Ordinal Embeddings of Minimum Relaxation: General Properties, Trees, and Ultrametrics

NOGA ALON
Tel Aviv University
MIHAI BĂDOIU
Google Inc.
ERIK D. DEMAINE
MIT
MARTIN FARACH-COLTON
Rutgers University
MOHAMMADTAGHI HAJIAGHAYI
AT&T Labs — Research
and
ANASTASIOS SIDIROPOULOS
MIT

Abstract. We introduce a new notion of embedding, called *minimum-relaxation ordinal embedding*, parallel to the standard notion of minimum-distortion (metric) embedding. In an ordinal embedding, it is the relative order between pairs of distances, and not the distances themselves, that must be preserved as much as possible. The (multiplicative) relaxation of an ordinal embedding is the maximum ratio between two distances whose relative order is inverted by the embedding. We develop several worst-case bounds and approximation algorithms on ordinal embedding. In particular, we establish that ordinal embedding has many qualitative differences from metric embedding, and we capture the ordinal behavior of ultrametrics and shortest-path metrics of unweighted trees.

Categories and Subject Descriptors: F.2.0 [**Analysis of Algorithms and Problem Complexity**]: General

---

A preliminary version of this article appeared in *Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms*, January 2005, pages 650–659. This work was done while M. Bădoiu and M. Hajiaghayi were at MIT. N. Alon was supported in part by a grant from the Israel Science Foundation, and by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University. E. Demaine and M. Hajiaghayi were supported in part by NSF under grant number ITR ANI-0205445. A. Sidiropoulos was supported in part by the Paris Kanellakis Fellowship Fund and by the Alexandros S. Onassis Public Benefit Foundation.
Author's addresses: N. Alon, Schools of Mathematics and Computer Science, Tel Aviv University, Tel Aviv, Israel, e-mail: nogaa@tau.ac.il; M. Bădiou, Google Inc., New York, NY, USA, e-mail: mihai@mit.edu; E. Demaine, A. Sidiropoulos, Computer Science and Artificial Intelligence Laboratory, MIT, 32 Vassar St., Cambridge, MA 02139, USA, e-mail: {edemaine,tasos}@mit.edu; M. Farach-Colton, Department of Computer Science, Rutgers University, Piscataway, NJ 08855, USA, e-mail: farach@cs.rutgers.edu; M. Hajiaghayi, AT&T Labs — Research, 180 Park Ave., Florham Park, NJ 07932, USA, e-mail: hajiagha@research.att.com.

General Terms: Algorithms, Theory

Additional Key Words and Phrases: metrics, ordinal embedding, relaxation, distortion

1 *Introduction*

The classical field of multidimensional scaling (MDS) has witnessed a surge of interest in recent years with a slew of papers on metric embeddings; see e.g. Indyk and Matoušek [2004]. The problem of multidimensional scaling is that of mapping points with some measured pairwise distances into some target metric space. Originally, the MDS community considered embeddings into an ℓ_p space with the goal of aiding in visualization, compression, clustering, or nearest-neighbor searching; thus, low-dimensional embeddings were sought. An *isometric embedding* preserves all distances, while more generally, *metric embeddings* trade off the dimension with the fidelity of the embeddings.

Note, however, that the distances themselves are not essential in nearest-neighbor searching and many contexts of visualization, compression, and clustering. Rather, the order of the distances captures sufficient information, that is, we might only need an embedding into a metric space with any monotone mapping of the distances. Such embeddings were heavily studied in the early MDS literature [Cunningham and Shepard 1974; Kruskal 1964a; Kruskal 1964b; Shepard 1962; Torgerson 1952] and have been referred to as *ordinal embeddings*, *nonmetric MDS*, or *monotone maps*. Here, we use the first term.

While the early work on ordinal embeddings was largely heuristic, there has been some work with mathematical guarantees since then. Define a *distance matrix* to be any matrix of pairwise distances, not necessarily describing a metric. Shah and Farach-Colton [2004] have shown that it is NP-hard to decide whether a distance matrix can be ordinally embedded into an additive metric, that is, the shortest-path metric in a tree. Define the *ordinal dimension* of a distance matrix to be the smallest dimension of a Euclidean space into which the matrix can be ordinally embedded. Bilu and Linial [2004] have shown that every matrix has ordinal dimension at most $n-1$. They also applied the methods of Alon et al. [1985] to show that (in a certain well-defined sense) almost every n-point metric space has ordinal dimension $\Omega(n)$. Because ultrametrics can be characterized by the order of distances on all triangles, they are closed under monotone mappings. Holman [1972] showed that every n-point ultrametric can be isometrically embedded into $(n-1)$-dimensional Euclidean space and that $n-1$ dimensions are necessary. Combined with the closure property just noted, this shows that the ordinal dimension of every ultrametric is exactly the maximal $n-1$.[1]

Relaxations of ordinal embeddings have involved problems of deciding the realization of partial orders. For example, Opatrny [1979] showed that it is NP-hard to decide whether there is an embedding into one dimension satisfying a partial order that specifies the maximum edge for some triangles. Such partial orders on triangles are called *betweenness constraints*. Chor and Sudan [1998] gave a 1/2-

[1] This observation settles an open problem posed by Bilu and Linial [2004] asking for the worst-case ordinal dimension of any metric on n points, which they showed was between $n/2$ and $n-1$. Ultrametrics show that the answer is $n-1$.

approximation for maximizing the number of satisfied constraints. It is also NP-hard to decide whether there is an embedding into an additive metric that satisfies a partial order defined by the total order of each triangle [Shah and Farach-Colton 2004].

1.1 Our Results. We take a different approach. We define a metric M' to be an *ordinal embedding with relaxation* $\alpha \geq 1$ of a distance matrix M if $\alpha M[i,j] < M[k,l]$ implies $M'[i,j] < M'[k,l]$. In other words, significantly different distances have their relative order preserved. Note that in an ordinary ordinal embedding, we must respect distance equality, while in an ordinal embedding with relaxation 1, we may break ties. It is now natural to minimize the relaxation needed to embed a distance matrix M into a target family of metric spaces. Here we optimize the confidence with which we make an ordinal assertion, rather than the number of ordinal constraints satisfied.

In this article, we prove a variety of results about the Ordinal Relaxation Problem. We show that the best relaxation achievable is always at most the best distortion of a metric embedding. Furthermore, while the optimal relaxation is bounded by the ratio between the largest and smallest distances in M, the optimal distortion can grow arbitrarily. Indeed, the ratio between the optimal relaxation and distortion can be arbitrarily large even when embedding into the line, and it can be infinite when embedding into cut metrics. (We also give a polynomial-time algorithm to compute the best ordinal embedding into a cut metric.) We show that, if the target class of the embedding is ultrametrics, the relaxation and distortion are equal, and the optimal embedding can be computed in polynomial time. More surprisingly, we show that ultrametrics are the only target metrics for which all distance matrices have a bounded ratio between the best distortion and the best relaxation.

We demonstrate many more differences between ordinal embedding relaxation and metric embedding distortion. While any metric can be isometrically embedded into ℓ_∞, there are four-point metrics that cannot be so embedded into any ℓ_p, $p < \infty$. In contrast, we show that it is possible to ordinally embed any distance matrix into ℓ_p for any fixed $1 \leq p \leq \infty$. We show that the shortest-path metric of an unweighted tree can be ordinally embedded into d-dimensional Euclidean space with relaxation $\tilde{O}(n^{1/d})$. We also show that relaxation $\Omega(n^{1/(d+1)})$ is sometimes necessary. In contrast, the best bounds on the worst-case distortion required are $O(n^{1/(d-1)})$ and $\Omega(n^{1/d})$ [Gupta 2000]. The proof techniques required for the ordinal case are also substantially different (in particular because the usual "packing" arguments fail) and lead to approximation algorithms described in the following. We show that ultrametrics can be ordinally embedded into the real line with relaxation 1. In contrast, simple ultrametrics (e.g., the uniform metric) require linear distortion in the worst case for any metric embedding into the line. Moreover, the best known metric embedding of ultrametrics into even $c \lg n$-dimensional Euclidean space has distortion $1 + \Omega(1/\sqrt{c})$ [Bartal and Mendel 2004], and ordinary (no-relaxation) ordinal embeddings require $n-1$ dimensions. For general metrics, we show a lower bound of $\Omega(\lg n/(\lg d + \lg\lg n))$ on the relaxation of any ordinal embedding into d-dimensional ℓ_p space for fixed integers p or $p = \infty$. In particular, for $d = \Theta(\lg n)$, this lower bound is $\Omega(\lg n/\lg\lg n)$, leaving a gap between the upper bound of $O(\lg n)$ which follows from Bourgain embedding. In contrast, for metric embeddings, there is an $\Omega(\lg n)$ lower bound on distortion for $d = \Theta(\lg n)$ [Linial et al. 1995; Matoušek 1997].

We also develop approximation algorithms for finding the minimum possible relaxation for an ordinal embedding of a specified metric. Specifically, we give a 3-approximation for ordinal embedding of the shortest-path metric of a specified unweighted tree into the line. In contrast, only $O(n^{1/3})$-approximation algorithms are known for the same problem with distortion [Bădoiu et al. 2005]. In general, approximation algorithms for embedding are a central challenge in the field and few are known [Håstad et al. 2003; Ivansson 2000; Bădoiu 2003; Agarwala et al. 1999; Farach and Kannan 1999; Bădoiu et al. 2004]. We also expect that our techniques will extend to obtain approximation algorithms for more general ordinal embedding problems.

2 *Definitions*

In this section, we define ordinal embeddings and relaxation, as well as the standard notions of metric embeddings and distortion.

Consider a finite metric $D : P \times P \to [0, \infty)$ on a finite point set P—the *source metric*—and a class $\mathcal{T}$ of metric spaces $(T, d) \in \mathcal{T}$, where d is the distance function for space T—the *target metrics*. An *ordinal embedding (with no relaxation)* of D into $\mathcal{T}$ is a choice $(T, d) \in \mathcal{T}$ of a target metric and a mapping $\phi : P \to T$ of the points into the target metric such that every comparison between pairs of distances has the same outcome: for all $p, q, r, s \in P$, $D(p, q) \leq D(r, s)$ if and only if $d(\phi(p), \phi(q)) \leq d(\phi(r), \phi(s))$. Equivalently, ϕ induces a monotone function $D(p, q) \mapsto d(\phi(p), \phi(q))$, and, for this reason, ordinal embeddings are also called *monotone embeddings*. An *ordinal embedding with relaxation* α of D into $\mathcal{T}$ is a choice $(T, d) \in \mathcal{T}$ and a mapping $\phi : P \to T$ such that every comparison between pairs of distances not within a factor of α has the same outcome: for all $p, q, r, s \in P$ with $D(p, q) > \alpha\, D(r, s)$, $d(\phi(p), \phi(q)) > d(\phi(r), \phi(s))$. Equivalently, we can view a relaxation α as defining a partial order on distances $D(p, q)$, where two distances $D(p, q)$ and $D(r, s)$ are comparable if and only if they are not within a factor of α of each other, and the ordinal embedding must preserve this partial order on distances.

An ordinal embedding with relaxation 1 is a different notion from ordinal embedding with no relaxation, because the former allows violation of equalities between pairs of distances. Indeed, we will show in Section 6.1 that the two notions have major qualitative differences. We define ordinal embedding with relaxation in this way, instead of making the $> \alpha$ inequality nonstrict, because otherwise our notion of relaxation 1 would have to be phrased as relaxation $1 + \varepsilon$ for any $\varepsilon > 0$. Another consequence is that we can define the minimum possible relaxation $\alpha^* = \alpha^*(D, \mathcal{T})$ of an ordinal embedding of D into $\mathcal{T}$, instead of having to take an infimum. (The infimum will be realized provided the space $\mathcal{T}$ is closed.)

We pay particular attention to contrasts between relaxation in ordinal embedding relaxation and distortion in standard embedding, which we call *metric embedding* for distinction. A *contractive metric embedding with distortion* c of a source metric D into a class $\mathcal{T}$ of target metrics is a choice $(T, d) \in \mathcal{T}$ and a mapping $\phi : P \to T$

such that no distance increases and every distance is preserved up to a factor of c: for all $p, q \in P$, $1 \le D(p,q)/d(\phi(p), \phi(q)) \le c$. Similarly, we can define an *expansive metric embedding with distortion c* with the inequality $1 \le d(\phi(p), \phi(q))/D(p,q) \le c$. When $c = 1$, these two notions coincide to require exact preservation of all distances; such an embedding is called a *metric embedding with no distortion* or an *isometric embedding*. In general, $c^* = c^*(D, \mathcal{T})$ denotes the minimum possible distortion of a metric embedding of D into $\mathcal{T}$. (This definition is equivalent for both contractive and expansive metric embeddings, by scaling.)

3 *Comparison between Distortion and Relaxation*

The following propositions relate α^* and c^*.

PROPOSITION 1. *For any source and target metrics, $\alpha^* \le c^*$.*

PROOF. Consider an expansive metric embedding ϕ into (T, d) with distortion c. We show that ϕ is also an ordinal embedding into (T, d) with relaxation $\alpha \le c$. Consider a pair of distances $D(p,q)$ and $D(p',q')$ with ratio $D(p,q)/D(p',q')$ larger than c. (Thus, in particular, we label p, q, p', q' so that $D(p,q) > D(p',q')$.) Then $d(\phi(p), \phi(q))/d(\phi(p'), \phi(q')) \ge D(p,q)/(cD(p',q'))$ by expansiveness of $D(p,q)$ and distortion of $D(p',q')$. Thus $d(\phi(p), \phi(q))/d(\phi(p'), \phi(q')) > 1$, so $d(\phi(p), \phi(q)) > d(\phi(p'), \phi(q'))$ as desired. $\square$

Next we show that c^* and α^* can have an arbitrarily large ratio, even when the target metric is the real line.

PROPOSITION 2. *Embedding a uniform metric (where $D(p,q) = 1$ for all $p \ne q$) into the real line has $c^* = n - 1$ and $\alpha^* = 1$.*

PROOF. The mapping $\phi(p) = 0$, for all $p \in P$, is an ordinal embedding with no relaxation, because every distance remains equal (albeit 0). Any expansive metric embedding into the real line must have distance of at least 1 between consecutively embedded points, so the entire embedding must occupy an interval of length at least $n - 1$. The two points embedded the farthest away from each other therefore have distance of at least $n-1$, for a distortion of at least $n-1$. On the other hand, any embedding in which consecutively embedded points have distance exactly 1 has distortion $n - 1$. $\square$

Next we give a general bound on α^* that is essentially always finite. Define the *diameter* $\mathrm{diam}(D)$ of a metric D to be the ratio of the maximum distance to the minimum distance. (If the minimum distance is zero and the maximum distance is positive, then $\mathrm{diam}(D) = \infty$; if both are zero, then $\mathrm{diam}(D) = 1$.)

PROPOSITION 3. *For any source metric D and any target metrics, $\alpha^* \le \mathrm{diam}(D)$.*

PROOF. The mapping $\phi(p) = 0$, for all $p \in P$, has ordinal relaxation $\mathrm{diam}(G)$, because all nonequal comparisons between distances are violated, and the largest ratio between any two distances is precisely $\mathrm{diam}(D)$. $\square$

No such general finite upper bound exists for c^*, as evidenced by cut metrics. A *cut metric* is defined by a partition $P = A \cup B$ of the point set P into two disjoint sets A and B. The metric assigns a distance of 0 between pairs of points in A

and pairs of points in B, and assigns a distance of 1 between other pairs of points. If the source metric D has no zero distances and the target metrics are the cut metrics, then $c^* = \infty$, because some distance must become 0 which requires infinite distortion.

In contrast, α^* remains at most $\mathrm{diam}(D)$, and in some sense measures the quality of a clustering of the points into two clusters. Furthermore, the optimal α^* and clustering can be computed efficiently:

PROPOSITION 4. *The minimum-relaxation ordinal embedding of a specified metric into a cut metric can be computed in polynomial time.*

PROOF. First we guess the optimal relaxation α^* among $O(n^4)$ possibilities (the ratio of the distance between any two pairs of points). Second we guess a pair (p, q) of points on different sides of the cut and with minimum distance $D(p, q)$. Thus all pairs (r, s) of points with smaller distance $D(r, s) < D(p, q)$ must have r and s on the same side of the cut. Also, if there is any ordinal embedding of relaxation α^*, there cannot be pairs (r, s) of points with distance larger by a factor of α^*, that is, with $D(r, s) > \alpha D(p, q)$, because such distances will be mapped to a distance smaller or equal to 1, the mapped distance of (p, q). Similarly, there cannot be pairs (r, s) and (r', s') of points with distance less than $D(p, q)$ and with $D(r, s) > \alpha D(r', s')$, because those pairs are forced to map to equal distances of 0. Finally, all pairs (r, s) of points with $D(p, q) \le D(r, s) \le \alpha D(p, q)$ must have r and s on different sides of the cut if there is another distance $D(r', s') < D(r, s)/\alpha$, and otherwise are unconstrained.

All constraints of the form "r and s must be on the same side of the cut" and "r and s must be on different sides of the cut" can be phrased as a 2-SAT instance. Each point r has a variable x_r which is 0 if it placed in set A and 1 if it placed in set B. Each constraint thus has the form $x_r = x_s$ or $x_r \neq x_s$, which can be phrased in 2-CNF. Thus we can find an ordinal embedding into a cut metric with relaxation at most the guessed value of α^*, if one exists. □

Next we consider the related problem of ordinal embedding into the real line, which is a generalization of cut metrics. First we show that we can decide whether ordinal embedding is possible without relaxation in this case. The algorithm requires more sophistication (namely, guessing and linear programming) than the trivial algorithm for isometric embedding, where one can incrementally build an embedding in any Euclidean space in linear time.

PROPOSITION 5. *In polynomial time, we can decide whether a given metric can be ordinally embedded into the line.*

PROOF. The algorithm guesses the leftmost point p in the desired ordinal embedding. The leftmost point determines the order of all points: q is nonstrictly left of r if and only if $D(p, q) \le D(p, r)$. Given this ordering, we can write the embedding problem as a linear program: construct a variable x_q for the embedded coordinate of each vertex, write $x_r - x_q$ for the embedded distance $d(q, r)$ if q is nonstrictly left of r (i.e., $D(p, q) \le D(p, r)$), and include all pairwise linear (in)equality constraints on embedded distances. Thus the given metric has an ordinal embedding if and only if the linear program has a feasible solution for some guessed leftmost point p. □

Next we consider the worst case for ordinal embedding into the line. We show in particular that the cycle requires large relaxation. The cycle also requires large distortion into the line—indeed, it satisfies the same lower bound as what we show for relaxation—but the proof technique for ordinal relaxation is different from the usual packing argument that suffices for metric distortion.

PROPOSITION 6. *Ordinal embedding of the shortest-path metric of an unweighted cycle of even length n into the line requires relaxation at least $n/2$.*

PROOF. Suppose to the contrary that there is an ordinal embedding ϕ of the cycle into the line with relaxation less than $n/2$. Label the vertices of the cycle 1 through n in cyclic order. Assume without loss of generality that $\phi(1) < \phi(n/2{+}1)$. We must also have $\phi(2) < \phi(n/2+1)$, because otherwise $|\phi(2) - \phi(1)| \geq |\phi(n/2+1) - \phi(1)|$, contradicting that $\alpha < n/2$. Similarly, $\phi(2) < \phi(n/2+2)$, because otherwise $|\phi(n/2+2) - \phi(n/2+1)| \geq |\phi(n/2+2) - \phi(2)|$, again contradicting that $\alpha < n/2$. Repeating this argument shows that $\phi(3) < \phi(n/2+3)$, etc., and finally that $\phi(n/2+1) < \phi(1)$, a contradiction. $\square$

Section 5 shows that some trees also require $\Omega(n)$ ordinal relaxation into the line.

4 ℓ_p *Metrics are Universal*

In this section, we show that every distance matrix can be ordinally embedded without relaxation into ℓ_p space of a polynomial number of dimensions, for any fixed $1 \leq p \leq \infty$. This result is surprising in comparison to metric embeddings. Every metric can be embedded into ℓ_p using $O(\lg n)$ distortion [Bourgain 1985; Linial et al. 1995], and in the worst case $\Omega(\lg n)$ distortion is necessary for any $p < \infty$, as proved by Linial et al. [1995] for $p = 2$ and by Matoušek [1997] for all other values of p. In particular, the shortest-path metric of a constant-degree expander graph requires $\Omega(\lg n)$ distortion.

THEOREM 1. *Every distance matrix can be ordinally embedded without relaxation into $O(n^5)$-dimensional ℓ_p space, for any fixed $1 \leq p \leq \infty$.*

The same result was established independently by Bilu and Linial [2004] using an algebraic proof. Specifically, they show that every distance matrix can be ordinally embedded into $(n-1)$-dimensional Euclidean space, and then use the property that any Euclidean metric can be isometrically embedded into any ℓ_p space with at most $\binom{n}{2}$ dimensions. In constrast, our proof is purely combinatorial.

We can also reduce the number of dimensions for some values of p. For example, for $p = 2$, a simple rotation reduces the number of dimensions to $n - 1$.

Our proof proceeds in two steps. First we show that 0/1 Hamming metrics are universal in the same sense as Theorem 1. To conclude the proof, we note that there is an ordinal embedding without relaxation from 0/1 Hamming metrics into any ℓ_p metric. In fact, the pth root of the distances in a 0/1 Hamming metric can be metrically embedded without distortion into ℓ_p with the same number of dimensions. This second part is merely an observation, so the main work is in showing that 0/1 Hamming metrics are universal:[2]

[2]Note that finite 0/1 Hamming metrics and finite Hamming metrics are essentially the same, because one can be converted into the other with a dimension blowup that is multiplicative in the

LEMMA 1. *Every distance matrix can be ordinally embedded without relaxation into a 0/1 Hamming metric with $O(n^5)$ dimensions. In other words, any desired ordering on the distances between pairs of n points can be realized by a 0/1 Hamming metric on those n points.*

PROOF. Given a partial order $\mathcal{P}$ on a set of distances, we construct a 0/1 Hamming metric H such that $P_{i,j} < P_{k,l}$ implies $H_{i,j} < H_{k,l}$. If $\mathcal{P}$ is nontotal, then we can take any topological sort of $\mathcal{P}$ and realize it as a Hamming metric. This ordinal embedding will satisfy the original partial order so, from now on, we assume that $\mathcal{P}$ is a total order. Because $\mathcal{P}$ is an order on distances, defined by pairs of points, we can define it as a sequence of pairs $\mathcal{P} = [(a_0, b_0), (a_1, b_1), \ldots, (a_{\binom{n}{2}}, b_{\binom{n}{2}})]$, where in each pair, we arbitrarily select which node is a and which is b.

We now must produce a 0/1 vector for each point of the space so that the Hamming metric induced preserves the order $\mathcal{P}$. We assume that n is a power of 2; otherwise we can simply round n up to the next power of two.

Our main tool will be Hadamard matrices, defined as follows. Let $H_0 = [0]$, and

$$H_{i+1} = \begin{pmatrix} H_i & H_i \\ H_i & \overline{H_i} \end{pmatrix},$$

where $\overline{H_i}$ is the bitwise negation of H_i. Notice that the first row is the all-0 vector, denoted $\vec{0}$. Also, each row other than the first row consists of half 0s and half 1s. More strikingly, any two rows of H_i have Hamming distance 2^{i-1}, that is, they differ in half their positions. Finally, $\vec{1}$ has Hamming distance 2^{i-1} from any row except the first row, with which it has Hamming distance 2^i.

We generate a set of dimensions that code for each distance $\mathcal{P}_i$ and concatenate all the dimensions at the end. To code for distance $\mathcal{P}_i = (a_i, b_i)$, we set a_i's bits to be 0^{in} and b_i's bits to be 1^{in}. Every other point in the space besides these two gets a distinct row from the Hadamard matrix, repeated i times. Now the induced distances are $in/2$ for any pair of points except for a_i and b_i, which are at distance in.

Let the total number of dimensions be $d = n\binom{n}{2}(\binom{n}{2} + 1)/2$.

Consider now the distances between any pair a and b resulting from the concatenation of all d dimensions, and assume that $a = a_i$ and $b = b_i$, that is, their pairwise distance is the ith in the list. Then their pairwise distance is $(d + in)/2$. Thus, this embedding assigns to the ith smallest distance in $\mathcal{P}$ the ith smallest distance in the Hamming metric. □

5 *Approximation Algorithm for Unweighted Trees into the Line*

In this section, we give a 3-approximation algorithm for ordinally embedding the shortest-path metric induced by an unweighted tree into the line with approximately minimum relaxation. In contrast, the best approximation algorithm known for metrically embedding trees into the line with approximately minimum distortion is a recently discovered $O(n^{1/3})$-approximation [Bădoiu et al. 2005].

number of points. Thus our result could have been established with general Hamming metrics. However, our construction directly yields a 0/1 Hamming metric, so we do not need this extra conversion detail.

First we find a structure for proving lower bounds on the optimal relaxation:

LEMMA 2. *Given n such that 3 divides $n-1$, ordinal embedding of the shortest-path metric of an unweighted 3-spider with $(n-1)/3$ vertices on each leg of the spider (i.e., a 3-star with each edge subdivided into a path of $(n-1)/3$ edges) into the line requires relaxation at least $(n-1)/3$.*

PROOF. Suppose to the contrary that there is an ordinal embedding ϕ of the 3-spider into the line with relaxation $\alpha < (n-1)/3$. Label the vertices as follows: 0 denotes the root, and $a_1, \ldots, a_{(n-1)/3}$, $b_1, \ldots, b_{(n-1)/3}$, and $c_1, \ldots, c_{(n-1)/3}$ denote the nodes on the legs of the spider in order of their distance from the root 0. Because $\alpha < (n-1)/3$, $|\phi(a_{(n-1)/3}) - \phi(0)| > 0$, and the same holds for $b_{(n-1)/3}$ and $c_{(n-1)/3}$. Because the spider has three legs, two of $a_{(n-1)/3}, b_{(n-1)/3}, c_{(n-1)/3}$ are on the same side of the root 0 on the line. By possible reflection and renaming, assume that the a and b legs are both to the right of 0, and that $\phi(a_{(n-1)/3}) \geq \phi(b_{(n-1)/3}) > \phi(0)$. Let k be such that $\phi(a_k) < \phi(b_{(n-1)/3}) < \phi(a_{k+1})$ (where the label a_0 refers to the root 0). Such a k exists because $\alpha < (n-1)/3$, so $\phi(a_k) \neq \phi(b_{(n-1)/3})$ for all k, and because $\phi(0) < \phi(b_{(n-1)/3}) < \phi(a_{(n-1)/3})$. It follows that $|\phi(b_{(n-1)/3}) - \phi(a_{k+1})| < |\phi(a_{k+1}) - \phi(a_k)|$. In contrast, in the 3-spider graph, $b_{(n-1)/3}$ and a_{k+1} have distance of at least $(n-1)/3$, and a_{k+1} and a_k have distance 1. Therefore $\alpha > (n-1)/3$. □

DEFINITION 1. *Given a tree T, a* tripod *(a, b, c) is the union of shortest paths in T connecting every pair of vertices among $\{a, b, c\}$. The* root *r of the tripod is the common vertex among all three shortest paths. The* length *of the tripod is $k = \min\{D(r, a), D(r, b), D(r, c)\}$.*

Any tripod of length k induces a 3-spider with k vertices on each leg, by truncating all longer arms of the tripod to length k. Thus by Lemma 2, any tree with a tripod of length k must have ordinal relaxation of at least k. Using this lower bound, we obtain a constant-factor approximation algorithm.

THEOREM 2. *Given a tree T, there is an ordinal embedding $\phi : T \to \mathbb{R}$ of T into the line with relaxation $2k+1$, where k is the length of the largest tripod of T. The embedding can be computed in polynomial time.*

PROOF. If there are at most two leaves in the tree T, then T can be trivially embedded into the line without distortion or relaxation. Otherwise, T has a tripod. Let (A, B, C) be a longest tripod, let r be its root, and let k be its length. We view T as rooted at r. Let (a, b, c) be a tripod rooted at r that maximizes $D(r, a) + D(r, b) + D(r, c)$. This tripod corresponds to taking the longest three paths starting from different neighbors of r. In particular, all three paths have length of at least k, so the tripod (a, b, c) has length k. Relabel $\{a, b, c\}$ so that $D(r, a) = k$.

CLAIM 1. *For any $d \in \{a, b, c\}$, for any $d' \neq r$ on the path from r to d, and for any descendant x of d', $D(d', x) \leq D(d', d)$.*

PROOF. Assume to the contrary that $D(d', x) > D(d', d)$. If $d = a$, then there would be a larger tripod (x, b, c) rooted at r. Otherwise, assume without loss of generality that $d = b$. Then there would be a tripod (a, x, c), of the same length, and such that $D(r, a) + D(r, x) + D(r, c) > D(r, a) + D(r, b) + D(r, c)$, a contradiction. □

Claim 2. *For any $d \in \{b, c\}$, for any $d' \neq r$ on the path from r to d, and for any descendant x of d', such that the path from x to d' intersects the path from r to d only at vertex d', $D(d', x) \leq k$.*

Proof. Suppose to the contrary that $D(d', x) > k$. By the definition of d', $D(d', a) > D(r, a) = k$. By Claim 1, $D(d', d) \geq D(d', x)$. If $D(d', d) \leq k$, then $D(d', x) \leq D(d', d) \leq k$, a contradiction. If $D(d', d) > k$, then the tripod (x, d, a) (rooted at d') has length of at least $k + 1$, which is again a contradiction. □

Now we construct the embedding ϕ as follows. For every vertex x on the shortest path between b and c, we contract every subtree that intersects the path only at x into the single vertex x. The resulting graph is the same path from b to c, but where each vertex represents several vertices of the original graph. We embed this path into the line, placing the ith vertex along the path at coordinate i. This embedding places several vertices of the original graph at the same point in the line.

We claim that the depth of each contracted tree is at most k. For each subtree rooted at r (e.g., the one containing a), no vertex x in the subtree can have $D(r, x) > k$ because then we could have chosen that vertex as a and increased the objective function $D(r, a) + D(r, b) + D(r, c)$, a contradiction. For each subtree rooted at another node $b' \neq r$ on the path from b to c, we can apply Claim 2 and obtain that $D(b', x) \leq k$ for any vertex x in the subtree rooted at b'. Therefore the depth of each contracted tree is at most k.

Finally we claim that the ordinal relaxation of this mapping is at most $2k + 1$. Consider two vertices x and y belonging to contracted subtrees rooted at s and t, respectively. Their original distance is at most $2k + D(s, t)$, and their new distance is $D(s, t)$. Therefore the distance changes order with respect to distances of at least $D(s, t)$, for a worst-case ratio of $(2k + D(s, t))/D(s, t)$. This ratio is maximized when $D(s, t) = 1$ in which case it is $2k + 1$.

Corollary 1. *There is a polynomial-time algorithm to find ϕ of Theorem 2. The algorithm is a 3-approximation algorithm for ordinally embedding trees into a line.*

Proof. The proof of Theorem 2 is constructive, thus it gives an algorithm. Since the length of the largest tripod is a lower bound of embedding ordinally the tripod into a line, we obtain that the algorithm is a $(2 + 1/k)$-approximation algorithm. □

6 *Ultrametrics*

In this section we establish several results about ordinal embedding when the source metric or the target metrics are ultrametrics. Recall that an *ultrametric* is a metric with the strong triangle inequality $D(p, r) \leq \max\{D(p, q), D(q, r)\}$. See Barthelemy and Guenoche [1991] for basic results about ultrametrics.

6.1 Ultrametrics into the Line.

First we demonstrate that ultrametrics can be ordinally embedded into the line with relaxation 1. Here we exploit the minor difference between relaxation 1 and no relaxation, namely, that equality constraints can be violated, because, as described in the introduction, any ordinal embedding without relaxation of any ultrametric into Euclidean space requires $n-1$ dimensions. Thus the ordinal dimension of an ultrametric is just barely $n - 1$; the slightest relaxation allows us to obtain a much better embedding. Our result also contrasts metric embeddings where ultrametrics can be embedded into Euclidean space with $1+\varepsilon$ distortion, but such an embedding requires $\varepsilon^{-2}\lg n$ dimensions [Bartal and Mendel 2004].

Our construction is based on monotone stretching of the discrepancy between different distances:

LEMMA 3. *For any $k > 1$ and for any ultrametric $M = (P, D)$, there is an ultrametric $M' = (P, D')$ such that, for any $p, q, r, s \in P$, if $D(p,q) = D(r,s)$, then $D'(p,q) = D'(r,s)$, and if $D(p,q) > D(r,s)$, then $D'(p,q) \geq kD'(r,s)$.*

PROOF. Because M is an ultrametric, we can construct a weighted tree T, with P forming the set of leaves, such that the weights are nondecreasing along any path of T starting from the root. Moreover, for any $u, v \in P$, the ultrametric distance $D(u,v)$ is equal to the maximum weight of an edge along the path from u to v in T.

For $u, v \in P$, define $r(D(u,v)) = i$ where $D(u,v)$ is equal to the ith smallest distance in M. Consider now the weighted tree T' obtained from T by replacing an edge of weight w by an edge of weight $k^{r(w)}$. Let M' be the resulting ultrametric induced by T'. If $D(p,q) = D(r,s)$, then $r(D(p,q)) = r(D(r,s))$, so $D'(p,q) = D'(r,s)$. Finally, if $D(p,q) > D(r,s)$, then $r(D(p,q)) \geq r(D(r,s))+1$, so $D'(p,q) \geq kD'(r,s)$. $\square$

We combine this lemma with a result for the metric case:

LEMMA 4 [MATOUŠEK 1990]. *Any n-point metric space can be embedded into the line with distortion $12n$.*

Now we are ready to prove the main result of this subsection:

THEOREM 3. *Any ultrametric can be ordinally embedded into the line with relaxation 1.*

PROOF. Given an ultrametric $M = (P, D)$, by Lemma 3, we can obtain an ultrametric $M' = (P, D')$ such that, for any $p, q, r, s \in P$, if $D(p,q) = D(r,s)$, then $D'(p,q) = D'(r,s)$, and if $D(p,q) > D(r,s)$, then $D'(p,q) \geq 13nD'(r,s)$. Applying Lemma 4, we obtain a noncontractive metric embedding ϕ of P into the line such that, for any $p, q, r, s \in P$, if $D(p,q) > D(r,s)$, then $|\phi(p) - \phi(q)| \geq D'(p,q) \geq 13nD'(r,s) \geq \frac{13}{12}\|\phi(r) - \phi(s)\|$. Therefore ϕ is an ordinal embedding with relaxation 1. $\square$

6.2 ARBITRARY DISTANCE MATRICES INTO ULTRAMETRICS. In this section, we give a polynomial-time algorithm for computing an ordinal embedding of an arbitrary metric into an ultrametric with minimum possible relaxation.

We will show that the optimal ordinal embedding of a distance matrix M into an ultrametric is the *subdominant* of M [Farach et al. 1995]. One recursive construction of the subdominant is as follows. First, we compute a partition $P = P_1 \cup P_2 \cup \cdots \cup P_k$, for some $k \geq 2$, such that the minimum distance between any P_i and P_j is maximized. Such a partition can be found by computing a minimum spanning tree T of M, and partitioning the points by removing all the edges of T of maximum length. Let Δ be the maximum distance between any two points in P. We create a hierarchical tree representation for an ultrametric by starting with a root v_P and

k children $v_{P_1}, \dots, v_{P_k}$. The length of the edge $\{v_P, v_{P_i}\}$ is equal to Δ for each $i \in \{1, 2, \dots, k\}$. We recursively compute hierarchical tree representations for the metrics induced by the point sets $P_1, P_2, \dots, P_k$, and then we merge these trees by identifying, for each $i \in \{1, 2, \dots, k\}$, the root of the tree for P_i with the node v_{P_i}. In fact this entire construction can be carried out with a single computation of the minimum spanning tree, and thus takes linear time.

LEMMA 5. *Let $\Delta = \max_{p,q \in P} D(p, q)$ and let δ be the minimum distance between two points in different sets P_i and P_j, maximizing over all partitions $P = P_1 \cup P_2 \cup \cdots \cup P_k$ with $k \geq 2$. Then any ordinal embedding into an ultrametric has relaxation of at least Δ/δ.*

PROOF. Suppose that the maximum distance Δ is attained by points u, v with $u \in P_i$ and $v \in P_j$, where $i \neq j$. Consider an optimal ordinal embedding ϕ of M into a hierarchical tree representation T of an ultrametric. Thus the distance between two leaves p and q is equal to the maximum length of an edge along the unique path between p and q. No matter how ϕ splits P into subsets at the root of T, there exist $r, s \in P$ such that $D(r, s) \leq \delta$ and the path from r to s in T visits the root of T. (Here we assume without loss of generality that the root has at least two children; otherwise, we can simply remove the root.) Thus the path from r to s passes through the maximum edge in T. Hence, the maximum distance along the path between u and v in T cannot be larger than the maximum distance along the path between r and s in T. Therefore $d(\phi(u), \phi(v)) \leq d(\phi(r), \phi(s))$, while $D(u, v) = \Delta \geq \delta \geq D(r, s)$, so the relaxation is at least Δ/δ. □

THEOREM 4. *Given any distance matrix M, we can compute in polynomial time an optimal ordinal embedding of M into an ultrametric.*

PROOF. Let ϕ be the ordinal embedding of $M = (P, D)$ computed by the algorithm, with a hierarchical tree representation T. The maximum relaxation α of ϕ is attained for some $p, q, r, s \in P$ such that $D(p, q) \geq \alpha D(r, s)$ and $d(\phi(p), \phi(q)) < d(\phi(r), \phi(s))$. It follows that there exists an internal node v of T, with children v_1 and v_2, such that leaves p and q are descendants of v_1, while only one of the leaves r or s is a descendant of v_1. Assume without loss of generality that r is a descendant of v_1 and s is a descendant of v_2.

Consider the recursive call of the algorithm on a subset of points $P' \subseteq P$ in which the node v was created. Because r and s are in different subtrees of v, it follows that, in the partition of the set P' of points computed by the algorithm, the minimum distance between distinct sets is at most $D(r, s)$. On the other hand, the maximum distance between pairs of points in P' is at least $D(p, q)$. Thus, by Lemma 5, the optimal relaxation for ordinal embedding of M into an ultrametric is at least $D(p, q)/D(r, s) \geq \alpha$. □

By a similar argument, it can be shown that the same algorithm also computes a metric embedding of M into an ultrametric with minimum possible distortion. Furthermore, the distortion is equal to the relaxation in this embedding. In the next section, we show that ultrametrics are essentially the only case where this can happen universally for all distance matrices.

6.3 When Distortion Equals Relaxation.

Finally we show that, in a certain sense, ultrametrics are the only target metrics that have equal values of α^* and c^*, or even a universally bounded ratio between α^* and c^*, for all distance matrices (not just metrics).

Theorem 5. *If a set $\mathcal{T}$ of target metrics is closed under inclusion (i.e., closed under taking the submetric induced on a subset of points), and there is a constant C such that every distance matrix D has $c^*/\alpha^* \leq C$ (when embedding D into $\mathcal{T}$), then every metric in $\mathcal{T}$ is an ultrametric.*

Proof. Consider any metric M in $\mathcal{T}$. We claim that M has more than one diameter pair. Suppose to the contrary that only p and q attain the maximum distance in M. Let M_{+d} be the distance matrix identical to M except for $M_{+d}(p,q) = M(p,q)+d$. Let d be any positive real number greater than the sum of the second- and third-largest distances. Then M_{+d} is not in $\mathcal{T}$ because it violates the triangle inequality and $\mathcal{T}$ is a family of metrics. Because no other distance in M is equal to $M(p,q)$, M_{+d} can be ordinally embedded with no relaxation into $\mathcal{T}$ simply by taking M. However, M_{+d} cannot be metrically embedded into $\mathcal{T}$ without distortion, because M_{+d} is not in $\mathcal{T}$. Furthermore M_{+cd} cannot be metrically embedded into $\mathcal{T}$ with distortion less than c, because any contractive metric embedding must reduce the distance between p and q by a factor of c. Therefore the ratio between the minimum metric distortion c^* and the minimum ordinal relaxation α^* cannot be bounded.

Now by inclusion, any submetric of M induced by three points is also in $\mathcal{T}$, and therefore has a nonunique maximum edge. Thus all triangles in M are tall isosceles, which is one characterization of M being an ultrametric. $\square$

In fact, this theorem needs only that the set $\mathcal{T}$ of target metrics is closed under the operation of taking the induced metric on any triple of points.

7 *Worst Case of Unweighted Trees into Euclidean Space*

In this section, we consider the worst-case relaxation required for ordinal embedding of the shortest-path metric of an unweighted tree T into d-dimensional ℓ_2 space. Our work is motivated by that of Gupta [2000] and Babilon et al. [2003]. We show that, for any $d \geq 2$, and for any unweighted tree T on n nodes, $\alpha^* = \tilde{O}(n^{1/d})$. We complement this result by exhibiting a family of trees with optimal ordinal relaxation $\Omega(n^{1/(d+1)})$. In contrast, the best bounds on the worst-case distortion required are $\tilde{O}(n^{1/(d-1)})$ and $\Omega(n^{1/d})$ [Gupta 2000]. These ranges overlap at the endpoint of $\tilde{\Theta}(n^{1/d})$, but it seems that ordinal embedding and metric embedding behave fundamentally differently, in particular because different proof techniques are required for both the upper and lower bounds.

First we prove the upper bound. At a high level, the algorithm finds nodes that can be contracted to a single point, which can be an effective ordinal embedding, unlike metric embedding where it causes infinite distortion.

Theorem 6. *Any unweighted tree can be ordinally embedded into d-dimensional ℓ_2 space with relaxation $\tilde{O}(n^{1/d})$.*

Proof. Let $T = (V(T), E(T))$ be an unweighted tree with $|V(T)| = n$. We show how to obtain an ordinal embedding of T into d-dimensional ℓ_2 space with relaxation $\tilde{O}(n^{1/d})$.

We construct a new tree T' as follows. Initially, we set $T'_0 := T$. For $i = 1, \ldots, n^{1/d}$, we repeat the following process: set $T'_i := T'_{i-1}$, and for any leaf v of T'_{i-1}, remove v from T'_i. Let $T' := T'_{n^{1/d}}$.

Define the function $p : V(T) \to V(T')$ such that, for any $v \in V(T) \setminus V(T')$, $p(v)$ is the node in $V(T')$, which is closest to v, and, for any $v \in V(T')$, $p(v) = v$. It is easy to see that for every leaf v of T', there are at least $n^{1/d}$ nodes $u \in V(T) \setminus V(T')$ with $p(u) = v$. Thus, the number of leaves of T' is at most $n^{\frac{d-1}{d}}$.

It follows that using Gupta's algorithm [Gupta 2000], we can compute an expansive metric embedding ϕ' of T' into d-dimensional ℓ_2 space with distortion of at most $kn^{1/d}$, for some $k = \mathrm{polylog}(n)$. To obtain an embedding ϕ of T, we simply set $\phi(v) = \phi'(p(v))$ for each $v \in V(T)$.

It remains to show that ϕ has ordinal relaxation $\tilde{O}(n^{1/d})$. Let $v_1, v_2, v_3, v_4 \in V(T)$, with $v_3 \neq v_4$ and

$$d_T(v_1, v_2) > (2 + k)n^{1/d} d_T(v_3, v_4).$$

We have

$$\begin{aligned}
\|\phi(v_1) - \phi(v_2)\| &= \|\phi'(p(v_1)) - \phi'(p(v_2))\| \\
&\geq d_{T'}(p(v_1), p(v_2)) \\
&\geq d_T(v_1, v_2) - 2n^{1/d} \\
&> (2 + k)n^{1/d} d_T(v_3, v_4) - 2n^{1/d} \\
&\geq kn^{1/d} d_T(v_3, v_4) \\
&\geq kn^{1/d} d_{T'}(p(v_3), p(v_4)) \\
&\geq \|\phi'(p(v_3)) - \phi'(p(v_4))\| \\
&= \|\phi(v_3) - \phi(v_4)\|.
\end{aligned}$$

Thus, we obtain that ϕ has ordinal relaxation at most $(2 + k)n^{1/d} = \tilde{O}(n^{1/d})$. □

Next we prove the worst-case lower bound. The main novelty here is a new packing argument for bounding relaxation. Let $F(m, L)$ denote the m-spider with arms of length L, that is, an m-star with each edge refined into a path of length L.

LEMMA 6. *Any ordinal embedding of $F(m, L)$ into d-dimensional ℓ_2 space requires relaxation* $\Omega(\min\{L, m^{1/d}\})$.

PROOF. Let $T = F(m, L)$, and let $r \in V(T)$ be the only vertex of T with degree greater than 2. For any i, with $0 \leq i \leq L$, let $U_i = \{v \in V(T) \mid d_T(r, v) = i\}$.

Consider an optimal embedding $\phi : V(T) \to \mathbb{R}^d$ with relaxation α. We define

$$\begin{aligned}
\mu_i &= \min_{u,v \in V(T)} \{\|\phi(u) - \phi(v)\| \mid d_T(u, v) = i\}, \\
\lambda_i &= \max_{u,v \in V(T)} \{\|\phi(u) - \phi(v)\| \mid d_T(u, v) = i\}.
\end{aligned}$$

Observe that, if $\mu_{2L} = 0$, then there exist $u, v \in U_L$ such that $\phi(u) = \phi(v)$. It follows that, if $\alpha < 2L$, then for any $\{x, y\} \in E(T)$, $\phi(x) = \phi(y)$, which implies that all the vertices are mapped to the same point, and thus $\alpha = \Omega(L)$.

It remains to show that the assertion is true in the case $\mu_{2L} > 0$. Consider the nodes of U_L. For any $u, v \in U_L$, we have $d_T(u,v) = 2L$, and thus $\|\phi(u) - \phi(v)\| \geq \mu_{2L}$. For any $v \in U_L$, let B_v be the ball of radius $\mu_{2L}/2$ centered at $\phi(v)$. It follows that, for any $u, v \in U_L$, the balls B_u, B_v can intersect only on their boundary. Thus,

$$\begin{aligned}\left|\bigcup_{v \in U_L} B_v\right| &= \sum_{v \in U_L} |B_v| \\ &= \Omega(m\mu_{2L}^d)\end{aligned}$$

By a packing argument, we obtain that there exist $u, v \in U_L$ such that $\|\phi(u) - \phi(v)\| = \Omega(m^{1/d}\mu_{2L})$, which implies

$$\lambda_{2L} = \Omega(m^{1/d}\mu_{2L}). \tag{1}$$

Now consider two nodes $u, v \in U_L$ such that $\|\phi(u) - \phi(v)\| = \lambda_{2L}$, and let p be the path from u to v in T. It follows that there exist nodes $x, y \in p$ with $d_T(x,y) = 2L/\alpha$ and $\|\phi(x) - \phi(y)\| \geq \lambda_{2L}/\alpha$. Thus

$$\lambda_{2L/\alpha} \geq \frac{\lambda_{2L}}{\alpha}. \tag{2}$$

Also, by the definition of the ordinal relaxation, we have

$$\mu_{2L} > \lambda_{2L/\alpha}. \tag{3}$$

Combining Equations (1), (2), and (3), we obtain $\alpha\lambda_{2L/\alpha} = \Omega(m^{1/d}\mu_{2L}) = \Omega(m^{1/d}\lambda_{2L/\alpha})$. Thus we have shown that, if $\mu_{2L} > 0$, then $\alpha = \Omega(m^{1/d})$. The lemma follows. $\square$

Theorem 7. *For any $n > 0$ and any $d \geq 2$, there is a tree T on n nodes for which every ordinal embedding has relaxation $\Omega(n^{1/(d+1)})$.*

Proof. The theorem follows from Lemma 6, for $T = F(n^{d/(d+1)}, n^{1/(d+1)})$. $\square$

8 *Arbitrary Metrics into Low Dimensions*

By Lemma 1, a general $O(\lg n)$ upper bound on relaxation carries over from metric embeddings of any n-point metric space into $O(\lg n)$-dimensional Euclidean space, using theorems of Bourgain [1985] and of Johnson and Lindenstrauss [1984]. For metric distortion, this bound is tight [Linial et al. 1995], but one might suspect that the ordinal relaxation can be smaller. Here we show that it cannot be much smaller: some n-point metric spaces require relaxation $\Omega(\log n/\log\log n)$. This claim is a special case of the following result.

Theorem 8. *There is an absolute constant $c > 0$ such that, for every d and n, there is a metric space T on n points such that the relaxation of any ordinal embedding of T into d-dimensional Euclidean space is at least $\frac{\log n}{\log d + \log\log n + c} - 1$.*

The proof is based on two known results. The first is a bound of Warren [1968] on the number of sign patterns of a system of real polynomials. The second is the existence of dense graphs with no short cycles. We first state these two results.

Let $P_j = P_j(x_1, \ldots, x_\ell)$, $j = 1, \ldots, m$, be m real polynomials. For a point $u = (u_1, \ldots, u_\ell) \in \mathbb{R}^\ell$, the *sign pattern* of the P_j's at u is the m-tuple $(\varepsilon_1, \ldots, \varepsilon_m) \in$

$(-1,0,1)^m$, where $\varepsilon_j = \text{sign } P_j(u)$. Let $s(P_1, P_2, \ldots, P_m)$ denote the total number of sign patterns of the polynomials $P_1, P_2, \ldots, P_m$, as u ranges over all points of $\mathbb{R}^\ell$.

The following result is derived by [Alon 1995] as a slight modification of a theorem of Warren [1968].

Theorem 9. *Let $P_1 \ldots P_m$ be m real polynomials in ℓ real variables, and suppose the degree of each P_j does not exceed k. If $2m \geq \ell$, then $s(P_1 \ldots P_m) \leq (8ekm/\ell)^\ell$.*

The following statement follows from a result of Erdős and Sachs [1963] and can also be proved directly by a simple probabilistic argument.

Lemma 7. *For every $g \geq 3$ and every $n \geq 3$, there are (connected) graphs on n vertices with at least $\frac{1}{4}n^{1+1/g}$ edges and with no cycle of length at most g.*

We note that there are slightly better known results based on certain algebraic constructions, but for our purpose here this estimate suffices.

We can now prove Theorem 8. Throughout the proof and the rest of the section, we assume that n is large, whenever this is needed, and omit all floor and ceiling signs whenever these are not crucial.

Proof of Theorem 8. Without trying to optimize the constants, define $g = \frac{\log n}{\log d + \log\log n + 5}$. We will show that some n-point metric spaces require relaxation of at least $g-1$. Without loss of generality, assume $g-1$ is bigger than 1, as otherwise there is nothing to prove. By Lemma 7, there is a graph $G=(V,E)$ on a set $V=\{1,2,\ldots,n\}$ of n labeled vertices, with $m \geq \frac{1}{4}n^{1+1/g} > 7nd\log n$ edges, and with no cycles of length at most g. For every subset $E' \subset E$ of precisely $m/2$ edges, the subgraph (V,E') of G defines a metric space $T(E')$ on the set V (if the subgraph is disconnected, some distances can be defined to be infinite; alternatively, we can fix a spanning tree in G and include it in all subgraphs to make sure they are all connected). This gives us a collection of $2^{(1+o(1))m}$ metric spaces on V with the following property.

(*) For every two distinct spaces (T,d) and (T',d') in the collection, there are two pairs of points x,y and z,w so that $d(x,y)=1$ and $d'(x,y) > g-1$, whereas $d'(z,w)=1$ and $d(z,w) > g-1$.

Indeed, this follows from the fact that, for every two distinct subgraphs in our collection, there is an edge $\{x,y\}$ belonging to the first one and not to the second, and vice versa. As the shortest cycle in G is of length exceeding g, the claim in (*) follows.

Fix a space T in our collection, and let ϕ_T be a minimum relaxation embedding of it into d-dimensional Euclidean space. Let $\phi_T(i) = (x^T_{i,1}, \ldots, x^T_{i,d})$. Then the square of the Euclidean distance between each two points in the embedding can be expressed as a polynomial of degree 2 in the dn variables $x^T_{i,j}$. The difference between two such squares of distances is thus also a polynomial of degree 2 in these variables. It follows that the order of all $\binom{n}{2}$ distances is determined by the signs of $\binom{n}{2}^2 < n^4/4$ polynomials, each of degree 2, in dn variables. By Theorem 9, the total number of such orders is at most

$$\left(\frac{16en^4}{4dn}\right)^{dn} \leq n^{(3+o(1))dn} = 2^{(3+o(1))nd\log n}.$$

This is smaller than the number of spaces in our collection, and hence, by the pigeonhole principle, there are two distinct spaces T and T' in our collection, so that the orders of the distances in their embeddings are the same. This, together with (*), implies that the relaxation in at least one of these embeddings is at least $g-1$, completing the proof. □

The last proof easily extends to embeddings into d-dimensional ℓ_p space for any even integer p. The only difference is that, in this case, the pth power of the distance between a pair of given points in the embedding is a polynomial of degree p in the dn variables describing the embedding. Working out the computation in the previous proof yields the following result.

THEOREM 10. *There is an absolute constant $c > 0$ such that, for every d and n, and for every even integer p, there is a metric space T on n points such that the relaxation in any ordinal embedding of T into d-dimensional ℓ_p space is at least* $\frac{\log n}{\log d + \log(\log n + \log p) + c} - 1$.

This argument, combined with an additional trick, can in fact be extended to handle ordinal embeddings into d-dimensional ℓ_p space for odd integers p, as well as embeddings into d-dimensional ℓ_∞ space.

THEOREM 11. *(i) For every $n \geq d$, there is a metric space T on n points such that the relaxation in any ordinal embedding of T in d-dimensional ℓ_∞ space is at least* $\frac{\log n}{\log d + \log\log n + O(1)} - 1$.

(ii) For every $n \geq d$, and for every odd positive integer p, there is a metric space T on n points such that the relaxation of any ordinal embedding of T into d-dimensional ℓ_p space is at least $\frac{\log n}{\log(2d^2 + 3d\log n + d\log p + O(d))} - 1$.

PROOF. As before, the result is proved by a counting argument: we prove that the number of possible orders between all distances in a set of n points in the relevant spaces is not too large, and use the fact that there are many significantly different metric spaces on n points, concluding that, for two such metric spaces, the embedding orders the distances identically, and hence derive the required lower bound on relaxation.

(i) We start by bounding the number of possible orders of all distances in a set X of n points in d-dimensional ℓ_∞ space. Given such a set, define, for each ordered set (x, y, z, w) of (not necessarily distinct) four points of X, and for each two indices i, j in $\{1, 2, \ldots, d\}$, the following linear polynomial in the dn variables representing the coordinates of the points: $(x_i - y_i) - (w_j - z_j)$. By Theorem 9, these d^2n^4 polynomials have at most $(O(1)dn^3)^{dn} \leq 2^{(4+o(1))dn\log n}$ sign patterns. (In fact, because the polynomials here are linear, there is a slightly better, and simpler, estimate than the one provided by Warren's Theorem [1968] (see Harding [1967]) but the asymptotic of the logarithm in this estimate is the same.) We claim that the signs of all these polynomials determine completely the order on all the $\binom{n}{2}$ distances between pairs of the points. Indeed, the signs of the polynomials $(x_i - y_i) - (x_j - y_j)$,

$(x_i-y_i)-(y_j-x_j)$ (and their inverses) determines a coordinate i such that $||x-y||_\infty$ is x_i-y_i or y_i-x_i (as this is simply the maximum of all $2d$ differences of the form $(x_i-y_i), (y_i-x_i)$). Suppose, now, that $||x-y||_\infty = x_i-y_i$ and $||w-z||_\infty = w_j-z_j$. Then the sign of $(x_i-y_i)-(w_j-z_j)$ determines which of the two distances is bigger. It follows that the total number of orders of the distances of n points in d-dimensional ℓ_∞ space is at most $2^{(4+o(1))dn\log n}$.

Define $g = \frac{\log n}{\log d+\log\log n+5}$, take a graph $G=(V,E)$ as in the proof of Theorem 8, and construct a collection of $2^{(1+o(1))7nd\log n}$ metric spaces on a set of n points satisfying (*). The desired result follows, just as in the proof of Theorem 8.

(ii) As in the proof of part (i), we first bound the number of possible orders of all distances in a set X of n points in d-dimensional ℓ_p space. Given such a set, define, for each two (not necessarily distinct) pairs $\{x,y\}$ and $\{z,w\}$ of points, and each two sign vectors

$$(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_d),(\delta_1,\delta_2,\ldots,\delta_d)\in\{-1,1\}^d,$$

the following polynomial in the dn coordinates of the points:

$$\sum_{i=1}^{d}\varepsilon_i(x_i-y_i)^p-\sum_{j=1}^{d}\delta_j(z_j-w_j)^p.$$

This is a set of $2^{2d}n^4$ polynomials, each of degree p, and thus, by Theorem 9, the number of their sign patterns is bounded by

$$2^{2d^2n+3dn\log n+dn\log p+O(dn)}. \tag{4}$$

As before, it is not difficult to see that the signs of all these polynomials determine completely the order of all distances between pairs of points. Therefore, the number of such orders does not exceed (4). The desired result now follows as before, by considering metrics induced by subgraphs with half the edges of a graph on n vertices with at least $\frac{1}{4}n^{1+1/g}$ edges and no cycles of length at most g, where $g=\frac{\log n}{\log(2d^2+3d\log n+d\log p+O(d))}$. □

9 *Conclusion and Open Problems*

We have introduced minimum-relaxation ordinal embeddings and shown that they have distinct and sometimes surprising behavior. Yet many problems remain to be explored in this context; our hope is that this article forms the foundation for a fruitful body of research. We highlight some of the more important directions for future exploration.

An important line of study is to continue comparing ordinal embeddings with metric embeddings. One interesting question is whether the dimensionality-reduction results of Bourgain [1985] and Johnson and Lindenstrauss [1984] can be improved upon for ordinal relaxation. From Theorem 8 and Proposition 1, we know that the optimal worst-case relaxation for an ordinal embedding of a general metric into $O(\lg n)$-dimensional Euclidean space is between $\Omega(\lg n/\lg\lg n)$ and $O(\lg n)$. Closing this $\Theta(\lg\lg n)$ gap is an intiguing open problem; a better upper bound would improve on Bourgain-based metric embeddings into $O(\lg n)$ dimensions. Another problem is how much relaxation is required for dimensionality reduction of a metric already embedded in arbitrary dimensional ℓ_p space. For $p = 2$, we obtain an ideal relaxation of $1 + \varepsilon$ using Johnson-Lindenstrauss combined with Proposition 1. For $p = \infty$, dimensionality reduction is impossible, by Theorem 11(i), because ℓ_∞ is universal in the metric sense. For $p \neq 2, \infty$, the problem is open; in contrast, it is known for metric embeddings that dimensionality reduction is impossible for ℓ_1 [Brinkman and Charikar 2003; Lee and Naor 2004].

Another important direction is to develop more approximation algorithms for minimum-relaxation ordinal embedding. Unlike general upper bounds on distortion, existing approximation algorithms for minimum-distortion metric embedding do not carry over to ordinal embedding because the optimum solution is generally smaller. Our $O(1)$-approximation result in Section 5, and the lack of a matching result for metric embedding despite much effort, shows that in some contexts ordinal embedding problems may prove more easily approximable than metric embedding. We expect that our approximation result can be generalized using similar techniques for unweighted graphs, weighted trees, and/or higher dimensions, and that it can be strengthened to a PTAS. A related open problem is to consider trees as target metrics, and find the tree metric into which a given metric can be ordinally embedded with approximately minimum relaxation. Another family of approximation problems arise with the related notion of *additive relaxation*, in contrast to (multiplicative) relaxation, where pairs of distances within an additive α must have their relative order preserved. In some cases, approximation results may be harder for ordinal embedding than metric embedding. For example, in the problem of approximating the minimum additive distortion/relaxation for an ordinal embedding of an arbitrary metric into the line, the simple greedy algorithm of Proposition 5 is a 3-approximation for metric embedding but can be arbitrarily bad for ordinal embedding.[3]

A final direction to consider is finding other applications of ordinal embedding. In particular, in the context of approximation algorithms for other problems, when are low-relaxation ordinal embeddings as useful as (and more powerful than) low-distortion metric embeddings? Nearest neighbor is a simple example where only the order of distances is relevant, but we expect there are several other such problems.

ACKNOWLEDGMENTS

We thank Amir Ben-Dor, Piotr Indyk, Howard Karloff, and Yuval Rabani for helpful discussions. In particular, Amir Ben-Dor contributed key ideas to the proof of universality (Theorem 1). Part of this work was done during a visit of N. Alon, E. D. Demaine, and M. Hajiaghayi at Microsoft Research in Redmond, WA; we thank our hosts at Microsoft for their hospitality.

REFERENCES

Agarwala, R., Bafna, V., Farach, M., Paterson, M., and Thorup, M. 1999. On the approximability of numerical taxonomy (fitting distances by tree metrics). *SIAM J. Comput. 28,* 3, 1073–1085.

[3]The example is as follows. The graph has four points a, b, p, q and $D(p, q) = 50$, $D(p, a) = 100$, $D(p, b) = 100 + \varepsilon$, $D[q, a] = 70$, $D(q, b) = 60$, and $D(a, b) = 10$. The optimum ordinal embedding can place the points in the order p, q, b, a, and has additive relaxation ε, while greedy places the points in the order p, q, a, b and has additive relaxation 10, resulting in an arbitrarily large ratio.

Alon, N. 1995. Tools from higher algebra. In *Handbook of combinatorics*, R. L. Graham, M. Grötschel, and L. Lovász, Eds. Vol. 2. MIT Press, Chapter 32, 1749–1783.

Alon, N., Frankl, P., and Rödl, V. 1985. Geometrical realization of set systems and probabilistic communication complexity. In *Proceedings of the 26th Annual Symposium on Foundations of Computer Science*. Portland, Oregon, 277–280.

Babilon, R., Matoušek, J., Maxová, J., and Valtr, P. 2003. Low-distortion embeddings of trees. *Journal of Graph Algorithms and Applications 7,* 4, 399–409.

Bădoiu, M., Dhamdhere, K., Gupta, A., Rabinovich, Y., Raecke, H., Ravi, R., and Sidiropoulos, A. 2005. Approximation algorithms for low-distortion embeddings into low-dimensional spaces. In *Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms*. Vancouver, British Columbia, Canada.

Bartal, Y. and Mendel, M. 2004. Dimension reduction for ultrametrics. In *Proceedings of the 15th ACM-SIAM Symp. on Discrete Algorithms*. 664–665.

Barthelemy, J.-P. and Guenoche, A. 1991. *Trees and Proximity Representations*. Wiley.

Bilu, Y. and Linial, N. 2004. Monotone maps, sphericity and bounded second eigenvalue. arXiv:math.CO/0401293.

Bourgain, J. 1985. On Lipschitz embedding of finite metric spaces in Hilbert space. *Israel J. Math. 52,* 1-2, 46–52.

Brinkman, B. and Charikar, M. 2003. On the impossibility of dimension reduction in ℓ_1. In *Proceedings of the 44th Symposium on Foundations of Computer Science*. 514–523.

Bădoiu, M. 2003. Approximation algorithm for embedding metrics into a two-dimensional space. In *Procedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms*. 434–443.

Bădoiu, M., Demaine, E. D., Hajiaghayi, M., and Indyk, P. 2004. Low-dimensional embedding with extra information. In *Proceedings of the 20th Annual ACM Symposium on Computational Geometry*. Brooklyn, New York.

Chor, B. and Sudan, M. 1998. A geometric approach to betweennes. *SIAM Journal on Discrete Mathematics 11,* 4, 511–523.

Cunningham, J. P. and Shepard, R. N. 1974. Monotone mapping of similarities into a general metric space. *Journal of Mathematical Psychology 11*, 335–364.

Erdős, P. and Sachs, H. 1963. Reguläre Graphen gegebener Taillenweite mit minimaler Knotenzahl. *Wiss. Z. Martin-Luther-Univ. Halle-Wittenberg Math.-Natur. Reihe 12*, 251–257.

Farach, M. and Kannan, S. 1999. Efficient algorithms for inverting evolution. *J. ACM 46,* 4, 437–449.

Farach, M., Kannan, S., and Warnow, T. 1995. A robust model for finding optimal evolutionary trees. *Algorithmica 13,* 1-2, 155–179.

Gupta, A. 2000. Embedding tree metrics into low-dimensional Euclidean spaces. *Discrete Comput. Geom. 24,* 1, 105–116.

Harding, E. F. 1966/1967. The number of partitions of a set of N points in k dimensions induced by hyperplanes. *Proceedings of the Edinburgh Mathematical Society, Series II 15*, 285–289.

Håstad, J., Ivansson, L., and Lagergren, J. 2003. Fitting points on the real line and its application to RH mapping. *Journal of Algorithms 49,* 62, 42–62.

Holman, W. 1972. The relation between hierarchical and euclidean models for psychological distances. *Psychometrika 37,* 4, 417–423.

Indyk, P. and Matoušek, J. 2004. Low-distortion embeddings of finite metric spaces. In *Handbook of Discrete and Computational Geometry*, Second ed., J. E. Goodman and J. O'Rourke, Eds. CRC Press, Chapter 8, 177–196.

Ivansson, L. 2000. Computational aspects of radiation hybrid. Ph.D. thesis, Department of Numerical Analysis and Computer Science, Royal Institute of Technology.

Johnson, W. B. and Lindenstrauss, J. 1984. Extensions of lipshitz mapping into hilbert space. *Contemporary Mathematics 26*, 189–206.

Kruskal, J. B. 1964a. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika 29*, 1–28.

Kruskal, J. B. 1964b. Non-metric multidimensional scaling. *Psychometrika 29*, 115–129.

LEE, J. R. AND NAOR, A. 2004. Embedding the diamond graph in L_p and dimension reduction in L_1. *Geometric and Functional Analysis 14,* 4, 745–747.

LINIAL, N., LONDON, E., AND RABINOVICH, Y. 1995. The geometry of graphs and some of its algorithmic applications. *Combinatorica 15,* 2, 215–245.

MATOUŠEK, J. 1990. Bi-Lipschitz embeddings into low-dimensional Euclidean spaces. *Commentationes Mathematicae Universitatis Carolinae 31,* 3, 589–600.

MATOUŠEK, J. 1997. On embedding expanders into l_p spaces. *Israel Journal of Mathematics 102*, 189–197.

OPATRNY, J. 1979. Total ordering problem. *SIAM J. Computing 8*, 111–114.

SHAH, R. AND FARACH-COLTON, M. 2004. On the complexity of ordinal clustering. *Journal of Classification*. To appear.

SHEPARD, R. N. 1962. Multidimensional scaling with unknown distance function I. *Psychometrika 27*, 125–140.

TORGERSON, W. S. 1952. Multidimensional scaling I: Theory and method. *Psychometrika 17,* 4, 401–414.

WARREN, H. E. 1968. Lower bounds for approximation by nonlinear manifolds. *Transactions of the American Mathematical Society 133*, 167–178.

RECEIVED DECEMBER 2006; ACCEPTED MARCH 2007