C \leftarrow C \cup B$.
    Set $i \leftarrow i + 1$.
    endwhile

LEMMA 9 (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 vertices of the tree will be visited. $\square$

The heart of the proof is to show that visiting the vertices in steps 3 and 4 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 10 (span of boundary). *The span of $N(C) \setminus C$ is at most $g(c)$.*

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 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 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 $N(C_i) \backslash C_i$ 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 $N(C_i) \backslash C_i$ requires only a path of length c. Hence, the total span of vertices in $N(C_i) \backslash C_i$ is bounded by $g(c)$. □

LEMMA 11. *The span of the set of vertices visited in any iteration is bounded by* $2g(c)$.

Proof. From Lemma 10, 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 4 of the algorithm is bounded by $g(c)$. Since, we visit a set of connected components, their span is bounded by $g(c) + \text{span}(N(C) \backslash C)$. Therefore, the span of the vertices visited in each iteration is bounded by $2g(c)$. □

LEMMA 12. *The distortion of the embedding produced by Algorithm Tree-Embed is at most* $O(g(c))$.

Proof. First we argue that the embedding is noncontracting. For each subset of vertices Y such that the algorithm visits Y at some iteration, the image of Y is given by (a translation of) the embedding computed in Lemma 1, and is thus noncontracting. By the description of the visiting process we have that the images of the maps for different such sets Y are combined inductively by leaving a sufficiently large gap to ensure that the resulting map is noncontracting. By induction on the number of subsets that the algorithm visits, it follows that the final map is noncontracting.

During any iteration, the algorithm visits two subsets of vertices. By Lemma 11, each of these subsets has span at most $O(g(c))$. Since the embedding obtained by visiting a subset Y is constructed by following a traversal of the induced spanning tree on Y, arguing as in the proof of Lemma 1, it follows that its image is contained inside some interval of length $O(g(c))$. It follows that the image of the set of vertices visited in the same iteration is contained in some interval of length $O(g(c))$. Therefore, the distortion of any pair of vertices that are visited in the same iteration is at most $O(g(c))$. 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 $O(g(c))$. Hence, the distortion is bounded by $O(g(c))$. □

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. Let $\{x, x'\}$ be this edge, with $x \in X_i$, $x' \in X_{i+1}$. By the definition of the light path ordering we have that x is visited during the last iteration of the Algorithm Tree-Embed. This makes sure that the distortion of the edge $\{x, x'\}$ is smaller than $O(g(c))$. Thus we get the following result.

THEOREM 3. *There is a polynomial time algorithm that finds an embedding of an unweighted tree with distortion* $O(\Delta\sqrt{c\log c} + c)$.

COROLLARY 2. *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 noncontracting 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 $f : V \to \{0, 1, \dots, c(n-1)\}$. 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* $g : V' \to \{0, 1, \dots, c(n-1)\}$.

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 13. *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. $\qquad\square$

We now introduce the notion of an *active region*. Intuitively, this is a small piece of information $a(f)$, defined for some plausible partial embedding f, such that $a(f)$ uniquely defines the domain of f, and is enough to decide whether f is feasible. This allows us to avoid enumerating the set of all plausible partial embeddings, and instead focus on the set of all active regions; as we shall see, this is a significant smaller set.

DEFINITION 4 (active region). *Let* f *be a partial embedding over* V'. *The* active region *of* f *is a pair* (X, Y), *where* $X = \{(u_1, f(u_1)), \dots, (u_{|X|}, f(u_{|X|}))\}$, *with* $|X| \leq \min\{2c+1, |V'|\}$, *where* $\{u_1, \dots, u_{|X|}\} \subseteq V'$, *such that*

$$f(u_i) = \max_{u \in V' \setminus \{u_{i+1}, \dots, u_{|X|}\}} f(u),$$

and Y *is the set of all edges in* E *having exactly one endpoint in* V'.

LEMMA 14. *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, for the sake of contradiction, that there exists $v \in V_1 \setminus V_2$. Let p be a path starting at v, and terminating at some node in $V_1 \cap V_2$; note that such a path exists because every vertex in V_1 has some path that leads to a vertex in the active region, and every vertex in the active region must be in $V_1 \cap V_2$. Let v'' be the first node in $V_1 \cap V_2$ visited by p, and $v' \in V_1 \setminus V_2$ be the node visited exactly before v''. Since $v' \in V_1 \setminus V_2$, it follows that v' does not appear in the active region, and 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 $f_2(v'') > L - c$. Since f_1 and f_2 have the same active regions, and v'' appears in this active region, it follows that $f_1(v'') = f_2(v'')$. We thus obtain that $|f_1(v') - f_1(v'')| = |f_1(v') - f_2(v'')| > c$, 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 c, 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 noncontracting 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, and f_2 is feasible. $\quad\square$

LEMMA 15. *For fixed values of c, the number of all possible active regions of 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 $\sum_{i=1}^{2c+1} \binom{n}{i}(nc)^i 2^{2ic} = O(n^{4c+2})$. $\quad\square$

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 16. *Let x_0 be the active region of the empty partial embedding. G admits a noncontracting 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 noncontracting embedding of G with distortion at most c.

If there exists a noncontracting 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 13, 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 $Y_{|V|} = \emptyset$. $\square$

Using Lemma 16, 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 14, 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 15, the above procedure runs in polynomial time when c is fixed.

THEOREM 4. *For any fixed integer c, we can compute in polynomial time a noncontracting 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 the traveling salesman problem (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 the 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 17. *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 $f : F(M) \to \mathbb{R}$ 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 noncontracting. In addition, the maximum distortion (of at most t) is achieved by the edges $\{o, v_1\}$ and $\{o, v_n'\}$. □

LEMMA 18. *If there is an embedding f of $F(M)$ into a line that has distortion c, then there is a tour in M of length at most $c + 1$.*

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 c, it follows that $|f(u_1) - f(u_{2n})| \le 2c$. Moreover, from noncontraction 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 $\mathbb{R}$ is still noncontracting (with respect to the metric induced by H). Therefore, there exists a mapping f' of $V \cup V'$ into $\mathbb{R}$ which is noncontracting, 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 $2c$. W.l.o.g., assume that the diameter δ of $f'(V)$ is not greater than the diameter of $f'(V')$. This implies that $\delta \le (2c - 2)/2 = c - 1$. Therefore, the ordering of the vertices in V induced by f' corresponds to a tour in M of length at most $\delta + 2 \le c + 1$. □

COROLLARY 3. *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. In this section we consider the problem of embedding an n-point subset of the unit sphere $\mathbb{S}^2$ equipped with the geodetic distance, into the (Euclidean) plane $\mathbb{R}^2$. Given such $M = (X, D)$, we efficiently embed it into the plane with distortion optimal up to a constant factor. We also provide a lower bound on the distortion in the worst case. While the forthcoming discussion is restricted to $\dim = 2$, it is not hard to extend both the arguments and the conclusions to any dimension d.

Let $M = (X, D)$ be a metric space, where $X \subset \mathbb{S}^2$, and D is induced by $d_{\mathbb{S}^2}$, the geodesic distance of $\mathbb{S}^2$. Let B the smallest closed spherical cup containing X. W.l.o.g., we may assume that the center of B is the south pole.

Consider the following mapping φ from $\mathbb{S}^2$ with a punctured north pole to $\mathbb{R}^2$. First, we parameterize the unit sphere in a manner similar to polar coordinates in the plane; i.e., a point $p \in \mathbb{S}^2$ will be represented in the form $p = (\rho, \theta)$, where $\rho \in [0, \pi]$ is the geodesic distance in $\mathbb{S}^2$ (i.e., the angle in radians) between the south pole and p, and $\theta \in [0, 2\pi)$ is the angle between the *reference plane* in $\mathbb{R}^3$, and the plane defined by the north pole, south pole, and p. The reference plane, containing the north and the south poles, is arbitrary. Using this parameterization, the mapping $\varphi : \mathbb{S}^2 \setminus north\ pole \to \mathbb{R}^2$ is simply

$$\phi(\rho, \theta) = (\rho, \theta),$$

where the second pair of coordinates (ρ, θ) are the standard polar coordinates in the

plane. Thus, φ maps the punctured $\mathbb{S}^2$ to an open ball or radius π with a center at the origin.

The main technical result of this section is that φ maps X into $\mathbb{R}^2$ with distortion optimal up to a constant multiplicative factor.

Clearly, φ is invertible, differentiable, and nowhere singular. Moreover, it isometrically maps the spherical meridians into (straight-line) rays with apex at the origin. The angle between meridians (as measured at the south pole) is also preserved under φ, i.e., it is equal to the angle between the corresponding rays.

CLAIM 1. *The local expansion of φ at the point $p = (\rho, \theta) \in \mathbb{S}^2$ is between 1 and $\rho/\sin(\rho)$. More concretely, the spherical ball $B(p, \epsilon)$ with center at p and an infinitesimal radius ϵ, is mapped by φ to the infinitesimal ellipse centered at $\varphi(p) = (\rho, \theta)$, with one main axis (corresponding to moving in the direction of the meridian) being of length 2ϵ, and the second main axis (corresponding to moving in the direction of the parallel) being of length $2\rho/\sin(\rho) \cdot \epsilon$.*

Proof. The fact that the infinitesimal ϵ-ball is mapped to an infinitesimal ellipse follows from the locally Euclidean structure of the sphere, and the differentiability, i.e., local linearity, of φ. The maximum and the minimum local expansion of φ are obtained by moving from p in the directions corresponding to the two main axes of this ellipse, respectively.

The directions of the axes can be deduced via a symmetry argument. Consider the large circle $C \subset \mathbb{S}^2$ containing the two poles and p. The image of C under φ is an open interval I, lying on the line ℓ via the origin and $\varphi(p)$. Let $h_C : \mathbb{S}^2 \to \mathbb{S}^2$ be the self-reflection of the sphere with respect to C, and let h_ℓ be the self-reflection of the plane with respect to ℓ. Then, $\varphi(h_C(q)) = h_\ell(\varphi(q))$ for any q in the punctured sphere. In addition, h_C and h_ℓ obviously are isometries. Consequently, since $\varphi(B(p, \epsilon))$ is reflection-symmetrical with respect to C, the ellipse $\varphi(B(p, \epsilon))$ is reflection-symmetrical with respect to ℓ, and thus one of its main axes lies on ℓ, while the other is necessarily orthogonal to ℓ. Moreover, the former axis corresponds to moving from p along the meridian, while the latter axis—by the symmetry argument— corresponds to moving along the parallel (i.e., orthogonally to the meridian).

The local expansion of φ at p obtained by moving on in the direction of the meridian is 1, since φ preserves the metric on the meridians. In the direction of the parallel, the local expansion of φ can be found by computing the ratio between the length of the ρ-parallel in $\mathbb{S}^2$, and its image in the plane. The latter is obviously $2\pi\rho$. The former is a planar circle, whose diameter is equal to the length of the chord bounding a cap of angle 2ρ (equivalently, $2\pi - 2\rho$) in the unit sphere $\mathbb{S}^1$. By basic trigonometry, the length of this chord is $2\sin(\rho)$, and therefore the length of the ρ-parallel is $2\pi\sin(\rho)$. Thus, the expansion of φ in the direction of the ρ-parallel is $\rho/\sin(\rho)$.   □

CLAIM 2. *The expansion of φ on the ρ-parallel of $\mathbb{S}^2$ equipped with the induced geodesic metric of $\mathbb{S}^2$, is at least 1 and at most $\rho/\sin(\rho)$. The upper bound is tight, and it is attained (approached) on infinitesimally closed points.*

Proof. Let B_ρ be the planar disc, being the section of the 3-dimensional unit ball by the plane containing the ρ-parallel of $\mathbb{S}^2$. Let also p, q be two points on this parallel, with $d_{\mathbb{S}^2}(p, q) = \theta \leq \pi$.

Let c be the center of B_ρ. A basic geometric observation is that the angle between the vectors $\vec{cp}, \vec{cq}$ is equal to the angle between the meridians via p and q, respectively, as measured at the south pole. Call this angle α. Keeping in mind the (already used)

fact that the length of the chord of an $\mathbb{S}^1$ cap of length (= angle) γ is $2\sin(\gamma/2)$, we obtain

$$
\begin{aligned}
\frac{\|\varphi(p) - \varphi(q)\|_2}{\theta} &= \frac{\|\varphi(p) - \varphi(q)\|_2}{\|p - q\|_2} \cdot \frac{\|p - q\|_2}{\theta} \\
&= \frac{\rho}{\sin(\rho)} \cdot \frac{2\sin(\theta/2)}{\theta} = \frac{\rho}{\sin(\rho)} \cdot \frac{\sin(\theta/2)}{\theta/2}.
\end{aligned}
$$

In the second inequality, the first term follows from the similarity of the triangles (c, p, q) and $(0, \varphi(p), \varphi(q))$. The second terms follow from considering the big circle containing the center of $\mathbb{S}^2$, p, and q. In this circle, $[p, q]$ is a chord of a cap of angle θ.

The maximum expansion is attained when θ tend to 0, and is equal to $\frac{\rho}{\sin(\rho)}$. The minimum expansion is attained at the antipode points of the ρ-parallel, i.e., when $\theta = 2\pi - 2\rho$. Assuming w.l.o.g., that $\rho \geq \pi/2$, it is equal to $\frac{\rho}{\pi - \rho} \geq 1$.   □

Returning to our metric space $M = (X, D)$, where $X \subset \mathbb{S}^2$, and D is induced by the geodesic distance of $\mathbb{S}^2$, we can now upper bound the metric distortion of the embedding $\varphi : X \hookrightarrow \mathbb{R}^2$. Recall that B, the smallest closed spherical cup containing X, is by our assumption centered at the south pole of the sphere. Let ρ be the geodesic radius of B.

LEMMA 19 (upper bound).   *The embedding $\varphi : X \hookrightarrow \mathbb{R}^2$ is noncontracting, and its expansion is bounded by $\rho/\sin(\rho)$. Thus, $\mathrm{dist}(\varphi) \leq \rho/\sin(\rho)$.*

Proof. It suffices to establish the lemma for $X = B$.

To show that φ is noncontracting on B is equivalent to showing that φ^{-1} is not expanding on $\varphi(B)$, i.e., the radius-ρ closed disc in the plane, with center at the origin. Observe that φ^{-1} is well defined and differentiable there, and therefore its local expansion, by reversed Claim 1, is between $\sin(\rho)/\rho$ and 1. Consequently, for any two points $q_1, q_2 \in \varphi(B)$, $\varphi^{-1}([q_1, q_2])$ is a path in $\mathbb{S}^2$ of length at most the length of $[q_1, q_2]$ times the maximum local expansion of φ^{-1}, i.e., 1. Thus, $d_{\mathbb{S}^2}(\varphi^{-1}(q_1), \varphi^{-1}(q_2) \leq \|q_1 - q_2\|_2$, as claimed.

To upper bound the expansion of φ on B, consider any two points $p_1, p_2 \in B$, and the spheric geodesic path γ between them. If $\gamma \subset B$, its image under φ, by Claim 1 is no longer than $|\gamma| \cdot \rho/\sin(\rho)$, where $|\gamma|$ denotes the length of γ. Otherwise, γ is composed of three parts $\gamma_1, \gamma_2, \gamma_3$, where γ_1, γ_3 (possibly degenerate) belong to B, while γ_2 is a geodetic path between two points r_1, r_2 on the ρ-parallel with interior disjoint from B. By Claim 1, the lengths of $\varphi(\gamma_1)$ and $\varphi(\gamma_1)$ are at most $|\gamma_1| \cdot \rho/\sin(\rho)$ and $|\gamma_3| \cdot \rho/\sin(\rho)$, respectively. By Claim 2,

$$
\|\varphi(r_1) - \varphi(r_2)\|_2 \leq d_{\mathbb{S}^2}(r_1, r_2) \cdot \frac{\rho}{\sin(\rho)} = |\gamma_2| \cdot \frac{\rho}{\sin(\rho)}.
$$

Consequently,

$$
\|\varphi(p_1) - \varphi(p_2)\|_2 \leq (|\gamma_1| + |\gamma_2| + |\gamma_3|) \cdot \frac{\rho}{\sin(\rho)} = d_{\mathbb{S}^2}(p_1, p_2) \cdot \frac{\rho}{\sin(\rho)} ,
$$

and the conclusion follows.   □

Next, we want to show that for *any* embedding $\phi : X \hookrightarrow \mathbb{R}^2$, its distortion $\mathrm{dist}(\phi)$ cannot be much less than $\rho/\sin(\rho)$. Instead of B, the smallest closed spherical cap containing X, it will be convenient now to work with its complement K, the largest open spherical cap disjoint from X. Clearly, the (spheric) radius of K is $\kappa = \pi - \rho$,

where ρ is the radius of B. Observe that the radius of K is κ if and only if X is a κ-net (where κ is tight) in the sphere, i.e., any $p \in \mathbb{S}^2$ is at (geodesic) distance $\leq \kappa$ from X.

LEMMA 20 (lower bound). *Assume that X forms a κ-net in $\mathbb{S}^2$. For any embedding $\phi : X \hookrightarrow \mathbb{R}^2$, it holds that $\mathrm{dist}(\phi) \geq \max\{1, \frac{\pi - 2\kappa}{\pi\kappa}\}$.*

Proof. The proof consists of two main ingredients: protucing a Lipschitz extension of ϕ to the entire sphere, then using the Borsuk–Ulam theorem. We start with the former. We refer the reader to [Mat03] for further information on the Borsuk–Ulam theorem, and to [LN04] for details and definitions regarding Lipschitz extensions.

We would like to have an extension $\widetilde{\phi}$ of ϕ from X to the entire $\mathbb{S}^2$, whose Lipschitz constant is the same as before, i.e.,

$$\max_{\mathbb{S}^2} \frac{\|\widetilde{\phi}(p_1) - \widetilde{\phi}(p_1)\|_2}{d_{\mathbb{S}^2}(p_1, p_2)} = \max_X \frac{\|\phi(p_1) - \phi(p_1)\|_2}{d_{\mathbb{S}^2}(p_1, p_2)}.$$

Unfortunately, such an expansion does not seem to exist for all $X \subset \mathbb{S}^2$. However, if we treat ϕ as an embedding of $\mathbb{S}^2$ equipped with the induced 3-dimensional *Euclidean* rather than geodesic distance, there indeed exists a Lipschitz extension $\widetilde{\phi} : \mathbb{S}^2 \hookrightarrow \mathbb{R}^2$ of ϕ on X, such that

$$\max_{\mathbb{S}^2} \frac{\|\widetilde{\phi}(p_1) - \widetilde{\phi}(p_1)\|_2}{\|p_1 - p_2\|_2} = \max_X \frac{\|\phi(p_1) - \phi(p_1)\|_2}{\|p_1 - p_2\|_2}.$$

This is a special case of Kirszbraun's theorem ([LS97, Kir34]; see also [LN04]). Since the ratio between the Euclidean and the geodesic metrics on $\mathbb{S}^2$ ranges in the interval $[1, \pi/2]$, we conclude that the Lipschitz constant of the same $\widetilde{\phi}$ with respect to the geodesic distance of $\mathbb{S}^2$ is at most $\pi/2$ times that of ϕ:

$$\max_{\mathbb{S}^2} \frac{\|\widetilde{\phi}(p_1) - \widetilde{\phi}(p_1)\|_2}{d_{\mathbb{S}^2}(p_1, p_2)} \leq \max_{\mathbb{S}^2} \frac{\|\widetilde{\phi}(p_1) - \widetilde{\phi}(p_1)\|_2}{\|p_1 - p_2\|_2} = \max_X \frac{\|\phi(p_1) - \phi(p_1)\|_2}{\|p_1 - p_2\|_2}$$
$$\leq \frac{\pi}{2} \cdot \max_X \frac{\|\phi(p_1) - \phi(p_1)\|_2}{d_{\mathbb{S}^2}(p_1, p_2)}.$$

Due to scalability of $\mathbb{R}^2$, we may, w.l.o.g., assume in what follows that the Lipschitz constant of ϕ is precisely 1; in particular, it is nonexpanding. Then, there exists an extension $\widetilde{\phi}$ of ϕ which is at most $\pi/2$-expanding (with respect to $d_{\mathbb{S}^2}$). In particular, $\widetilde{\phi}$ is continuous.

Next, we apply the Borsuk–Ulam theorem [Bor33] to $\widetilde{\phi}$ to conclude that there exist antipodals $z, z' \in \mathbb{S}^2$, such that $\widetilde{\phi}(z) = \widetilde{\phi}(z')$. Since X is a κ-net in $\mathbb{S}^2$, there exist points $p, p' \in X$, such that $d_{\mathbb{S}^2}(p, z) \leq \kappa$ and $d_{\mathbb{S}^2}(p', z') \leq \kappa$. Since $\widetilde{\phi}$ is at most $\pi/2$-expanding, it follows that $\|\phi(p) - \phi(p')\|_2 \leq \pi/2 \cdot 2\kappa = \pi\kappa$. On the other hand, we have

$$d_{\mathbb{S}^2}(p, p') \geq d_{\mathbb{S}^2}(z, z') - d_{\mathbb{S}^2}(z, p) - d_{\mathbb{S}^2}(z', p') \geq \pi - 2\kappa.$$

Thus, ϕ has expansion 1, and has contraction $\geq \frac{\pi - 2\kappa}{\pi\kappa}$. Hence, the distortion of ϕ is at least

$$\mathrm{dist}(\phi) \geq \max\left\{1, \frac{\pi - 2\kappa}{\pi\kappa}\right\}.$$

$\square$

THEOREM 5. *There exists a polynomial-time 4.411-approximation algorithm for the problem of embedding a metric $M = (X, D)$, where $X \subset \mathbb{S}^2$, and D is induced by the geodesic distance of the sphere, into the Euclidean plane.*

Proof. Combining the results of Lemmas 19 and 20, one concludes that the embedding of Lemma 19 is optimal up to the factor of $(\pi - \kappa)/\sin(\kappa)$ divided by $\max\{1, \frac{\pi - 2\kappa}{\pi \kappa}\}$, where κ, as always, is the spherical radius of the largest open K disjoint with X. This ratio, as a function of κ, is maximized when the second expression becomes 1, i.e., at $\kappa = \frac{\pi}{\pi + 2}$. For this κ, the value of the ratio is slightly under 4.411.

In order to implement the construction of Lemma 19 in polynomial time, one needs to find the maximum K. This can be done in a trivial manner in time $O(n^3)$ observing that the maximal K must have at least 3 points of X on its boundary, and check all caps defined by different triples of X. The running time can be improved to $O(n \log n)$ using the algorithm from [MJSG99] which is based on the observation that the maximum cap is defined by the hyperplane supporting one of the facets of the convex hull of the points in $\mathbb{R}^3$. Alternatively, one can obtain a linear-time algorithm for computing the largest empty cap by observing that this is an LP-type problem with constant combinatorial dimension; we refer the reader to [HP11, section 15.5] for further details on LP-type problems. □

Remark 1. The constant 4.411 in Theorem 5 is not optimal. The lower bound of Lemma 20 could be somewhat strengthened if instead of using the Euclidean approximation for the sphere and employing the Kirschbaum theorem, one used an ℓ_∞^2 approximation of the Euclidean plane, and employed the existence of a Lipschitz-preserving extension for ℓ_∞ host spaces. This would yield another extension $\widetilde{\phi}$ of ϕ, with $\|\widetilde{\phi}\|_{\mathrm{Lip}} \leq \sqrt{2} \cdot \|\phi\|_{\mathrm{Lip}}$. Using this extension, the constant in Theorem 5 becomes 4.113. However, for higher dimensions the first extension is much superior to the second, which was the reason for our choice. Finding the $\widetilde{\phi}$ of the minimum possible $\|\widetilde{\phi}\|_{\mathrm{Lip}}$ is an interesting open problem.

Remark 2. Since the geodesic metric on the unit sphere is $\pi/2$-close to the Euclidean metric inherited from $\mathbb{R}^3$, Theorem 5 implies a similar result for $(X, \| * \|_2)$. Here, we do not attempt to obtain a good constant for this setting, since getting an analogue of Lemma 19 apparently requires a messy numerical optimization.

As a corollary of Theorem 5, we obtain the following structural result.

THEOREM 6. *The metric space $M = (X, D)$ as above, with $|X| = n$, can always be embedded into the Euclidean plane with distortion $O(\sqrt{n})$. Conversely, there exist X's requiring distortion $\Omega(\sqrt{n})$.*

Proof. For the first part, in view of Lemma 20, it suffices to show that there exist K as above of sufficiently large radius κ. Indeed, assume that no such K exists. Then, the union of radius-κ spherical caps centered at the points of X cover the entire $\mathbb{S}^2$. Since by elementary geometry, the surface area of such a cap is 2π times its height, i.e., $2\pi(1 - \cos(\kappa)) = 4\pi \sin^2(\kappa/2)$, and the area of the sphere is 4π, it must hold that

$$4\pi \sin^2(\kappa/2) \cdot n \geq 4\pi \quad \Longrightarrow \quad \kappa \geq 2\arcsin(n^{-1/2}) = 2n^{-1/2} + O(n^{-3/2}).$$

The tightness is implied by the well-known existence of ϵ-nets on $\mathbb{S}^2$ of size $O(\epsilon^{-2})$. Such a net can be produced, e.g., greedily, by adding each time, as long as possible, a new point, ϵ-far from the previously added points. Clearly this procedure results in an ϵ-net. Let n be its size. The spherical caps of radius $\epsilon/2$ constitute a disjoint

packing, hence, by a volume argument,

$$4\pi \sin^2(\epsilon/2) \cdot n \le 4\pi \quad \implies \quad n \le \sin^{-2}(\epsilon/2) = 4\epsilon^{-2} + O(1). \qquad \square$$

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)*, SIAM J. Comput., 28 (1999), pp. 1073–1085.

[BCIS05]   M. Badoiu, J. Chuzhoy, P. Indyk, and A. Sidiropoulos, *Low-distortion embeddings of general metrics into the line*, in Proceedings of the 37th Annual ACM Symposium on Theory of Computing, Baltimore, MD, ACM, New York, 2005, pp. 225–233.

[BCIS06]   M. Badoiu, J. Chuzhoy, P. Indyk, and A. Sidiropoulos, *Embedding ultrametrics into low-dimensional spaces*, in Proceedings of the 22nd ACM Symposium on Computational Geometry, Sedona, AZ, ACM, New York, 2006, pp. 187–196.

[BIR03]   M. Badoiu, P. Indyk, and Y. Rabinovich, *Approximate Alorithms for Embedding Metrics into Low-Dimensional Spaces*, manuscript.

[BIS07]   M. Badoiu, P. Indyk, and A. Sidiropoulos, *Approximation algorithms for embedding general metrics into trees*, in Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2007, New Orleans, LA, SIAM, Philadelphia, 2007, pp. 512–521.

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

[Bou85]   J. Bourgain, *On Lipschitz embedding of finite metric spaces into Hilbert space*, Israel J. Math., 52 (1985), pp. 46–52.

[Bǎ03]   M. Bǎdoiu, *Approximation algorithm for embedding metrics into a two-dimensional space*, in 14th Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, 2003, pp. 434–443.

[CK11]   A. Chakrabarti and S. Khot, *Combinatorial theorems about embedding trees on the real line*, J. Graph Theory, 67 (2011), pp. 153–168.

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

[ESZ10]   J. Edmonds, A. Sidiropoulos, and A. Zouzias, *Inapproximability for planar embedding problems*, in Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, 2010, pp. 222–235.

[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, ACM, New York, 1993.

[Fei00]   U. Feige, *Approximating the bandwidth via volume respecting embeddings*, J. Comput. System Sci., 60 (2000), pp. 510–539.

[FFL+13]   M. R. Fellows, F. V. Fomin, D. Lokshtanov, E. Losievskaja, F. A. Rosamond, and S. Saurabh, *Distortion is fixed parameter tractable*, ACM Trans. Comput. Theory, 5 (2013), 16.

[FLS09]   F. V. Fomin, D. Lokshtanov, and S. Saurabh, *An exact algorithm for minimum distortion embedding*, in Graph-Theoretic Concepts in Computer Science, Springer, 2009, pp. 112–121.

[Gup00]   A. Gupta, *Improved bandwidth approximation for trees*, in Proceedings of the Eleventh Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, 2000, pp. 788–793.

[HIL98]   J. Hastad, L. Ivansson, and J. Lagergren, *Fitting points on the real line and its application to rh mapping*, in Algorithms–ESA '98, Lecture Notes in Comput. Sci. 1461, Springer, Berlin, 1998, pp. 465–467.

[HP05]   A. Hall and C. H. Papadimitriou, *Approximating the distortion*, in Approximation, Randomization and Combinatorial Optimization, Algorithms and Techniques, Proceedings of the 8th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, APPROX 2005 and 9th International Workshop on Randomization and Computation, RANDOM 2005, Berkeley, CA, Springer, Berlin, 2005, pp. 111–122.

[HP11] S. HAR-PELED, *Geometric Approximation Algorithms*, Math. Surveys Monogr. 173, American Mathematical Society, Providence, RI, 2011.

[Ind01] P. INDYK, *Tutorial: Algorithmic applications of low-distortion geometric embeddings*, in Proceedings of the Symposium on Foundations of Computer Science, IEEE Computing Society, Los Alamitos, CA, 2001, pp. 10–33.

[Iva00] L. IVANSSON, *Computational Aspects of Radiation Hybrid*, Doctoral Dissertation, Department of Numerical Analysis and Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden, 2000.

[Kir34] M. D. KIRSZBRAUN, *Über die zusammenziehenden und Lipschitzschen Transformationen*, Fund. Math., 22 (1934), pp. 77–108.

[KRS09] C. KENYON, Y. RABANI, AND A. SINCLAIR, *Low distortion maps between point sets*, SIAM J. Comput., 39 (2010), pp. 1617–1636.

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

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

[KS07] S. KHOT AND R. SAKET, *Hardness of embedding metric spaces of equal size*, in Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, Proceedings of the 10th International Workshop, APPROX 2007, and 11th International Workshop, RANDOM 2007, Princeton, NJ, Springer, Berlin, 2007, pp. 218–227.

[LLR94] N. LINIAL, E. LONDON, AND Y. RABINOVICH, *The geometry of graphs and some of its algorithmic applications*, in Proceedings of 35th Annual IEEE Symposium on Foundations of Computer Science, IEEE Computer Society, Los Alamitos, CA, 1994, pp. 577–591.

[LMS18] D. LOKSHTANOV, D. MARX, AND S. SAURABH, *Slightly superexponential parameterized problems*, SIAM J. Comput., 47 (2018), pp. 675–702.

[LN04] J. R. LEE AND A. NAOR, *Absolute Lipschitz extendability*, C. R. Acad. Sci. Ser. I Math., 338 (2004), pp. 859–862.

[LS97] U. LANG AND V. SCHROEDER, *Kirszbraun's theorem and metric spaces of bounded curvature*, Geom. Funct. Anal., 7 (1997), pp. 535–560.

[Mat90] J. MATOUŠEK, *Bi-Lipschitz embeddings into low-dimensional Euclidean spaces*, Comment. Math. Univ. Carolin., 31 (1990), pp. 589–600.

[Mat96] J. MATOUŠEK, *On the distortion required for embedding finite metric spaces into normed spaces*, Israel J. Math., 93 (1996), pp. 333–344.

[Mat03] J. MATOUŠEK, *Using the Borsuk–Ulam Theorem: Lectures on Topological Methods in Combinatorics and Geometry*, Springer, New York, 2003.

[MJSG99] J. MAJHI, R. JANARDAN, M. SMID, AND P. GUPTA, *On some geometric optimization problems in layered manufacturing*, Comput. Geom., 12 (1999), pp. 219–239.

[MS10] J. MATOUŠEK AND A. SIDIROPOULOS, *Inapproximability for metric embeddings into* $\mathbb{R}^{\{\}}$, Trans. Amer. Math. Soc., 362 (2010), pp. 6341–6365.

[NR15] A. NAYYERI AND B. RAICHEL, *Reality distortion: Exact and approximate algorithms for embedding into the line*, in IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS 2015, Berkeley, CA, IEEE, Piscataway, NJ, 2015, pp. 729–747.

[OS08] K. ONAK AND A. SIDIROPOULOS, *Circular partitions with applications to visualization and embeddings*, in Proceedings of the 24th ACM Symposium on Computational Geometry, College Park, MD, ACM, New York, 2008, pp. 28–37.

[PS05] C. H. PAPADIMITRIOU AND S. SAFRA, *The complexity of low-distortion embeddings between point sets*, in Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2005, Vancouver, Canada, SIAM, Philadelphia, 2005, pp. 112–118.

[Sax80] J. B. SAXE, *Dynamic-programming algorithms for recognizing small-bandwidth graphs in polynomial time*, SIAM J. Algebr. Discrete Methods, 1 (1980), pp. 363–369.

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

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

[Wor] Working Group on Algorithms for Multidimensional Scaling. *Workshop on Algorithms for Multidimensional Scaling*, http://dimacs.rutgers.edu/Workshops/Algorithms/.