Greedy Embeddings, Trees, and Euclidean vs. Lobachevsky Geometry

Petar Maymounkov*
Massachusetts Institute of Technology
petar@csail.mit.edu

Abstract

A *greedy embedding* of an unweighted undirected graph $G = (V, E)$ into a metric space (X, ρ) is a function $f : V \to X$ such that for every *source-sink* pair of different vertices $s, t \in V$ it is the case that s has a neighbor v in G with $\rho(f(v), f(t)) < \rho(f(s), f(t))$.

Finding greedy embeddings of connectivity graphs helps to build distributed routing schemes with compact routing tables. In this paper we take a refined look at greedy embeddings, previously addressed in [1, 2], by examining their description complexity as a key parameter in conjunction with their dimensionality. We give arguments showing that the dimensionality lower-bounds for monotone maps do not extend to greedy embeddings. We prove a *unified* $O(\log n)$ lower-bound on the dimension of no-stretch greedy embeddings when the host metric is Euclidean or Lobachevsky geometry. The essence of the lower bound entails showing that low-dimensional spaces lack the topological capacity to realize the embeddings of certain graphs with "hard crossroads." This technique might be of independent interest. We develop new methods for building *concise* embeddings of trees (and some other graphs) in 3-dimensional Lobachevsky spaces using recursive applications of hyperbolic isometries guided by caterpillar-like decompositions. Our embeddings improve over prior work [1] by achieving $O(\kappa(T) \cdot \log n)$ description complexity, where $\kappa(T)$ is the caterpillar dimension. We further demonstrate concise $O(\log n)$-dimensional greedy embeddings of trees into Euclidean space using techniques inspired by [3], thereby strengthening our belief and intuition that all graphs can be embedded with no stretch in $\ell_2^{O(\log n)}$.

*PhD candidate.

1. Introduction

A *greedy embedding* of an unweighted undirected graph $G = (V, E)$ into a metric space (X, ρ) is a function $f : V \to X$ such that for every *source-sink* pair of different vertices $s, t \in V$ it is the case that s has a neighbor v in G with $\rho(f(v), f(t)) < \rho(f(s), f(t))$. From here on $n = |V|$ and the word "embedding" will refer to a greedy one unless otherwise stated. This definition implies that routing greedily (with respect to the host metric) in G always succeeds. In particular, a routing algorithm induced by a given greedy embedding works as follows. To deliver a letter from $s \in V$ to $t^* \in X$, the algorithm recursively forwards the letter to a neighbor of minimal embedding distance to t^* (ties are broken arbitrarily in a deterministic manner which is universally fixed for the purposes of our discussion), provided that such neighbor is closer to t^* than the current vertex. Otherwise, routing halts and it is assumed that the target has been reached. If the embedding is greedy and $t^* = f(t)$ for some $t \in V$, then the letter is guaranteed to reach t.

The notion of a greedy embedding is motivated by its applications to routing-with-local-information in large distributed systems (discussed in more detail later). In this context one is particularly concerned with three properties of the embedding algorithm. From here on we will loosely use the term *embedding* to refer to f itself or to the algorithm that finds an embedding for a given input graph.

i. For a given embedding algorithm, the maximum (over $v \in G$) number of bits that the algorithm uses to describe $f(v)$ is called the *description complexity* of the embedding (algorithm). Note that in a typical application, the computer at node v stores its own coordinate $f(v)$ in order to be able to perform routing tasks. Embeddings with $\Omega(n)$ description complexity are not interesting in light of application constraints. Our primary interest is in embeddings with polylog(n) description complexity, heretofore referred to as *concise* embeddings.

ii. Every embedding f defines a unique path in G between all pairs of unequal vertices (s, t), which is the path realized by the greedy routing algorithm (when routing from s to $f(t)$ and vice-versa) with respect to f (after resolving ties deterministically, as noted above). The length of this path is denoted by $d_f(s, t)$, which is to be distinguished from the length of the shortest-path between s and t in G denoted by $d_G(s, t)$. With this notation in hand, the *stretch* of a greedy embedding is defined as $D = \max_{s \neq t \in V} \frac{d_f(s,t)}{d_G(s,t)}$. An embedding with stretch $D = 1$ is called a *no-stretch* embedding. Note that no-stretch greedy embeddings are not equivalent to no-stretch distance-preserving embeddings. (Examples are given below.)

iii. The *congestion* of a greedy embedding is defined as the edge-congestion of the set of routes realized by greedy routing (with respect to the embedding) between all pairs of vertices.

1.1. History of the problem

The power of geometric interpretation for routing problems was initially recognized in a sequence of papers [4, 5, 6, 7] from the ad-hoc, wireless and sensor networks communities. These papers consider the problem of routing messages in ad-hoc wireless networks where participating nodes are aware of their physical planar location on Earth; and, additionally, the connectivity graph (induced by the nature of radio communications) is close to planar. The papers describe routing algorithms that make local forwarding decisions based on the geographic location of the target node and the current node's neighbors. The algorithms have a common framework. First, a planarization of the connectivity graph is obtained; consequently, routing consists of greedy approach towards the target, combined with face routing around the perimeter of obstacles (when greedy approach is not possible).

Routing with geographic location, however, has unsurmountable shortcomings. In particular, [7] shows that the best possible routing algorithm (based only on geographic location) may result in routing costs that are quadratic in the size of the optimal ones. This negative result is due to the arbitrary geometric complexity that the obstacles can have. Two other shortcomings are the assumption that the connectivity graphs are close to planar and that geographic location information is available. These two assumptions render the routing schemes under consideration useless for more complex networks like the Internet or P2P networks, where the connectivity graphs are significantly more complex: Such graphs are often modeled by scale-free or preferential-attachment random graphs [8, 9].

In light of these shortcomings, the non-strictly-theoretical approaches of [10, 11, 12] consider assigning virtual coordinates in $\mathbb{R}^d$ to network nodes, so that basic greedy routing works with little modification. The assignment methods investigated entail variants of the rubber-band algorithm (applied to the connectivity graphs) and multi-dimensional scaling techniques (applied to the shortest-path metric of the connectivity graphs). The experimental results of these papers are generally promising but far from perfect or well-understood for large and realistic classes of graphs.

The first rigorously theoretic attempt at the problem was made by Papadimitriou et al. in [2], where the notion of greedy embedding was defined. This paper concerns the question of mere existence of greedy embeddings for graphs (irrespective of stretch or congestion). The paper shows that any graph containing a 3-connected planar subgraph has a greedy embedding in $\mathbb{R}^3$, and conjectures that every such graph has a greedy embedding in $\mathbb{R}^2$ as well. This conjecture was proven correct for graphs containing a triangulated planar subgraph [13].

Following the work of [2], and perhaps motivated by the application of greedy embeddings, Kleinberg [1] asks the more general question: *Are there (nice) host metric spaces that accommodate greedy embeddings of all connected graphs?*. He answers this question affirmatively using the fact that a greedy embedding of a graph spanning tree is also a greedy embedding of the graph (albeit with possibly arbitrary stretch), and showing that all trees can be embedded (not concisely) into $\mathbb{H}^2$, the 2-dimensional Lobachevsky space. Additionally, his paper highlights the importance of the stretch and congestion parameters of greedy embeddings in view of applications. For embeddings of star graphs on n vertices into $\mathbb{R}^d$ endowed with a Minkowski norm, it is shown that $d = \Omega(\log n)$. Kleinberg concludes his work with a list of open problems regarding existential and algorithmic aspects of greedy embeddings with various levels of stretch.

1.2. Our results

This work is a continuation of the study of greedy embeddings. The primary theme in our paper is addressing the existence of no-stretch embeddings for all graphs. Our findings provide evidence that such embeddings may exist. Our emphasis on finding embeddings with no stretch is in line with the fact that in real-world routing applications even small amounts of stretch are prohibitive.

First, in the spirit of keeping the applications in mind, we address the bit complexity of greedy embeddings (defined above). We improve Kleinberg's result by showing that all trees (as well as some other tree-like graphs) have concise (also defined above) greedy embeddings into $\mathbb{H}^3$, the 3-dimensional Lobachevsky space. We complement this result by exhibiting concise low-dimensional greedy embeddings of trees into ℓ_2. The latter construction sheds some light on the "shape" of possible greedy embeddings of general graphs in Euclidean spaces.

Second, we give arguments and a theorem that strongly suggest that no-stretch embeddings do not require high dimension and, in fact, we believe that all connected graphs have concise no-stretch low-dimensional embeddings in ℓ_2. Therefore, we begin a systematic attempt to understand the structure of no-stretch embeddings. As a first step, we develop a unified technique with a strong topological flavor that demonstrates that a certain family of graphs with "hard crossroads" requires $\Omega(\log n)$ dimensions to embed into Lobachevsky or Euclidean space. This technique motivates an interesting topological question regarding Minkowski normed spaces and manifolds. Our lower-bound can be interpreted as saying that Lobachevsky geometry is no more powerful than Euclidean geometry when it comes to harder graphs, contrary to intuition. We complement this lower-bound with a theorem stating that every no-stretch greedy embedding into ℓ_2^d can be used to derive a corresponding embedding into $\mathbb{H}^{d+1}$.

This paper is not concerned with congestion since we believe that this parameter is of secondary importance. This belief is supported by the fact that standard techniques like using a distribution over embeddings or routing through randomly selected intermediaries can be used to reduce congestion.

The paper is organized as follows. Section 2 positions our work with respect to the related class of ordinal and proximity embeddings. Section 3 discusses the preliminaries of hyperbolic geometry, greedy embeddings, and tree decompositions. Section 4 proves a lower-bound on embedding dimension for a family of graphs with rich combinatorial structure. Sections 5 and 6 explain our concise embeddings for trees in Lobachevsky and Euclidean space, respectively. Finally, Section 7 contains concluding remarks and open problems motivated by our work.

2. Related work

Greedy embeddings are a type of *ordinal embeddings*, i.e. embeddings that preserve the relative order of pairwise vertex distances. The latter have enjoyed significant attention in the multi-dimensional scaling community in view of their applications to visualization, compression, nearest-neighbor search, etc. (see [14] for details). The strictest kind of ordinal embeddings are monotone maps, which are discussed below. Monotone maps provably require $\Omega(n)$ dimensions to realize almost all distance orders on n-point metrics (see [15]). To address this problem [14] considers *ordinal embeddings of minimum relaxation*, a variant that enforces order preservation of well-separated points only. In this vein, greedy embeddings are a variant of ordinal embeddings that require order preservation only among pairs of points of the form (x, z) and (y, z) where x and y must be neighbors or share a neighbor in the original graph.

In [15] Linial and Bilu study embeddings that preserve relative pairwise distances. They define a *monotone map* as function $f : X \to Y$ mapping a finite metric (X, d_X) into a (usually normed or otherwise nice) host metric (Y, d_Y) such that $\forall\ a, b, c, d \in X :\ d_X(a, b) < d_X(c, d) \Leftrightarrow d_Y(f(a), f(b)) < d_Y(f(c), f(d))$. They show that any ordering on $\binom{[n]}{2}$ can be realized as a metric on n points with a matching ordering on the pairwise distances. Furthermore, they show that any such metric can be embedded in ℓ_2^n and almost no such metrics can be embedded in $\ell_2^{o(n)}$.

We prove a theorem showing that the relationship between monotone maps and greedy embeddings is weak at best. The proof of the following theorem is deferred to the full version:

Definition 2.1. A *reduction* of the problem of finding a monotone map for a given order $\pi \in S_{\binom{n}{2}}$ on the pairwise distances between n points, is a function R from $S_{\binom{n}{2}}$ to the set of unweighted undirected graphs. Additionally, there is a subset of vertices $H \subseteq V(R(\pi))$ such that $|H| = n$ and for every no-stretch greedy embedding f of $R(\pi)$ the restriction to H of f is a monotone map for π.

Theorem 2.1. *The problem of finding a monotone map for a given order on the pairwise distances between n points cannot be reduced to a problem of finding a no-stretch greedy embedding of a graph on $o(e^n)$ vertices for $1 - o(1)$ fraction of orders.*

It is obvious that a monotone map for a graph metric is also a greedy embedding for this metric. On the other hand, there is an abundance of examples where the greedy embedding of a graph metric is not necessarily a monotone map. Perhaps the best such example is that of $K_{n,n}$. The complete bipartite graph $K_{n,n}$ has no monotone map into $\ell_2^{o(n)}$ while it has a no-stretch greedy embedding into ℓ_2^2: Position all vertices of $K_{n,n}$ uniformly along the unit circle S^1 while interlacing vertices from opposite sides of $K_{n,n}$. The techniques used in [15] to derive the $\Omega(n)$ bound on the monotone dimension of most graphs were borrowed from [16] and these techniques produce only trivial lower-bounds for the greedy embedding dimension.

Another related notion is that of a proximity embedding defined in [17]. A *proximity embedding* $f : G \to \ell_2^d$ of an undirected unweighted graph G is one for which $\forall v, w \in G : \ \|f(v) - f(w)\| < 1 \Leftrightarrow (v, w) \in G$. The *sphericity* of a graph is the minimum dimension d for which such embedding exists. As pointed out in [15], the sphericity of a graph is a lower-bound on its minimum monotone map dimension. Sphericity also bears only a weak connection to greedy embeddings. On the one hand, $K_{n,n}$ has sphericity $\Omega(n)$ while it embeds greedily into ℓ_2^2 with no stretch. On the other hand, trees and graphs of bounded degree have easy proximity embeddings in $\ell_2^{O(\log n)}$ using standard constructions involving the Johnson-Lindenstrauss lemma and only local graph structure considerations (see [18]). In contrast, Appendix A proves that standard techniques based on flattening or Bourgain's embedding cannot be used to construct greedy embeddings.

3. Preliminaries

3.1. Hyperbolic geometry

Hyperbolic geometry is a vast and complex area with applications in various branches of Mathematics. In this section we give a brief and somewhat self-sufficient introduction to the properties of hyperbolic spaces used in this paper. More comprehensive expositions can be found in the classical texts of Thurston [19] and Do Carmo [20]. A considerably more concise and self-contained introduction is the one by Katok [21], which we recommend for the beginner. Hyperbolic geometry has found little attention in Computer Science, but for a few notable exceptions [22, 1, 23].

Hyperbolic spaces, also known as Lobachevsky spaces (not to be confused with the more general Gromov δ-hyperbolic spaces [24, 25]; of course, Lobachevsky spaces are also Gromov $\log 3$-hyperbolic), can be constructed either axiomatically [26] (much like classical Euclidean geometry) or more explicitly using the language of Differential Geometry [20]. For the benefit of the reader's intuition we give the latter construction. We then state a few simple facts (while omitting proofs) which will enable us to reason about hyperbolic geometry in terms of its model via the more familiar Euclidean space.

3.1.1. The half-plane model

The d-dimensional real hyperbolic space, denoted $\mathbb{H}^d$, is modeled by the *upper-half plane* $\mathbb{R}^{d+} = \{(x_1, \ldots, x_d)^T \in \mathbb{R}^d \,|\, x_d > 0\}$ in $\mathbb{R}^d$ endowed with the Riemannian metric:

$$ds^2 = \frac{dx_1^2 + \cdots + dx_d^2}{x_d^2}$$

By construction $\mathbb{H}^d$ is geodesic. The Euclidean hyperplane $\partial\mathbb{H}^d = \{(x_1, \ldots, x_d)^T \in \mathbb{R}^d \,|\, x_d = 0\}$ plays a special role and is called the *boundary at infinity*. The following few facts establish the basic properties of $\mathbb{H}^d$.

Theorem 3.1 (See Proposition 3.1 in [20]). *Infinite geodesics, also called lines, in $\mathbb{H}^d$ (i.e. isometric maps of the form $g : \mathbb{R} \hookrightarrow \mathbb{H}^d$) correspond to Euclidean circles and lines orthogonal to the boundary at infinity and restricted to the upper half-plane, collectively referred to as generalized circles.*

Fact 3.1 (See [20] p.177). Hyperplanes in $\mathbb{H}^d$ (i.e. isometric maps $h : \mathbb{H}^{d-1} \hookrightarrow \mathbb{H}^d$) correspond to $(d-1)$-dimensional Euclidean spheres and planes orthogonal to the boundary at infinity and restricted to the upper half-plane, collectively called *generalized spheres*.

The isometries of $\mathbb{H}^d$ are modeled by conformal (unoriented-angle preserving) transformations of $\mathbb{R}^d$ that map the upper half-plane to itself, restricted to the upper half-plane. More notably though:

Theorem 3.2 (See Theorem 5.2 and Theorem 5.3 in [20]). *Let $f : \mathbb{H}^d \to \mathbb{H}^d$ be an isometry. Then, f is the restriction to $\mathbb{R}^{d+}$ of a composition of Euclidean isometries, dilations or inversions that map $\mathbb{R}^{d+}$ onto itself, at most one of each.*

We assume that the reader is already familiar with the isometries and dilations of $\mathbb{R}^{d+}$. An *inversion about a Euclidean hyperplane* is defined as Euclidean reflection with respect to that hyperplane. An *inversion* of a point p *about a Euclidean hypersphere* centered at c with radius r is defined as the unique point q on the ray $\overrightarrow{cp}$ for which $|cp| \cdot |cq| = r^2$. For convenience, we define shorthand notation for the Euclidean hemisphere $S_{c,r} = \{u \in \mathbb{R}^{d+} : \|u - c\| = r\}$ and the corresponding half-ball $B_{c,r} = \{u \in \mathbb{R}^{d+} \,|\, \|u - c\| < r\}$, where in both cases $c \in \partial\mathbb{H}^d$.

In this paper we use three specific isometries to construct greedy embeddings. Since we need to keep track of the bit complexity of point coordinates after application of isometric transformations, we give explicit formulas for them here:

- An *inversion* about a hyperbolic hyperplane corresponding to a Euclidean hemisphere $S_{c,r}$ is given by $\alpha_{c,r}(v) = (v - c) \cdot r^2 / \|v - c\|^2 - c$
- A *translation* by a vector $w \in \partial\mathbb{H}^d$ is given by $\beta_w(v) = v + w$
- A *dilation* at the origin by a factor $D > 0$ is given by $\gamma_D(v) = D \cdot v$

Finally, we will need an expression for the pairwise distance function of $\mathbb{H}^d$. We will denote the geodesic segment between points v and w in $\mathbb{H}^d$ by $[v, w]$. The Riemannian metric on $\mathbb{H}^d$ naturally induces a pairwise metric function $\rho(\cdot, \cdot)$. Let $\hbar = \{x \in \mathbb{R}^d \,|\, x_2 = \cdots = x_{d-1} = 0, x_d > 0\}$. We give an expression for $\rho(v, w)$ for the case when $v, w \in \mathbb{H}^d \cap \hbar$. (It should be clear that hyperbolic isometries can position any two points in this manner.) $\hbar$ can be viewed as a copy of $\mathbb{H}^2$, and v and w can be located in $\hbar$ using only two coordinates, namely the 1-st and the d-th. We shall now view v and w as complex numbers in the following way $v = v_1 + iv_d$ (similarly for w). With this notation in hand, the following theorem gives the pairwise distance between v and w:

Theorem 3.3 (See Theorem 1.2.6 in [21]). *Let $v, w \in \mathbb{H}^2$, then:*

$$\rho(v, w) = \ln \frac{|v - \overline{w}| + |v - w|}{|v - \overline{w}| - |v - w|}$$

3.1.2. The Klein model

The Klein model of hyperbolic space will be instrumental in our lower-bound proof. In the Klein model $\mathbb{H}^d$ is modeled by the d-dimensional Euclidean disc $D^d = \{x \in \mathbb{R}^d \mid \|x\| < 1\}$. In particular, the Klein model can be viewed as a homeomorphism $h : \mathbb{H}^d \to D^d$. We shall make use of one simple property of this model; for further information, we refer the interested reader to [19]:

Fact 3.2. Hyperbolic hyperplanes in the Klein model correspond to Euclidean hyperplanes restricted to the unit disc.

3.1.3. Bisecting hyperplanes

To prove correctness of our constructions, we will use a lemma from [1] for $\mathbb{H}^2$ whose proof translates to $\mathbb{H}^d$ unchanged:

Lemma 3.1. Let v and w be different points in $\mathbb{H}^d$, and let b be the hyperbolic hyperplane that bisects the geodesic $[v, w]$, then for all $u \in \mathbb{H}^d$ it holds that $\rho(v, u) < \rho(w, u)$ if and only if v and u are in the same half-space with respect to b .

In order to apply this lemma in our constructions, we will need to identify the bisecting hyperplane between vertices with equal d-th coordinates:

Lemma 3.2. Let $u, v \in \mathbb{H}^d$ such that $u_d = v_d$. Then the hyperbolic hyperplane bisecting $[u, v]$ coincides with the Euclidean hyperplane bisecting u and v (in Euclidean sense).

3.2. Distance-preserving embeddings

Definition 3.1. A map $f : X \to Y$, where (X, d_X) and (Y, d_Y) are metric spaces, is a distance-preserving embedding of X into Y with distortion $D > 0$, if there exists a constant $r > 0$ such that:

$$\forall\, v, w \in X \qquad r \cdot d_X(v, w) \le d_Y(f(v), f(w)) \le D \cdot r \cdot d_X(v, w)$$

3.3. Tree decomposition and heavy paths

This section describes a variant of the well-known caterpillar decomposition of trees [27, 28], also recognized as Tarjan and Harel's [29] heavy-path decomposition. Let T be an arbitrary unrooted tree on n vertices. A *path decomposition* of T into k paths is a collection of vertex-disjoint line subgraphs of T which covers T's vertices completely, i.e. $T = P_1 \uplus \cdots \uplus P_k$. A *hierarchical path decomposition* is a path decomposition which is additionally endowed with a hierarchical relationship among the paths. In particular, this relationship is represented by a rooted tree H whose vertex set is $P_1, \ldots, P_k$. Furthermore, (P_i, P_j) is an edge in H iff P_i and P_j are connected by an edge in T. A *heavy-path decomposition* of a (rooted/unrooted) tree is a particular hierarchical path decomposition which has depth at most $2\kappa(T) \le 2 \log L$, where L is the number of leaves of T. The quantity $\kappa(T)$ is the *caterpillar dimension* of T. It is easily verified that for an unrooted tree T, a caterpillar decomposition of T using an arbitrary root can be modified to produce a heavy-path decomposition of depth at most $2\kappa(T)$. A heavy-path decomposition of a bounded degree-3 tree is illustrated in Figure 1.

4. The unified lower bound

In this section we develop a "dual" representation of greedy embeddings in terms of "bisecting sets," which allows us to prove a unified lower bound on the Euclidean and Lobachevsky dimensionality of a certain family of graphs. The Lobachevsky bound is new, while the Euclidean was already known from [1]. Nevertheless, the unified framework of the proof seems to be of value.

We now define a family of graphs *with hard crossroads* that have rich combinatorial structure on the set of all-pairs shortest paths. Let $\mathcal{Q}_d$ be the graph on $n = d + 3 \cdot 2^d$ vertices and $m = 2^d \cdot (d+2)$ edges, defined as follows. The vertex set consists of $\{s_i\}_{i \in [2^d]}$, $\{t_j\}_{j \in [d]}$, $\{w_{i,q}\}_{i \in [2^d], q \in \{0,1\}}$. The edge set consist of two types of edges:

i. For every $i \in [2^d]$, include the edges $(s_i, w_{i,0})$ and $(s_i, w_{i,1})$.

ii. Let $i = b_1 b_2 \cdots b_d$ be the binary representation of i. Then for every $i \in [2^d]$, include the edges $(w_{i,b_1}, t_1), (w_{i,b_2}, t_2), \ldots, (w_{i,b_d}, t_d)$.

The main result of this section is the following theorem:

Theorem 4.1. *Every no-stretch greedy embedding of $\mathcal{Q}_d$ into Euclidean or Lobachevsky space requires d dimensions.*

This theorem implies a $\log(n)$ lower bound on the dimension of no-stretch greedy embeddings of graphs on n vertices.

Sketch of Proof: (See Appendix B for complete proof.) Let $f : \mathcal{Q}_d \to (X, d_X)$ be an embedding into a geodesic metric space with a continuous pairwise distance metric. We shall use v and $f(v)$ interchangeably. Let us now cover some topological preliminaries.

Let $a \neq b \in X$ be two different points in X. Define the *bisecting set* of $[a,b]$ to be $\mathsf{Bisect}(a,b) = \{c \in X \mid d_X(c,a) = d_X(c,b)\}$. In the spaces under consideration, every bisecting set is non-trivial and furthermore it separates the space X in at least two disjoint sets, called *chambers*, one for each endpoint of the bisected geodesic segment. We use the notation $(c\,{}^{a}|^{b}\,d)$ to indicate that c (respectively d) lies in the chamber of a (respectively b) with respect to the bisecting set of $[a,b]$. We also use that chambers are preserved by homeomorphisms. More generally, let $S \subset X$ separate X into a collection of chambers $\{C_\alpha\}_\alpha$ and let $h : X \to Y$ be a homeomorphism, then $h(S)$ separates Y exactly into $\{h(C_\alpha)\}_\alpha$.

It is easily seen that the correctness of f (as a no-stretch greedy embedding) can be expressed by a set of inequalities of the form $d_X(f(x), f(z)) < d_X(f(y), f(z))$ where $x, y, z \in V(\mathcal{Q}_d)$. Every such inequality implies a (weaker) *separation constraint* $(z\,{}^{x}|^{y}\,)$. The collection of separation constraints partitions X into a set of chambers and establishes combinatorial constraints regarding the position of $f(x)$, for every $x \in V(\mathcal{Q}_d)$, with respect to containment in chambers.

The idea of the proof is to find a homeomorphism h that sends X to $\mathbb{R}^d$ while mapping all bisecting sets (or at least subsets thereof) to hyperplanes. Then using Linear Algebra we establish that the required (by the separation constraints) geometric set system (See [30] for definition), formed by the points $\{(h \circ f)(x)\}_{x \in V(\mathcal{Q}_d)}$ and the hyperplanes $\{h(\mathsf{Bisect}(f(x), f(y)))\}_{x \neq y \in V(\mathcal{Q}_d)}$, cannot be realized in low dimensions.

When X is Lobachevsky, the bisecting sets are hyperbolic hyperplanes, and the Klein model is the required homeomorphism. When X is Euclidean the homeomorphism is simply the identity. ∎

It is an interesting (as far as we know open) question to find homeomorphisms that work for Minkowski normed spaces. Some thought will convince the reader that such homeomorphisms will have to be input-specific, unlike the universal Klein model homeomorphism for the Lobachevsky case. We also believe that this approach can be extended to nice classes of manifolds (but we do not dabble in this here).

The following theorem (whose proof is sketched in Appendix C) complements our lower-bound result:

Theorem 4.2. *If a graph G has a no-stretch greedy embedding into ℓ_2^d, then it has a no-stretch greedy embedding into $\mathbb{H}^{d+1}$.*

5. Concise Hyperbolic Embeddings of Trees

Theorem 5.1. *Every tree T on n vertices has a concise greedy embedding in $\mathbb{H}^3$ with $O(\kappa(T)\cdot\log n)$-bit vertex coordinates.*

Proof of Theorem: It is sufficient to exhibit embeddings for bounded degree-3 trees. This follows from the fact that if $T^* \supseteq T$ is a super-tree, then every greedy embedding of T^* restricts to a greedy embedding of T, and the fact that every tree T on n vertices is found as a subtree of a ternary tree of size no larger than $2n$.

We begin by obtaining a heavy-path decomposition $T = P_1 \uplus \cdots \uplus P_k$ with a hierarchical relationship on the P_j's represented by a tree H (as in Section 3.3 and Figure 1).

Some notation is due now. Let P_j be a path in T, viewed as a vertex in H, and let $\mathsf{parent}(P_j)$ denote its parent path in H (if it exists). Denote by $\mathsf{apex}(P_j)$ be the unique vertex in $\mathsf{parent}(P_j)$ that P_j connects to, and by $\mathsf{exit}(P_j)$ the unique vertex in P_j that connects to $\mathsf{apex}(P_j)$. Let $\mathsf{subtree}(\mathsf{apex}(P_j))$ denote the subtree of T consisting of P_j, all of its descendants in H, as well as $\mathsf{apex}(P_j)$. For a vertex v in P_j that is not the apex of any P_i we will let $\mathsf{subtree}(v)$ denote the singleton subtree of T consisting of v itself.

A *canonical embedding* $f_v : \mathsf{subtree}(v) \to \mathbb{H}^3$ of $\mathsf{subtree}(v)$, where $v \in T$, is one for which all relevant vertices are embedded in the interior of $B_{(0,0)^T,1}$, and v is embedded at the unique location inside the ray e_z such that $\rho(f_v(v), S_{(0,0)^T,1}) = \alpha$, where α is any fixed positive real number for which $e^\alpha \in \mathbb{Q}$. (Later we will see that we can also use $\alpha = 1$.) We will describe a recursive (on H) procedure that canonically embeds each $\mathsf{subtree}(v)$ until all of T is embedded.

In the base case, canonically embedding a single vertex v is trivial. We simply embed v at the unique point on the ray e_z that has hyperbolic distance to $S_{(0,0)^T,1}$ equal to α and is "inside" $S_{(0,0)^T,1}$ (i.e. on the same side as the origin). Explicitly $f_v(v) = (0,0,1/e^\alpha)^T$. The bit complexity of this embedding is $O(1)$ due to our choice of α. We should note however that since the rest of the embedding will be obtained via isometric transformations, we can view the quantity $1/e^\alpha$ as an irreducible (or free) variable and describe all coordinates as polynomials over it. Either approach works.

Let us now proceed to the recursive step of embedding w where $w = \mathsf{apex}(P_j)$ for some P_j consisting of vertices $v_1, \ldots, v_k$. And let $\mathsf{exit}(P_j) = v_q$ for some $q \in [k]$. From the recursion, we have embeddings $f_{v_1}, \ldots, f_{v_k}$ with $f_{v_i} : \mathsf{subtree}(v_i) \to B_{(0,0)^T,1}$. We shall first define an embedding $g_w : \mathsf{subtree}(w) \to \mathbb{H}^3$ which is not canonical. Later we will transform g_w into a canonical one:

$$g_w(u) = \begin{cases} \left(\beta_{(i-q,0)^T} \circ f_{v_i}\right)(u), & \text{if } u \in \mathsf{subtree}(v_i) \\ (0,1,1/e^\alpha)^T, & \text{otherwise, i.e. if } u = w \end{cases} \tag{5.1}$$

The embedding $g_w(u)$ is illustrated in Figure 2. Our first order of business will be to check that it is correct. Afterwards we will apply the necessary isometric transformations to reshape it into canonical form. Notice that isometric transformations do not violate correctness.

To check correctness (i.e. that greedy routing works), we have to identify a subregion $R \subset \mathbb{H}^3$ where we plan to position the rest of T, i.e. $T \backslash \mathsf{subtree}(w)$, later in the recursion. We will let $R = B_{(0,1)^T, 1/e^{2\alpha}}$. Note that R is intentionally chosen so that $\rho(g_w(w), R) = \alpha$.

The inductive (as in recursive) hypothesis is that all points in $\mathsf{subtree}(v_i)$ for $i \in [k]$ are embedded in such a way (by f_{v_i}) that (a) greedy routing works among themselves, and (b) if routing is attempted to any location outside of $B_{(0,0)^T,1}$, it will reach v_i. Therefore our task is to check that under g_w:

i. Routing from any v_i to $u \in \mathsf{subtree}(v_j)$ reaches destination, and

ii. Routing from any v_i to w or any location inside R reaches destination.

To prove the first part, it is sufficient to show that routing to $u \in \mathsf{subtree}(v_j)$ reaches at least v_j, then by inductive hypothesis we know that u will be reached under f_{v_j} (translated by $\beta_{(i-q,0)^T}$). Assume without loss of generality that $i < j$. Indeed, since $g_w(v_i)$ and $g_w(v_{i+1})$ have the same z coordinate, their bisecting hyperplane separates $g_w(v_i)$ from all points $g_w(u)$ where $u \in \mathsf{subtree}(v_j)$. Therefore routing will progress to v_{i+1}. The second assertion is also easy to check. In particular, one simply verifies that at a vertex v_i the bisecting plane of the edge that leads to w separates v_i from w and all of R. This is illustrated in Figure 2 where bisecting hyperbolic hyperplanes are pictured as dotted lines.

The canonical embedding f_w is obtained from g_w by:

i. Applying a spherical hyperbolic inversion with respect to R. (This transformation takes all of g_w and "squeezes" it inside R.)

ii. Translating R to R' so that R' is centered at the origin

iii. Isometrically expanding R' to $R'' = S_{(0,0)^T,1}$

Formally, $f_w = \left(\gamma_{e^{2\alpha}}\right) \circ \left(\beta_{(0,-1)^T}\right) \circ \left(\alpha_{(0,1)^T,1/e^{2\alpha}}\right) \circ \left(g_w\right)$.

To calculate the bit-description complexity per vertex, we will trace out what happens to a vertex's coordinates throughout the recursion. At the lowest level of the recursion, a vertex starts off with $O(1)$-bit coordinates (namely $(0, 0, 1/e^{\alpha})^T$). At each level of the recursion, the vertex is translated by at most n positions along the x axis. This step adds at most $O(\log n)$ bits to its x-coordinate. Observe that the canonization step is a fixed isometric transformation, so it contributes $O(1)$ additional bits. There are $\kappa(T)$ recursive levels, amounting to a total of $O(\kappa(T) \cdot \log n)$ bits per vertex coordinate. ∎

We will briefly note (without proof) that since $\mathbb{H}^d$ is Gromov $(\log 3)$-hyperbolic (for every $d \geq 2$), if we scale our embedding procedure so that the hyperbolic distance between vertices sharing an edge is Δ, then the greedy embedding is also a distance-preserving embedding (in the sense of Definition 3.1) with distortion $1 + \log 3/\Delta$.

The techniques described in this section can be used to embed slightly more general classes of graphs. In particular, let G be a graph that can be decomposed into a vertex-disjoint family of subgraphs, i.e. $G = H_1 \uplus \cdots \uplus H_k$. Let G^* be a graph with a vertex set $[k]$ where $(i, j) \in E(G^*)$

iff there is an edge in G between H_i and H_j. Then if G^* is a tree and each H_i can be embedded canonically with no stretch, all of G can be embedded canonically with no stretch.

It is easily seen, for example, that graphs that can be decomposed into lines and *cycles* succumb to the same embedding procedure. The canonical embedding of a cycle is illustrated in Figure 3. More complicated examples can be derived by using higher hyperbolic dimension and/or a cleverer arrangement of the canonical embeddings from lower levels of the recursion. The limitation of this technique, however, is that it is inherently recursive and therefore it applies to graphs that at large scale look like trees.

6. Low dimensional Euclidean embeddings of trees

Inspired by ideas from [3], in this section we construct low dimensional greedy embeddings of trees into Euclidean spaces. Note that our construction is somewhat different than Gupta's and surprisingly it does not require the use of a hierarchical path decomposition to accomplish conciseness.

Theorem 6.1. *Every tree T on n vertices has a concise greedy embedding in $\ell_2^{O(\log n)}$ with $O(\log^2 n)$-bit vertex coordinates.*

Sketch of Construction: We begin by picking an arbitrary root v_0 for T. Using the Johnson-Lindenstrauss lemma, or alternatively using sphere packing constructions as in [3, 31], we obtain a bundle $\mathcal{B}$ of $n-1$ unit vectors such that (a) each vector has positive 1-st coordinate; (b) the angle between any two vectors is a constant slightly larger than $\pi/3$, say $\pi/3 + \pi/180$; and (c) $\mathcal{B}$ is realized in $\ell_2^{O(\log n)}$.

The embedding algorithm assigns a vector $g(v, w) \in \mathcal{B}$ to each edge $(v, w) \in T$ in a manner to be specified shortly. The embedding $f : V(T) \to \ell_2$ is then defined as $f(v) = g(v_0, v_1) + \cdots + g(v_{k-1}, v)$, where $v_0, v_1, \ldots, v_{k-1}, v$ is the path from v_0 to v in T. The matching $g : E(T) \to \mathcal{B}$ is chosen as follows. For vertex $v \in T$ let $\tilde{g}(v) = \{g(u, w) \in \mathcal{B} : (u, w) \in \mathsf{subtree}(v)\}$. Then $g(\cdot, \cdot)$ is such that for every $v \in T$ and all pairs of children v_a and v_b of v it holds that $\mathsf{Cone}\big(\tilde{g}(v_a)\big) \cap \mathsf{Cone}\big(\tilde{g}(v_b)\big) = \{\mathbf{0}\}$. Such a matching exists and can be found algorithmically using the sweeping-hyperplane method of [3].

The correctness of this construction is sketched in Appendix C. ∎

7. Open Problems and Closing Remarks

An abundance of open problems arises from the notion of greedy embeddings and their applications. We only mention two.

The main open problem is that of finding no-stretch greedy embeddings of any graph into $\ell_2^{O(\log n)}$. We shortly describe a promising strategy for attacking this problem, which we have not yet investigated thoroughly. Let $\mathcal{Y}_{p,q,r}$ be the graph consisting of edge-disjoint copies of L_p, L_q and L_r (where L_l is the undirected line graph on l edges), exactly one of each, where also all three line subgraphs share a starting vertex and a (different) ending vertex. Let a *gadget* be a procedure for embedding $\mathcal{Y}_{p,q,r}$ into $\ell_2^{O(1)}$ for any p, q and r. We believe that using such a gadget in conjunction with dimensionality reduction, can lead to the desired embeddings of arbitrary graphs. Our intuition is based on the following lemma (proof deferred to full version):

Lemma 7.1. Every unweighted undirected graph can be decomposed into a collection of (not necessarily disjoint) sub-trees and (irreducible) sub-cycles such that (i) the shortest paths between vertices on a sub-cycle lie entirely in the sub-cycle, and (ii) the intersection of any two sub-cycles is a connected arc, a vertex, or the empty set.

In view of applications, one other particularly important question concerns the existence of algorithms for finding greedy embeddings in Peleg's message-passing model of distributed network computation [32], where message-cost (in addition to time) is of central importance. Furthermore, algorithms with good incremental properties and resilience to small changes in the input graph are desired.

References

[1] R. Kleinberg, "Geographic routing using hyperbolic space," *To appear in Proc. 32nd IEEE INFOCOM (INFOCOM 2007)*, 2007.

[2] Papadimitriou and Ratajczak, "On a conjecture related to geometric routing," *TCS: Theoretical Computer Science*, vol. 345, 2005.

[3] A. Gupta, "Embedding tree metrics into low dimensional euclidean spaces," in *STOC*, pp. 694–700, 1999.

[4] B. Karp and H. T. Kung, "GPSR: greedy perimeter stateless routing for wireless networks," in *MOBICOM*, pp. 243–254, 2000.

[5] P. Bose, P. Morin, I. Stojmenovic, and J. Urrutia, "Routing with guaranteed delivery in ad hoc wireless networks," *Wireless Networks*, vol. 7, no. 6, pp. 609–616, 2001.

[6] J. Gao, L. J. Guibas, J. Hershberger, L. Zhang, and A. Zhu, "Geometric spanner for routing in mobile networks," in *MobiHoc '01: Proceedings of the 2nd ACM international symposium on Mobile ad hoc networking & computing*, (New York, NY, USA), pp. 45–55, ACM Press, 2001.

[7] F. Kuhn, R. Wattenhofer, Y. Zhang, and A. Zollinger, "Geometric ad-hoc routing: of theory and practice," in *PODC*, pp. 63–72, 2003.

[8] Bollobas, Riordan, Spencer, and Tusnady, "The degree sequence of a scale-free random graph process," *RSA: Random Structures and Algorithms*, vol. 18, 2001.

[9] B. Bollobás and O. Riordan, "The diameter of a scale-free random graph," *Combinatorica*, 2002. In press.

[10] A. Rao, C. Papadimitriou, S. Shenker, and I. Stoica, "Geographic routing without location information," in *MobiCom '03: Proceedings of the 9th annual international conference on Mobile computing and networking*, (New York, NY, USA), pp. 96–108, ACM Press, 2003.

[11] R. Fonseca, S. Ratnasamy, D. Culler, S. Shenker, and I. Stoica, "Beacon vector routing: Scalable point-to-point in wireless sensornets," July 09 2004.

[12] A. Caruso, S. Chessa, S. De, and A. Urpi, "GPS free coordinate assignment and routing in wireless sensor networks," in *INFOCOM*, pp. 150–160, IEEE, 2005.

[13] R. Dhandapani, "Greedy drawings of triangulations," *Submitted*, 2006.

[14] Alon, Badoiu, Demaine, Farach-Colton, Hajiaghayi, and Sidiropoulos, "Ordinal embeddings of minimum relaxation: General properties, trees, and ultrametrics," in *SODA: ACM-SIAM Symposium on Discrete Algorithms (A Conference on Theoretical and Experimental Analysis of Discrete Algorithms)*, 2005.

[15] Bilu and Linial, "Monotone maps, sphericity and bounded second eigenvalue," *JCTB: Journal of Combinatorial Theory, Series B*, vol. 95, 2005.

[16] N. Alon, P. Frankl, and V. Rödl, "Geometrical realization of set systems and probabilistic communication complexity," in *FOCS*, pp. 277–280, IEEE, 1985.

[17] H. Maehara, "Space graphs and sphericity," *Discrete Appl. Math.*, vol. 7, pp. 55–64, 1984.

[18] P. Frankl and H. Maehara, "The johnson-lindenstrauss lemma and the sphericity of some graphs," *J. Comb. Theory, Ser. B*, vol. 44, no. 3, pp. 355–362, 1988.

[19] W. P. Thurston, *Three-Dimensional Geometry and Topology*, vol. 35 of *Princeton Mathematical Series*. Princeton University Press, 1997.

[20] Do Carmo, Manfredo, *Riemannian Geometry*. Boston: Birkhäuser Verlag, 1992.

[21] S. Katok, *Fuchsian groups*. Chicago Lectures in Mathematics, Chicago, IL: University of Chicago Press, 1992.

[22] Sleator, Tarjan, and Thurston, "Rotation distance, triangulations, and hyperbolic geometry," *JAMS: Journal of the American Mathematical Society*, vol. 1, 1988.

[23] R. Krauthgamer and J. Lee, "Algorithms on negatively curved spaces," in *FOCS*, 2006.

[24] J. M. Alonso and et al., "Notes on word hyperbolic groups," in *Group theory from a geometrical viewpoint (Trieste, 1990)*, pp. 3–63, World Sci. Publ., River Edge, NJ, 1991. Edited by H. Short.

[25] M. Gromov, "Hyperbolic groups," in *Essays in group theory*, vol. 8 of *Math. Sci. Res. Inst. Publ.*, pp. 75–263, New York: Springer, 1987.

[26] H. S. M. Coxeter, *Non-Euclidean geometry*. MAA Spectrum, Washington, DC: Mathematical Association of America, sixth ed., 1998.

[27] Linial, Magen, and Saks, "Trees and euclidean metrics," in *STOC: ACM Symposium on Theory of Computing (STOC)*, 1998.

[28] J. Matousek, "On embedding trees into uniformly convex banach spaces," 1999.

[29] Harel and Tarjan, "Fast algorithms for finding nearest common ancestors," *SICOMP: SIAM Journal on Computing*, vol. 13, 1984.

[30] J. Matousek, "Geometric computation and the art of sampling," in *IEEE Symposium on Foundations of Computer Science*, p. 2, 1998.

[31] J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices and groups*, vol. 290 of *Grundlehren der math. Wissenschaften*. Springer, 1993.

[32] D. Peleg, *Distributed computing: a locality-sensitive approach*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2000.

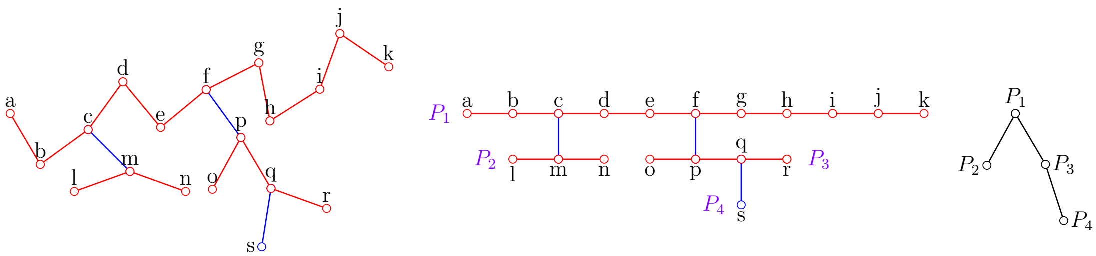

Figure 1. Shown: (a) an unrooted tree, (b) its heavy-path decomposition, and (c) the hierarchical path relationship.

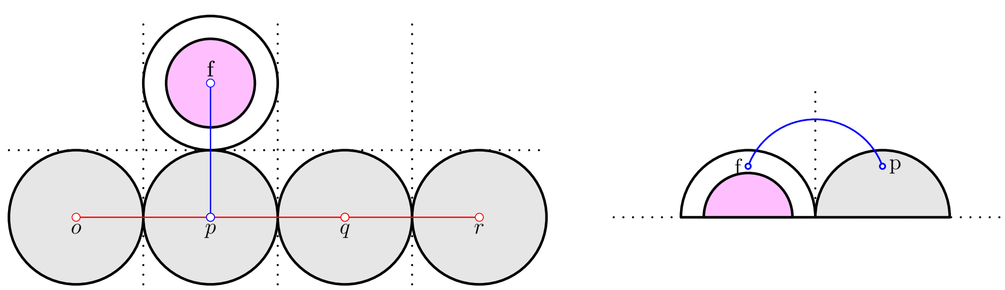

Figure 2. Illustrated is the g_f embedding of P_3 from Figure 1. On the left is a view from the z-axis looking down towards the origin. On the right is a planar section defined by $g_f(f), g_f(p)$ and the origin.

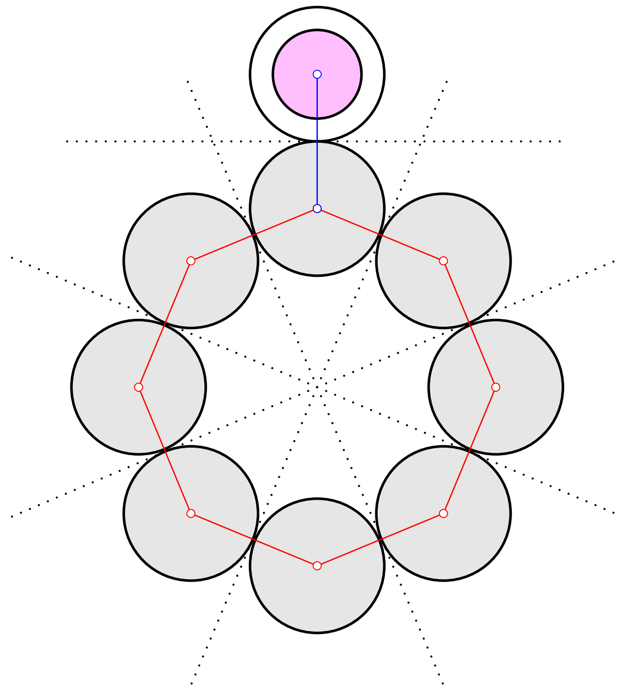

Figure 3. An illustration of the canonical embedding of a cycle.

A. Inappicability of standard techniques

Finding low-dimensional greedy embeddings into ℓ_2 is hard: In this section we explain that standard dimensionality reduction techniques alone are of no use for constructing greedy embeddings. When seeking low-dimensional Euclidean embeddings it is common to use one of the following two approaches:

- *The Bourgain approach:* Find an embedding into an arbitrary metric that realizes the required embedding properties. Then squeeze this embedding into $\ell_2^{O(\log n)}$ using Bourgain's $O(\log n)$-distortion embedding, while "making sure that the necessary embedding properties are preserved." Alternatively,

- *The Johnson-Lindenstrauss approach:* Find an embedding into ℓ_2 (of arbitrary dimension) that realizes the required embedding properties. Then reduce the dimension using Johnson-Lindenstrauss Flattening Lemma, while again "making sure that the necessary embedding properties are preserved."

In the case of greedy embeddings "making sure that the necessary embedding properties are preserved" boils down to requiring that pairwise distances whose relative magnitudes must be preserved by the embedding have a sufficiently large margin ϵ. The minimum margins are $\epsilon = O(\log n)$ and $\epsilon = 1/\log^{O(1)} n$, respectively, for the above two approaches.

We are going to show that neither of these requirements are achievable for almost any graph G. Start with two technical observations:

Lemma A.1. Let $s = v_1, v_2, \ldots, v_{k+1} = t$ be the unique shortest path between s and t in a graph

G. Let $f : (G, d_G) \to (X, d_X)$ be a no-stretch greedy embedding with margin ϵ, and also let $x_i = d_X(f(v_i), f(v_{i+1}))$. Then $x_1 + \cdots + x_k \geq (1+\epsilon)^{k-1} \cdot \max\{x_1, \ldots, x_k\}$

Sketch of Proof: The margin requirement says that for any three vertices u, v and w such that u and v are adjacent and v is on the unique shortest path from u to w, it must hold that $d_X(f(u), f(w)) > (1+\epsilon)d_X(f(v), f(w))$. Observe that for any $1 \leq i < j \leq k$ the unique shortest-path between v_i and v_j is $v_i, v_{i+1}, \ldots, v_j$. Now the lemma follows by induction. ∎

Corollary A.1. *Using the setup from the previous lemma, the following must hold* $k \geq (1+\epsilon)^{k-1}$

Proof. Let $x_{\max} = \max\{x_1, \ldots, x_k\}$, then simply: $k \cdot x_{\max} \geq x_1 + \cdots + x_k \geq (1+\epsilon)^{k-1} \cdot x_{\max}$. ∎

When we substitute for the margin ϵ in the above corollary we get:

Corollary A.2. *No graph that has a unique shortest path of length* $O(1)$*, respectively* $\log^{O(1)} n$*, can be greedily embedded in* $\ell_2^{\text{polylog}(n)}$ *using the Bourgain approach, respectively the Johnson-Lindenstrauss approach.*

In other words, both approaches are futile for almost every graph, and in particular for nice classes of graphs like trees, cycles, random graphs, etc.

B Proof of lower bound

B.1 Dual Representation

Let $G = (V, E)$ be a graph and let $f : V \to (X, \|\cdot\|)$ be an embedding, where $(X, \|\cdot\|)$ is one of ℓ_2^d or $\mathbb{H}^d$ and d is referred to as the *dimension of* X. The conditions for f being a no-stretch greedy embedding of G are described by a collection of inequalities, heretofore called *greedy constraints*, of the form:

$$\|f(x) - f(z)\| < \|f(y) - f(z)\| \tag{B.1}$$

where $x, y, z \in V$ are pairwise unequal. When the host metric is Euclidean or Lobachevsky, we can rephrase the constraints using the language of hyperplanes. We shall use the notation $(a\,{}^c|^d\,b)$, where a, b, c and d are points in the host metric space, to mean that the bisecting hyperplane of $[c, d]$ separates a, c and b, d. More generally, in this notation we allow arbitrary lists (including the empty list) of points in place of a or b. Furthermore, we abuse notation a little by using v to refer both to a vertex $v \in V$ and its image $f(v)$. It is now easily seen that the constraint (B.1) can be rewritten in the form $(z\,{}^x|^y)$. (The hyperbolic case follows from Lemma 3.1.)

B.2 Proof Outline

The idea of the proof is to examine the constraints of $\mathcal{Q}_d$ and show that when X is low-dimensional there is no configuration of points and *any* hyperplanes that satisfy the constraints. What makes this proof manageable is that we seek to realize the greedy constraints using arbitrary hyperplanes, rather than strictly bisecting ones. In order to unify our analysis of the normed and Lobachevsky cases, we make the following observation. In the Klein model of $\mathbb{H}^d$, the hyperbolic hyperplanes correspond to Euclidean hyperplanes restricted to the unit disc $D^d = \{x \in \mathbb{R}^d : \|x\| < 1\}$ in $\mathbb{R}^d$. Therefore for both types of geometries it suffices to show that the hyperplane/point configurations required by $\mathcal{Q}_d$ cannot be realized in $\ell_2^{d'}$ with $d' < d$. This will be established using simple Linear Algebra.

B.3 Linear Algebra and Point/Hyperplane Configurations

We shall only concern ourselves now with Euclidean geometry. If M is a matrix, we will let m_i, M_j and $M_{i,j}$ refer to the i-th row, j-th column and the (i,j)-th entry of M, respectively.

A *(point/hyperplane) configuration* Ψ is a collection of points $V_1, \ldots, V_n \in \mathbb{R}^d$ and hyperplanes $(a_1^T, b_1), \ldots, (a_k^T, b_k) \in \mathbb{R}^d \times \mathbb{R}$, where a point V_j is on the "positive" side of (a_i^T, b_i) iff $a_i V_j - b_i > 0$. The left-hand side of the latter inequality is referred to as the *polarity* of V_j with respect to (a_i^T, b_i).

Let $A \in M_{k,d}(\mathbb{R})$ be the matrix whose rows are $a_1, \ldots, a_k$, $V \in M_{d,n}(\mathbb{R})$ be the matrix whose columns are $V_1, \ldots, V_n$, and $b \in \mathbb{R}^k$ be the vector whose entries are $b_1, \ldots, b_k$. Define the *signature* of Ψ to be the matrix $\chi(\Psi) = AV - b \cdot \mathbf{1}^T$, where $\mathbf{1}^T = (1, \ldots, 1) \in \mathbb{R}^d$. Observe that $\chi(\Psi)$ has a natural interpretation; in particular, $\mathsf{sign}(\chi(\Psi)_{i,j})$ indicates the polarity of V_j with respect to (a_i^T, b_i). Furthermore, $\chi(\Psi)$ can be interpreted as a configuration where the points are represented by the columns of $\chi(\Psi)$ and the hyperplanes are the canonical Euclidean hyperplanes through the origin, orthogonal to the unit vectors e_i in the i-th direction. In that sense, $\chi(\Psi)$ is a "straighten-out" version of Ψ which is more amenable to dimension analysis. In the rest of the proof, we will make use of the following property of $\chi(\Psi)$:

Lemma B.1 (See Section C)**.** For Ψ as above, $\mathsf{dim}\left(\mathsf{span}(V_1, \ldots, V_n)\right) \geq \mathsf{rank}\left(\chi(\Psi)\right) - 1$.

For every greedy embedding of a graph G, we can can view the image of $V(G)$ and the corresponding bisecting hyperplanes between all pairs of points as a configuration. The greedy constraints will impose certain sign-constraints on the entries of the signature of this configuration. In the case of $\mathcal{Q}_d$, these sign-constraints will help us derive a lower bound on the rank of the signature and hence on the dimensionality of the embedding.

B.4 The Constraints of $\mathcal{Q}_d$

It is easily checked that the following is a subset of the no-stretch greedy constraints of $\mathcal{Q}_n$, involved along the routes from s_i, for $i \in [2^d]$, to t_j, for $j \in [d]$. Let $i = b_1 b_2 \cdots b_d$ be the binary representation of i and let $\pi \in S_d$ be a permutation such that $b_{\pi(1)} = \cdots = b_{\pi(q)} = 0$ and $b_{\pi(q+1)} = \cdots = b_{\pi(d)} = 1$ where $0 \leq q \leq d$. The constraints are:

$$\forall i \in [2^d] \qquad (t_{\pi(1)}, \ldots, t_{\pi(q)} \,{}^{w_{i,0}}|^{w_{i,1}}\, t_{\pi(q+1)}, \ldots, t_{\pi(d)}) \tag{B.2}$$

Let Ψ be the configuration corresponding to a no-stretch greedy embedding of $\mathcal{Q}_d$. As noted earlier, the rows of $\chi(\Psi)$ correspond to (and are indexed by) the bisecting hyperplanes, and the columns correspond to (and are indexed by) the vertices of $\mathcal{Q}_d$. Let $C \in M_{2^d,d}(\mathbb{R})$ be the sub-matrix of $\chi(\Psi)$ defined by the hyperplanes (rows) that appear in (B.2) and the vertices $\{t_j\}_{j\in[d]}$. It is clear that $\mathsf{rank}(\chi(\Psi)) \geq \mathsf{rank}(C)$. Next, we are going to show that $\mathsf{rank}(C) = d$. This will imply $\mathsf{rank}(\chi(\Psi)) \geq d$, and by Lemma B.1 we will get the desired lower bound $d = \Omega(\log |V(\mathcal{Q}_d)|)$.

B.5 Rank of the signature

The constraints of (B.2) impose that the set of d-tuples

$$\left\{(\mathsf{sign}(\sigma \cdot C_{i,1}), \mathsf{sign}(\sigma \cdot C_{i,2}), \ldots, \mathsf{sign}(\sigma \cdot C_{i,d}))\ :\ i \in [2^d], \sigma \in \{-1, +1\}\right\}$$

contains all 2^d sign patters on d slots. Then the following lemma (whose proof is found in Appendix C) implies that $\mathsf{rank}(C) = d$:

Lemma B.2. Let $C \in M_{2^d,d}(\mathbb{R})$ be a matrix whose rows realize all 2^d sign patterns over d columns, then $\mathsf{rank}(C) = d$.

C. Proofs and sketches thereof

Proof of Lemma B.1. Note that $\mathsf{rank}(b \cdot \mathbf{1}^T) \in \{0, 1\}$, then:

$$\begin{aligned}
\mathsf{dim}(\mathsf{span}(V_1, \dots, V_n)) &= \mathsf{rank}(V) \\
&\geq \mathsf{rank}(AV) \\
&= \mathsf{rank}(\chi(\Psi) + b \cdot \mathbf{1}^T) \\
&\geq |\,\mathsf{rank}(\chi(\Psi)) - \mathsf{rank}(b \cdot \mathbf{1}^T)| \\
&\geq \mathsf{rank}(\chi(\Psi)) - 1
\end{aligned}$$

∎

Proof of Lemma B.2. Induct on d. The base case $d = 1$ is straightforward. Without loss of generality let $C \in M_{2^d,d}(\mathbb{R})$ be such that $\mathsf{sign}(C_{i,j}) = \mathsf{sign}(b_{i,j} - 1/2)$, where $b_{i,j}$ is the j-th bit in the binary representation of i. Let U be the sub-matrix of C consisting of the first 2^{d-1} rows. From the induction hypothesis, U has rank $d - 1$. Let $U' \in M_{d-1,d}(\mathbb{R})$ be a diagonalized version of U. In particular:

i. $U'_{i,i+1} = 1$ for $i \in [d - 1]$,

ii. $U'_{i,j+1} = 0$ for $i \neq j \in [d - 1]$, and

Pictorially:

$$U' = \begin{pmatrix} U_{1,1} & 1 & 0 & \cdots & 0 \\ U_{2,1} & 0 & 1 & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ U_{d-1,1} & 0 & \cdots & 0 & 1 \end{pmatrix}$$

By definition, C must have a row c_l where $l \in [2^d] \backslash [2^{d-1}]$ such that:

i. $\mathsf{sign}(C_{l,1}) = +$, and

ii. $\mathsf{sign}(C_{l,j+1}) = \mathsf{sign}(-U_{j,1})$ for all $j \in [d - 1]$; if $U_{j,1} = 0$ then $\mathsf{sign}(C_{l,j+1})$ can be arbitrary.

It is now easily verified that c_l is linearly independent from all rows in U' thereby proving the inductive step. ∎

Sketch of proof of Theorem 4.2. Given a greedy embedding f of G into ℓ_2^d, the set $f(G)$ can be embedded onto the d-dimensional unit sphere $S^d \in \mathbb{R}^{d+1}$ such that the relative distance between all pairs of points is preserved. Let $g : V(G) \to S^d$ be this embedding. The bisecting hyperplanes between all pairs of points in $g(G)$ in $\mathbb{R}^{d+1}$ go through the origin. The disc $2D^d = \{x \in \mathbb{R}^{d+1} : \|x\| < 2\}$ together with $g(G)$ inside it, can be interpreted as an embedding of G into $\mathbb{H}^d$ via the Klein model. This is the required embedding. ∎

Sketch of correctness for Theorem 6.1. For any $v \in \mathcal{B}$ the bisecting hyperplane of v separates v from all other vectors in $\mathcal{B}$. Using this, and the fact that any two vectors in $\mathcal{B}$ form an angle of roughly $\pi/3 + \pi/180$, one can show correctness by induction. The induction is guided by the "growing" process of creating the tree, similarly to the one in [3]. ∎