Metric Embedding, Hyperbolic Space, and Social Networks*

Kevin Verbeek† Subhash Suri†

ABSTRACT

We consider the problem of embedding an undirected graph into hyperbolic space with minimum distortion. A fundamental problem in its own right, it has also drawn a great deal of interest from applied communities interested in empirical analysis of large-scale graphs. In this paper, we establish a connection between distortion and *quasi-cyclicity* of graphs, and use it to derive lower and upper bounds on metric distortion. Two particularly simple and natural graphs with large quasi-cyclicity are n-node cycles and $n \times n$ square lattices, and our lower bound shows that any hyperbolic-space embedding of these graphs incurs a multiplicative distortion of at least $\Omega(n/\log n)$. This is in sharp contrast to Euclidean space, where both of these graphs can be embedded with only *constant* multiplicative distortion. We also establish a relation between quasi-cyclicity and δ-hyperbolicity of a graph as a way to prove upper bounds on the distortion. Using this relation, we show that graphs with small quasi-cyclicity can be embedded into hyperbolic space with only constant *additive* distortion. Finally, we also present an efficient (linear-time) randomized algorithm for embedding a graph with small quasi-cyclicity into hyperbolic space, so that with high probability at least a $(1-\varepsilon)$ fraction of the node-pairs has only constant additive distortion. Our results also give a plausible theoretical explanation for why social networks have been observed to embed well into hyperbolic space: they tend to have small quasi-cyclicity.

1. INTRODUCTION

Metric embeddings have been a topic of intense research in many areas of computer science, including geometry, theory, databases and machine learning, during the past few decades. Besides their fundamental nature and rich mathematical content, these embeddings have also become a major tool in the analysis of high-dimensional data: the general idea is to embed a high-dimensional metric space into a much smaller-dimensional space without creating large distortion in pair-wise distances. A celebrated result in this area is the Johnson-Lindenstrauss Lemma [7], showing that metrics induced by n points in an Euclidean space can be embedded with $(1+\varepsilon)$ distortion into $O(\varepsilon^{-2}\log n)$-dimensional space. The reader may consult the book by Matoušek [13] for an excellent survey and exposition to many elegant results on metric embeddings.

In a different context, researchers in the networking communities have also explored geometric embeddings for network measurements and geographic routing [8, 17, 18, 20]. A number of network services use shortest path distances to implement application level multicast trees or routes [8, 14, 17]. While maintaining the full matrix of pairwise distances is infeasible in Internet scale networks, a geometric embedding gives a highly compact representation of those distances. Similarly, virtual coordinates and geographic routing are an elegant way to avoid the need for full network topology at each network router. There is also a great deal of interest in estimating pair-wise distances in *social* networks, as a means to measure influence, to recommend friends, and to show selective profiles, but running all-pair shortest path algorithms is infeasible at their scale. Fortunately, a number of empirical studies have shown that simple geometric embeddings are good enough for estimating (most) pair-wise network distances in large graphs [19, 20]. The embedding dimension is typically small (e.g. 10), and therefore the distance between any two nodes can be calculated in constant time. Interestingly, these experimental results also show that *hyperbolic-space* embeddings lead to higher accuracy (lower distortion) than Euclidean space [18, 20]. Kleinberg's well-known result [8] that every connected finite graph has a greedy embedding into the *hyperbolic plane*, while many graphs do not admit such an embedding in the Euclidean plane, also underscores the relative advantage of hyperbolic space. Some researchers have even argued that the "hidden space" of the Internet and social networks is naturally the negatively curved hyperbolic space [12].

Our work is a theoretical counterpart to this line of empirical research. We are interested both in establishing the worst-case distortion bounds for hyperbolic space embedding in general, and in discovering useful characterizations of graphs that may explain the observed low distortion in hyperbolic embedding of large-scale social networks. We begin with some formal definitions necessary to formulate the problem and state our results.

*This research was partially supported by the National Science Foundation Grant CNS-1035917 and by DARPA GRAPHS (BAA-12-01).

†Department of Computer Science, University of California Santa Barbara [kverbeek|suri]@cs.ucsb.edu

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
SoCG'14, June 8–11, 2014, Kyoto, Japan.
Copyright 2014 ACM 978-1-4503-2594-3/14/06 ...$15.00.

1.1 Metric Spaces and Embeddings

A *metric space* is a tuple $(\mathcal{X}, d)$ where $\mathcal{X}$ is the underlying space and $d\colon \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ is the metric (or distance function) measuring the distance between two elements of $\mathcal{X}$. A *metric embedding* (or simply *embedding*) of $(\mathcal{X}_1, d_1)$ into $(\mathcal{X}_2, d_2)$ is a map $f\colon \mathcal{X}_1 \to \mathcal{X}_2$. In this paper, we are primarily concerned with embedding the *graph metric* (V, d_G) of an undirected graph $G = (V, E)$ into k-dimensional hyperbolic space $(\mathbb{H}^k, d_{\mathbb{H}})$, where d_G measures the length of the shortest path between two nodes in G, and $d_{\mathbb{H}}$ is the standard metric of $\mathbb{H}^k$. In all our embeddings, the dimension k of the target space is a constant, independent of the size of the input graph. When the metric is clear from the context, we simply refer to it as $\mathcal{X}$ or G.

The quality of an embedding is measured by the worst-case *distortion* over all the distances in the metric, either in relative or absolute terms. In particular, let f be an embedding of the metric space $(\mathcal{X}_1, d_1)$ into the metric space $(\mathcal{X}_2, d_2)$. We say that f has *multiplicative distortion* c, for $c \geq 1$, if

$$d_2(f(x), f(y)) \leq d_1(x, y) \leq c\, d_2(f(x), f(y)) \;\; \forall x, y \in \mathcal{X}_1 \quad (1)$$

The embedding f has *additive distortion* c if

$$|d_1(x, y) - d_2(f(x), f(y))| \;\leq\; c \qquad \forall x, y \in \mathcal{X}_1 \qquad (2)$$

All our lower bounds in Section 3 are shown using multiplicative distortion, which implies similar (in fact, stronger) lower bounds on the additive distortion as well. Our upper bounds and algorithms, however, produce results with additive distortion.

Remark. Some papers on embeddings allow the input distances to be *scaled* before constructing the embedding, which may produce smaller distortion especially for the additive measure. We also utilize such a scaling by a *constant factor* for our upper bounds and algorithms. We do not consider scaling of distances by a non-constant factor because our intention is to investigate the intrinsic difference between hyperbolic and Euclidean space: after all, hyperbolic space behaves like Euclidean space in the limit at small distances. For the sake of simplicity, we do not use scaling in our lower bounds, although the results hold with any constant factor scaling. Finally, we point out that, by definition, our embeddings (for multiplicative distortion) are *non-expanding*: the distances in the target space are upper-bounded by the input space distances. This restriction does not influence our lower bounds but simplifies the presentation.

1.2 Results

We begin with a simple lower bound result: any embedding of an n-node cycle in hyperbolic space has multiplicative distortion at least $\Omega(n/\log n)$. This may seem surprising since the cycle is trivially embedded in Euclidean space with distortion $\pi/2$. We extend this lower bound using the concept of *quasi-cyclicity*: any graph with quasi-cyclicity n has multiplicative distortion $\Omega(n/\log n)$ in hyperbolic space. An example graph with large quasi-cyclicity but no large induced cycle is the $\sqrt{n} \times \sqrt{n}$ regular lattice. This graph also embeds with constant distortion in Euclidean space, but is shown to require $\Omega(\sqrt{n}/\log n)$ distortion in hyperbolic space.

Next, we establish a relation between quasi-cyclicity and δ-hyperbolicity (see Section 2 for a formal definition) of a graph as a way to prove upper bounds on the distortion. Using this relation, we show that graphs with small quasi-cyclicity can be embedded into hyperbolic space with only constant *additive* distortion. Finally, we present an efficient (linear-time) randomized algorithm for embedding a graph with small quasi-cyclicity into hyperbolic space, so that at least a $(1 - \varepsilon)$ fraction of the node-pairs have constant additive distortion, with high probability.

1.3 Related Work

The body of research on hyperbolic space and its connections to various network properties and dynamics is large. Due to lack of space, we just mention a few papers that are algorithmic in nature and explore questions similar to the ones discussed in our paper. In [11], Krauthgamer and Lee present algorithms for approximate nearest-neighbors and low-stretch routing, as well as a PTAS for the Traveling Salesman Problem in hyperbolic space. Their results also apply to δ-hyperbolic spaces. In [3], Chepoi *et al.* present schemes for computing an additive approximation of the diameter, center, and radius of δ-hyperbolic spaces and graphs. They also show that several graph classes are δ-hyperbolic and present a linear-time algorithm for approximating trees of n-node δ-hyperbolic graphs with $O(\delta \log n)$ additive distortion. In [16], Sarkar shows that every finite tree admits an embedding into $\mathbb{H}^2$ with only $1 + \varepsilon$ multiplicative distortion, for any $\varepsilon > 0$; for this result, however, the input metric needs to be scaled by a factor that depends on both ε and the maximum degree of the tree, which may be non-constant. Nonetheless, this suggests a natural approach to embed graph metrics into hyperbolic space: first embed the graph metric into a tree metric, and then embed the tree metric into $\mathbb{H}^2$ using the result of [16]. Some graph metrics (e.g. the n-cycle), however, require $\Omega(n)$ distortion when embedded into a tree metric, as shown in [15]. On the other hand, under a relaxed model where the target metric is not a single tree but rather a *distribution* over several tree metrics, one can achieve an *expected* distortion $O(\log n)$, as shown by Fakcharoenphol *et al.* [4]. In our paper, we only consider embedding the graph into a single target metric (hyperbolic space), with the goal of minimizing the worst-case distortion.

The hyperbolic-space embeddings have also proved useful for *greedy routing*: in [8], Kleinberg proves that every finite graph admits a greedy embedding into hyperbolic space, which is known not to be possible in Euclidean space.

2. PRELIMINARIES

We begin by establishing several technical preliminaries necessary for our proofs, including formal definitions of hyperbolic space. Given two points $x, y \in \mathcal{X}$ of a metric space $(\mathcal{X}, d)$, we use $[x, y]$ to denote a shortest path (geodesic) from x to y. A metric space is *geodesic* if, for every $x, y \in \mathcal{X}$, all points on the shortest path from x to y are in $\mathcal{X}$, where the shortest path is seen as a continuous curve. A graph metric is not geodesic, but can be made geodesic by adding all edges as continuous curves to the metric space. Hence, for a graph metric, $[x, y]$ actually denotes a shortest path in the geodesic version of the graph metric.

The distance between two sets $A, B \subseteq \mathcal{X}$ is written as $d(A, B)$, that is, $d(A, B) = \min_{x \in A, y \in B} d(x, y)$.

The hyperbolic space is a type of non-Euclidean geometry with negative curvature. Formally speaking, real hyperbolic k-space $\mathbb{H}^k$ is a k-dimensional Riemannian manifold with

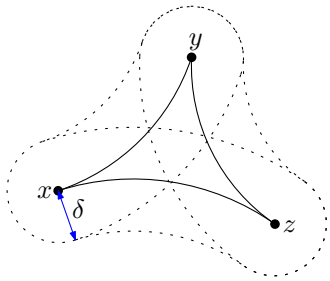

Figure 1: A δ-thin triangle.

constant sectional curvature -1. There are several models of hyperbolic space, including the hyperboloid model, Klein model, and Poincaré disc/half-plane models. Our results are model-independent and rely only on general properties of hyperbolic space. Because $\mathbb{H}^k$ is a manifold, angles locally behave as in Euclidean space but, in contrast to Euclidean space, the angles of a triangle satisfy $(\alpha+\beta+\gamma) < \pi$ in hyperbolic space. The difference $\pi-(\alpha+\beta+\gamma)$ is called the *defect* of a triangle, and it equals the *area* of the triangle in $\mathbb{H}^k$.

One can define a more general notion of hyperbolic spaces called *δ-hyperbolic spaces*. The following definition is due to Gromov [6]. For a metric space $(\mathcal{X}, d)$ and $p, x, y \in \mathcal{X}$, the *Gromov product* of x and y at p, denoted by $(x|y)_p$ is defined as:

$$(x|y)_p = \frac{1}{2}(d(p,x) + d(p,y) - d(x,y)) \tag{3}$$

A metric space $(\mathcal{X}, d)$ is δ-hyperbolic for some $\delta \geq 0$ if the following holds for all $p, x, y, z \in \mathcal{X}$:

$$(x|z)_p \geq \min((x|y)_p, (y|z)_p) - \delta \tag{4}$$

Another definition for geodesic metric spaces is due to Rips. For three points $x, y, z \in \mathcal{X}$, a *geodesic triangle* consists of $[x, y]$, $[y, z]$, and $[x, z]$. A geodesic triangle is *δ-thin* if for every $p \in [x, y]$, we have $d(p, [y, z] \cup [x, z]) \leq \delta$ (and symmetrically for $p \in [x, z]$ and $p \in [y, z]$). See Fig. 1. A metric space is δ-hyperbolic if all geodesic triangles are δ-thin. These two definitions are equivalent *upto a constant factor* of δ, do not change the asymptotic bounds of our theorems, and therefore we use them interchangeably.

We call G a δ-hyperbolic graph if its metric is δ-hyperbolic. For instance, trees are 0-hyperbolic, cliques are $\frac{1}{2}$-hyperbolic (if we include points on edges), and a graph with diameter D is $\frac{D}{2}$-hyperbolic. Unless stated otherwise, if we say a metric space is δ-hyperbolic, then we assume that δ is a constant. The definitions of δ-hyperbolic metric spaces give only an upper bound on δ. When proving lower bounds, we say a metric space or graph is *strictly δ-hyperbolic* if it is δ-hyperbolic and not δ'-hyperbolic for $\delta' < \delta$. In this definition we generally allow δ to depend on the input size (e.g., the number of vertices of the graph metric).

3. LOWER BOUNDS ON DISTORTION

In this section, we prove that some simple graphs require large distortion in any hyperbolic space embedding. In the process we introduce the notion of *quasi-cyclicity*, which turns out to be closely related to the distortion. Let C_n be the cycle graph with n nodes $v_0, \ldots, v_{n-1}$, with edges of the form (v_i, v_{i+1}). We refer to the graph metric of C_n as the *cycle metric* of size n, denoted by (C_n, d_C). Our proof uses the following result of Knaster, Kuratowski and Mazurkiewicz [10].

LEMMA 3.1 ([10]). *Consider an n-dimensional simplex Δ_n with vertices $V = \{v_0, \ldots, v_n\}$, and a closed set S_v for each $v \in V$. If, for each subset $A \subseteq V$, the union $\bigcup_{v \in A} S_v$ covers the face of Δ_n spanned by the vertices in A, then $\bigcap_{v \in V} S_v \neq \emptyset$.*

LEMMA 3.2. *Let $p \in \mathbb{H}^k$ and let $P\colon [0,1] \to \mathbb{H}^k$ be a curve such that $\angle P(0)pP(1) = \alpha$ and $d_{\mathbb{H}}(p, P(t)) \geq R$ for $0 \leq t \leq 1$. Then the length of P is at least $\alpha \sinh(R)$.*

PROOF. Our proof, which is straightforward but technical, is presented in Appendix A, to maintain the conceptual flow of the discussion. □

The following lemma presents our key lower bound, namely, that an n-cycle suffers a large distortion in any hyperbolic space embedding.

LEMMA 3.3. *Any embedding of C_n into $\mathbb{H}^k$ has multiplicative distortion $\Omega(\frac{n}{\log n})$.*

PROOF. Let us assume n is a multiple of 3, and consider the nodes $u_1 = v_0$, $u_2 = v_{\frac{n}{3}}$, and $u_3 = v_{\frac{2n}{3}}$. We have $d_C(u_i, u_j) = \frac{n}{3}$, for $1 \leq i < j \leq 3$. Let f be an embedding of C_n into $\mathbb{H}^k$, for some $k \geq 2$, with $p_i = f(u_i)$ as the image of u_i, for $1 \leq i \leq 3$. We focus on the geodesic triangle T formed by p_1, p_2, and p_3 in $\mathbb{H}^k$, and let α_i be the angle of T at p_i. See Fig. 2(a) for illustration.

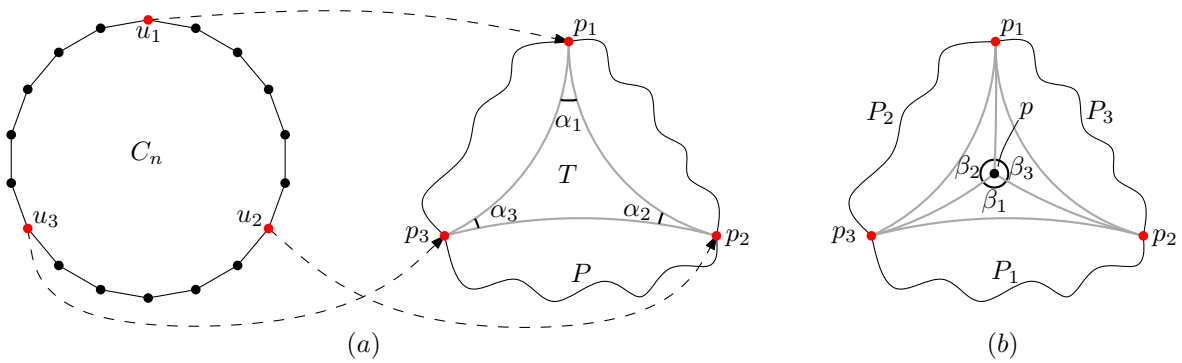

Figure 2: (a) A metric embedding of C_n into $\mathbb{H}^k$. (b) Subdividing T into three triangles.

First, suppose one of the angles, say α_1, is at least $\frac{2\pi}{3}$. Consider the polygonal curve $P\colon [0,1] \to \mathbb{H}^k$ that is the image of the geodesic segment $[u_2, u_3]$ in $\mathbb{H}^k$ under f, where $P(0) = p_2$ and $P(1) = p_3$. Since distances do not expand under f, the length of P is at most $\frac{n}{3}$. On the other hand, by Lemma 3.2, the length of P is at least $\alpha_1 \sinh(R)$, where R is the minimum distance between p_1 and P. Since $\sinh(x) = (e^x - e^{-x})/2$, we get that if $\alpha_1 \sinh(R) = O(n)$, then $R = O(\log n)$. Let $p \in P$ be the point such that $d_{\mathbb{H}}(p_1, p) = R$. Although p may not be the image of one of the nodes of C_n, there must exist a node $u \in [u_2, u_3]$ such that $d_{\mathbb{H}}(p, f(u)) \leq 1$ (distances do not expand under f), and so $d_{\mathbb{H}}(p_1, f(u)) = O(\log n)$ by the triangle inequality. This leads to the desired lower bound: the node-pair u, u_1 has graph distance $d_C(u, u_1) \geq \frac{n}{3}$, but hyperbolic-space distance $O(\log n)$, and therefore suffers a multiplicative distortion of at least $\Omega(\frac{n}{\log n})$.

The above argument also works if $\alpha_1 = \Omega(1)$. If none of the angles of triangle T is at least of size $\Omega(1)$, then we need a different argument (recall that the sum of the angles of a hyperbolic triangle is less than π, and possibly $o(1)$). In this case, consider a point p in the interior of T, and the three triangles formed by the geodesic segments $[p, p_i]$. See Fig. 2(b). Let β_i be the angle at p opposite of p_i. We define the closed set S_i $(1 \leq i \leq 3)$ as the set of points p such that $\beta_i \leq \frac{2\pi}{3}$. It is easy to verify that the sets S_i satisfy the conditions of Lemma 3.1, and therefore there exist a point p such that $\beta_i \leq \frac{2\pi}{3}$ for all $1 \leq i \leq 3$. Since those three angles sum to 2π, we must have $\beta_1 = \beta_2 = \beta_3 = \frac{2\pi}{3}$. Let $P_1, P_2, P_3\colon [0,1] \to \mathbb{H}^k$ be the polygonal curves obtained by embedding $[u_2, u_3]$, $[u_1, u_3]$, and $[u_1, u_2]$ into $\mathbb{H}^k$, respectively. Using the same line of arguments as above, there must be nodes $v_1 \in [u_2, u_3]$, $v_2 \in [u_1, u_3]$, and $v_3 \in [u_1, u_2]$ such that $d_{\mathbb{H}}(p, f(v_i)) = O(\log n)$ for $1 \leq i \leq 3$. By the triangle inequality, this implies that $d_{\mathbb{H}}(f(v_i), f(v_j)) = O(\log n)$ for $1 \leq i < j \leq 3$. Finally, since there must be a pair of nodes (v_i, v_j) such that $d_C(v_i, v_j) \geq \frac{n}{3}$, we again find a node-pair that suffers a multiplicative distortion of at least $\Omega(\frac{n}{\log n})$ under the embedding f. This concludes the proof. $\square$

The lower bound of the preceding lemma is asymptotically tight: embed the nodes of C_n on a hyperbolic circle of radius $O(\log n)$ by ensuring that each adjacent pair of points on the circle has distance 1. The lower bound is independent of the dimension of the target hyperbolic space, and the result extends to any metric space that includes C_n as an induced submetric. However, there are graphs that do not contain large cycle submetrics and yet require large distortion. One such metric space is the $\sqrt{n} \times \sqrt{n}$ lattice graph L_n whose largest cycle submetric has size only 4, but we show below that any embedding of L_n into $\mathbb{H}^k$ has $\Omega(\frac{\sqrt{n}}{\log n})$ multiplicative distortion. In fact, we introduce a generalization of cycle metrics, called *quasi-cycles*, which allows us to extend the lower bound to a much larger class of graphs.

Quasi-cycles. A *quasi-cycle* is a metric space that is similar to the cycle metric in which some shortcuts are allowed, but those shortcuts do not reduce the distances by more than a constant factor[1]. More precisely, a quasi-cycle of size n is a metric space (C_n, d) for which there is a constant $\alpha > 0$,

[1]We should note that our quasi-cycles are unrelated to the λ-quasi-circles defined in [2].

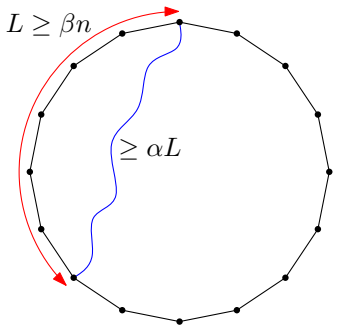

Figure 3: An (α, β)-quasi-cycle.

such that the following inequality holds:

$$\alpha d_C(u,v) \leq d(u,v) \leq d_C(u,v) \qquad \forall u, v \in C_n, \tag{5}$$

where α defines the *quasiness* of the quasi-cycle. For technical reasons we also introduce the notion of *weak quasi-cycles* where the bound of Inequality (5) holds only if $d_C(u,v) \geq \beta n$, for some constant $0 \leq \beta \leq \frac{1}{2}$. See Figure 3 for illustration. The constant β is the *weakness* of a quasi-cycle (for strong quasi-cycles, $\beta = 0$). We refer to strong quasi-cycles as α-quasi-cycles and weak quasi-cycles as (α, β)-quasi-cycles.

LEMMA 3.4. *Any embedding of an (α, β)-quasi-cycle (C_n, d), where $\alpha > 0$ and $\beta \leq \frac{1}{3}$, into $\mathbb{H}^k$ has multiplicative distortion at least $\Omega(\frac{n}{\log n})$.*

PROOF. The proof is analogous to that of Lemma 3.3. Given any embedding f of (C_n, d) into $\mathbb{H}^k$, there are two nodes $u, v \in C_n$ whose distance in the cycle is $d_C(u,v) \geq \frac{n}{3}$ but whose distance in hyperbolic space is $d_{\mathbb{H}}(f(u), f(v)) = O(\log n)$. By definition of quasi-cycles, we have $d(u,v) \geq \alpha \frac{n}{3}$ if $\beta \leq \frac{1}{3}$, and so the multiplicative distortion of node-pair u, v in f is at least $\Omega(\frac{n}{\log n})$. $\square$

Returning to the $\sqrt{n} \times \sqrt{n}$ lattice graph L_n, consider the submetric induced by the outer cycle of L_n. It has $O(\sqrt{n})$ nodes and it is a $\frac{1}{2}$-quasi-cycle. Therefore, by Lemma 3.4, any embedding of L_n into $\mathbb{H}^k$ has $\Omega(\frac{\sqrt{n}}{\log n})$ multiplicative distortion.

If we define the *quasi-cyclicity* of a metric space $(\mathcal{X}, d)$ as the size of the largest induced (α, β)-quasi-cycle with $\alpha > 0$ and $\beta \leq \frac{1}{3}$, then we have the following general lower bound.

THEOREM 3.5. *Let m be the quasi-cyclicity of a metric space $(\mathcal{X}, d)$. Then any embedding of $(\mathcal{X}, d)$ into $\mathbb{H}^k$ has multiplicative distortion at least $\Omega(\frac{m}{\log m})$.*

Remark. Naturally, the quasi-cyclicity m of a metric space depends on the value of α, and to a lesser extent on the value of β. As a function of α, the lower bound for the multiplicative distortion is $\Omega(\frac{\alpha m}{\log m})$, which matches $\Omega(\frac{m}{\log m})$ for any constant value of α.

4. QUASI-CYCLES AND δ-HYPERBOLIC SPACES

In this section we show a counterpart to Theorem 3.5: every metric space with small quasi-cyclicity admits an embedding with small distortion into $\mathbb{H}^k$. We do this by establishing a relation between quasi-cyclicity and δ-hyperbolicity.

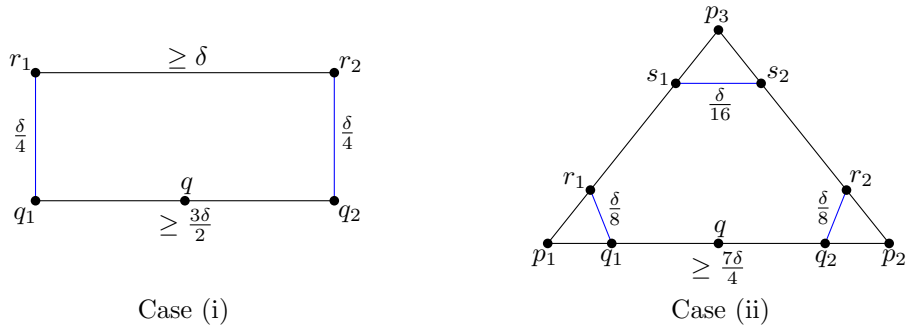

Figure 4: The cases of Theorem 4.1.

THEOREM 4.1. *If a geodesic metric space $(\mathcal{X}, d)$ is strictly δ-hyperbolic, then $(\mathcal{X}, d)$ must contain an $(\alpha, \frac{1}{3})$-quasi-cycle of size $\Omega(\delta)$ for some constant $\alpha > 0$.*

PROOF. We prove the result using Rips' definition of δ-hyperbolic metric spaces. By definition of strict δ-hyperbolicity, there must be three points $p_1, p_2, p_3 \in \mathcal{X}$ and a point $q \in [p_1, p_2]$ such that $d(q, [p_1, p_3] \cup [p_2, p_3]) = \delta$. Then, for any given $0 \leq x \leq \delta$, there must exist points $q_1 \in [p_1, q]$ and $q_2 \in [q, p_2]$ such that $d(q_1, [p_1, p_3] \cup [p_2, p_3]) = d(q_2, [p_1, p_3] \cup [p_2, p_3]) = x$. There are two cases to consider.

Case (i). There exists a point $q_1 \in [p_1, q]$ such that $d(q_1, [p_2, p_3]) = \frac{\delta}{4}$. Note that we can switch p_2 and p_1 without loss of generality, so this also covers the symmetric case. Let $q_1 \in [p_1, q]$ be the point closest to q such that $d(q_1, [p_2, p_3]) = \frac{\delta}{4}$. Define q_2 in the same way, where $q_2 \in [q, p_2]$. Let $r_1 \in [p_2, p_3]$ be the point such that $d(q_1, r_1) = \frac{\delta}{4}$ and let $r_2 \in [p_2, p_3]$ be the point such that $d(q_2, r_2) = \frac{\delta}{4}$ (see Fig. 4 left). The points q_1, r_1, r_2, and q_2 form a geodesic quadrilateral C. Since $d(q, [p_2, p_3]) \geq \delta$, the triangle inequality implies that $d(q, q_1) \geq \frac{3\delta}{4}$ and $d(q, q_2) \geq \frac{3\delta}{4}$, and hence $d(q_1, q_2) \geq \frac{3\delta}{2}$. This also means that $d(r_1, r_2) \geq \delta$, since otherwise $d(q_1, q_2) < \frac{3\delta}{2}$. Hence the length of C is at least 3δ. For now we assume that the length of C is also $O(\delta)$. We show that if the distance along C between two points $x, y \in C$ is at least δ, then $d(x, y) \geq \frac{\delta}{4}$. This implies that C is an $(\alpha, \frac{1}{3})$-quasi-cycle. By construction we directly get that $d([q_1, q_2], [r_1, r_2]) \geq \frac{\delta}{4}$. By the triangle inequality we get that $d([q_1, r_1], [q_2, r_2]) \geq \delta$. Finally consider points $x \in [q_1, q_2]$ and $y \in [q_1, r_1]$ such that the distance between x and y along C is at least δ. In that case $d(x, q_1) \geq \frac{3\delta}{4}$. Again, by the triangle inequality we get that $\frac{3\delta}{4} \leq d(x, q_1) \leq d(x, y) + d(y, q_1) \leq d(x, y) + \frac{\delta}{4}$ and hence $d(x, y) \geq \frac{\delta}{2}$. The other cases are symmetric. Thus C is an $(\alpha, \frac{1}{3})$-quasi-cycle of size $\Omega(\delta)$.

Case (ii). We have $d([p_1, q], [p_2, p_3]) > \frac{\delta}{4}$ and $d([p_2, q], [p_1, p_3]) > \frac{\delta}{4}$. Let $q_1 \in [p_1, q]$ be the point closest to q such that $d(q_1, [p_1, p_3]) = \frac{\delta}{8}$. Define q_2 in the same way, where $q_2 \in [q, p_2]$ and $d(q_2, [p_2, p_3]) = \frac{\delta}{8}$. Let $r_1 \in [p_1, p_3]$ be the point such that $d(q_1, r_1) = \frac{\delta}{8}$ and let $r_2 \in [p_2, p_3]$ be the point such that $d(q_2, r_2) = \frac{\delta}{8}$. As in Case (i) we can derive that $d(q_1, q_2) \geq \frac{7\delta}{4}$. We also get that $d(r_1, r_2) \geq \frac{3\delta}{2}$. Instead of using the geodesic $[r_1, r_2]$, we find points $s_1 \in [r_1, p_3]$ and $s_2 \in [r_2, p_3]$ such that $d(s_1, s_2) = \frac{\delta}{16}$ and $d([r_1, s_1], [r_2, s_2]) \geq \frac{\delta}{16}$. Note that these points must exist. Now the points q_1, r_1, s_1, s_2, r_2, and q_2 form a geodesic hexagon C with length at least $\frac{7\delta}{2}$ (see Fig. 4 right). Again we assume that the length of C is $O(\delta)$. We now show that if the distance along C between two points $x, y \in C$ is at least $\frac{\delta}{2}$, then $d(x, y) \geq \frac{\delta}{16}$. This implies that C is an $(\alpha, \frac{1}{3})$-quasi-cycle (actually, it is an $(\alpha, \frac{1}{7})$-quasi-cycle). It is easy to see that $d([q_1, q_2], [r_1, s_1] \cup [r_2, s_2]) \geq \frac{\delta}{8}$ and $d([r_1, s_1], [r_2, s_2]) \geq \frac{\delta}{16}$. The following cases remain (symmetric cases are ignored).

(a) $x \in [q_1, q_2]$ and $y \in [q_1, r_1]$. Because the distance along C is at least $\frac{\delta}{2}$, we get that $d(x, q_1) \geq \frac{3\delta}{8}$. Now, by the triangle inequality, $d(x, y) \geq \frac{\delta}{4}$.

(b) $x \in [q_1, q_2]$ and $y \in [s_1, s_2]$. By construction and the triangle inequality we get that $\frac{\delta}{8} \leq d(x, s_1) \leq d(x, y) + d(y, s_1) \leq d(x, y) + \frac{\delta}{16}$. Thus $d(x, y) \geq \frac{\delta}{16}$.

(c) $x \in [q_1, r_1]$ and $y \in [r_1, s_1]$. Similar to Case (a).

(d) $x \in [q_1, r_1]$ and $y \in [s_1, s_2]$. If $d(x, y) \leq \frac{\delta}{16}$ then $d(q_1, s_2) \leq d(q_1, x) + d(x, y) + d(y, s_1) \leq \frac{\delta}{4}$. That means that we should be in Case (i), so $d(x, y) \geq \frac{\delta}{16}$.

(e) $x \in [q_1, r_1]$ and $y \in [r_2, s_2]$. Similar to Case (d).

(f) $x \in [q_1, r_1]$ and $y \in [q_2, r_2]$. By the triangle inequality we get that $\frac{7\delta}{4} \leq d(q_1, q_2) \leq d(q_1, x) + d(x, y) + d(y, q_2) \leq \frac{\delta}{4} + d(x, y)$. Thus $d(x, y) \geq \frac{3\delta}{2}$.

(g) $x \in [r_1, s_1]$ and $y \in [s_1, s_2]$. Similar to Case (a).

Thus, C is an $(\alpha, \frac{1}{3})$-quasi-cycle of size $\Omega(\delta)$.

Finally we consider the case that the length of C in Case (i) or Case (ii) is $\omega(\delta)$, or in particular, at least 37δ. This can happen only if $d(q_1, q_2)$, $d(r_1, s_1)$, or $d(r_2, s_2)$ is at least 12δ. Without loss of generality we assume that $d(q_1, q_2) \geq 12\delta$. Now it is easy to see that we can find points $q_1', q_2' \in [q_1, q_2]$ with the following properties: (1) $6\delta \leq d(q_1', q_2') \leq 12\delta$, and (2) there exist $r_1', r_2' \in [p_1, p_3]$ or $r_1', r_2' \in [p_2, p_3]$ such that $\frac{\delta}{8} \leq d(q_1', r_1'), d(q_2', r_2') \leq \delta$. (Note that, for each $x \in [q_1, q_2]$, either $d(x, [p_1, p_3]) \leq \delta$ or $d(x, [p_2, p_3]) \leq \delta$.) Hence, the cycle C formed by q_1', r_1', r_2', q_2' has length $\Theta(\delta)$ and is a scaled version of Case (i). Thus, C is an $(\alpha, \frac{1}{3})$-quasi-cycle of size $\Omega(\delta)$. $\square$

Remark. The result easily extends to graph metrics of unweighted graphs (which are not geodesic). This follows because the proof works for the geodesic version of the graph metric, and all distances involved in the proof change by at

most 1 when considering the original graph metric (since each edge has length 1).

In the preceding theorem, we considered (α, β)-quasi-cycles with $\beta = \frac{1}{3}$, which is sufficient to bound the quasi-cyclicity of a strictly δ-hyperbolic metric space. But we can play with the constants to obtain the same result for other values of β, leading to a tradeoff between α, β, and the size of the quasi-cycle. Theorem 4.1 also implies that, if the quasi-cyclicity of a metric space $(\mathcal{X}, d)$ is $O(1)$, then $(\mathcal{X}, d)$ is δ-hyperbolic. (We remind the reader again that by convention δ is constant for δ-hyperbolicity.)

4.1 An Upper Bound for Distortion

Unfortunately, not all δ-hyperbolic metric spaces embed well into $\mathbb{H}^k$, and we need some additional assumptions to achieve low distortion embeddings into hyperbolic space. In Euclidean space, good embeddings are associated with bounded *doubling dimension*: for example, the beacon-based embedding of [9] requires bounded doubling dimension to achieve constant (multiplicative) distortion. However, bounded doubling dimension alone does not guarantee low distortion: the cycle metric has bounded doubling dimension yet it embeds very poorly into hyperbolic space. On the other hand, if we assume δ-hyperbolicity then a weaker condition than doubling dimension, called *bounded growth at some scale* and discussed below, suffices for a good upper bound on distortion.

Given a metric space $(\mathcal{X}, d)$ and a point $x \in \mathcal{X}$, define the *ball* $B_x(r) = \{y \in \mathcal{X} \mid d(x, y) \leq r\}$. The *doubling dimension* of $(\mathcal{X}, d)$ is the smallest k such that every ball in $(\mathcal{X}, d)$ of radius r can be covered by at most 2^k balls of radius $\frac{r}{2}$. We also say $(\mathcal{X}, d)$ is 2^k*-doubling*, or simply *doubling*, if k is constant.[2] A metric space $(\mathcal{X}, d)$ has *bounded growth at some scale* if there exist constants $R > r > 0$ such that every ball of radius R can be covered by K balls of radius r for some constant K. It is easy to see that this is a weaker condition than doubling.

In the context of graph metrics, it turns out that *bounded node degree* implies bounded growth at some scale, and so that is the assumption we make for our graphs. (The bounded degree assumption is typically satisfied in most practical graphs, including social networks, but can also be theoretically justified: we argue in Lemma A.1 (Appendix A) that without the bounded degree assumption, one cannot achieve constant (multiplicative) distortion, in constant-dimensional hyperbolic space using only a constant factor scaling.) Finally, for our upper bound, we also allow a constant-factor scaling of the input metric: this affects neither our lower bounds nor the intrinsic behavior of hyperbolic space.

For the sake of brevity, we define a *good additive (multiplicative) embedding* as an embedding with constant additive (multiplicative) distortion subject to rescaling of the metric by only a constant factor. We can now state the main result of this section.

THEOREM 4.2. *Let G be a graph metric with bounded degree and constant quasi-cyclicity. Then, there exists a good additive embedding of G into $\mathbb{H}^k$.*

PROOF. The proof follows from Theorem 4.1 in combination with the following result of Bonk and Schramm [2] (with our notation[3]): Any δ-hyperbolic metric space $(\mathcal{X}, d)$ with bounded growth at some scale admits a good additive embedding into $\mathbb{H}^k$. $\square$

5. A LINEAR TIME ALGORITHM FOR HYPERBOLIC EMBEDDING

We now address the algorithmic question of efficiently embedding graph metrics into hyperbolic space. Throughout this section we assume that the input is a δ-hyperbolic graph metric $G = (V, E)$ with bounded degree. Our goal is to compute an embedding of G into $\mathbb{H}^k$ with small distortion, for some constant k.

The result by Bonk and Schramm [2] is constructive but not algorithmic. Since we are interested in embedding large-scale graphs, such as social networks, their size rules out the feasibility of quadratic-time algorithms. Our main result in this section is a simple randomized algorithm that runs in linear $O(|V| + |E|)$ time, and achieves constant additive distortion for all but an ε fraction of the node-pairs of G. The idea of ensuring good distortion for most of the node-pairs is widely used in practice [17, 18, 20], with a rigorous theoretical treatment by Kleinberg *et al.* [9] for Euclidean space. We also follow this beacon-approach, however, we cannot apply their algorithm directly because our input graphs do not have bounded doubling dimension, and our target space is hyperbolic space instead of Euclidean space.

Our algorithm is inspired by the results of [2]. The main idea is to consider the so-called *Gromov boundary* of a metric space. The boundary ∂G of G is doubling, and the boundary $\partial \mathbb{H}^k$ of $\mathbb{H}^k$ resembles $\mathbb{R}^{k-1}$. We can therefore use the beacon-approach of Kleinberg *et al.* to compute an embedding f of ∂G into $\partial \mathbb{H}^k$. We then use results of [2] to extend f to an embedding F of G into $\mathbb{H}^k$ (see Fig. 5; more details are explained in Section 5.1). Intuitively, the boundary of a metric space $(\mathcal{X}, d)$ can be seen as the set of geodesic rays of $\mathcal{X}$ originating from a fixed base point in $\mathcal{X}$. Thus, f is an embedding of the set of geodesic rays of G into the set of geodesic rays of $\mathbb{H}^k$, and F extends this embedding to points on the geodesic rays. Interestingly, if f is a good *multiplicative* embedding, then F is a good *additive* embedding.

Our algorithm consists of the following steps:

1. Compute the boundary ∂G.
2. Compute an embedding f of ∂G into Euclidean space ($\mathbb{R}^{k-1}$).
3. Extend f to an embedding F of G into $\mathbb{H}^k$.

For step (3) we use the results of [2], which we describe in Section 5.1. Since ∂G may contain a quadratic number of node-pairs, we cannot compute ∂G in linear time. Therefore, the key challenge is to compute f without fully constructing ∂G. Before we can describe this algorithm in detail, we first need to review the key concepts used in the result of [2].

[2]The doubling dimension is closely related to the Assouad dimension, see [1]. A metric space is doubling if and only if it has finite Assouad dimension.

[3]The result of [2] actually requires the metric space to be geodesic, but this restriction can be ignored using [2, Theorem 4.1]. We point out that our graph metrics are 1-almost geodesic using their notation.

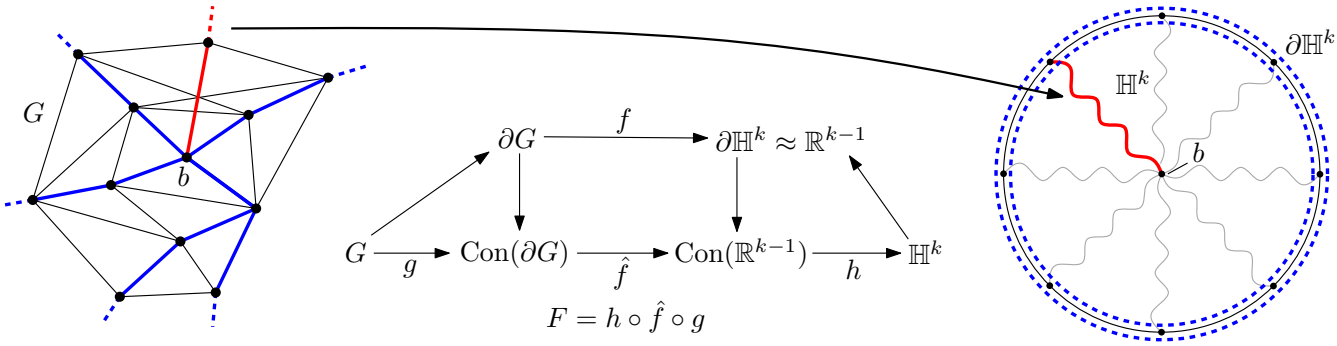

Figure 5: A schematic representation of the embedding of G into $\mathbb{H}^k$.

5.1 Embedding δ-hyperbolic Metric Spaces into $\mathbb{H}^k$

Gromov boundary. The *boundary* $\partial\mathcal{X}$ of a δ-hyperbolic metric space $(\mathcal{X}, d)$ is defined as follows. Given a point $b \in \mathcal{X}$, we say a sequence $\{x_i\} \subseteq \mathcal{X}$ *converges at infinity* if $\lim_{i,j\to\infty}(x_i|x_j)_b = \infty$, where $(y|z)_x$ is the Gromov product of y and z at x. Two sequences $\{x_i\} \subseteq \mathcal{X}$ and $\{y_i\} \subseteq \mathcal{X}$ that converge at infinity are called *equivalent* if $\lim_{i\to\infty}(x_i|y_i)_b = \infty$. The elements of $\partial\mathcal{X}$ are the equivalence classes of sequences that converge at infinity. As mentioned above, the elements of $\partial\mathcal{X}$ correspond (roughly) to geodesic rays originating from b. An example with $\mathcal{X} = \mathbb{H}^2$ (in the Poincaré disc model) is shown on the right of Figure 5. Using limits we can easily extend the definition of the Gromov product $(y|z)_x$ to the boundary $\partial\mathcal{X}$, i.e., $y, z \in \mathcal{X} \cup \partial\mathcal{X}$ and $x \in \mathcal{X}$ (see [2] for details).

The notion of a boundary does not make sense for finite graphs, which are our primary interest, and so we extend a graph metric G as follows. Let $b \in V$ be an arbitrary node, called the *base* of G, and consider the shortest path tree of G rooted at b. Let $L(G) \subset V$ be the set of leaves of this tree. To each leaf $u \in L(G)$ we attach a chain of nodes γ_u that extends to infinity, and let G' be the resulting metric space (see Fig. 5 left). Each element of $\partial G'$ can be identified with γ_u for some $u \in L(G)$. The definition of the Gromov product implies that $(\gamma_u|\gamma_v)_b = (u|v)_b$, and so, considering only the Gromov product, we can simply identify the boundary ∂G with $L(G)$.[4]

Returning to general boundaries, we can define a metric $d_{\partial\mathcal{X}} = d_{b,\epsilon}$ on a boundary $\partial\mathcal{X}$. This metric depends on the choice of the base point $b \in \mathcal{X}$ and a constant $\epsilon > 0$, and is defined as follows:

$$d_{\partial\mathcal{X}}(x, y) = \inf\left\{\sum_{i=1}^{m} e^{-\epsilon(x_{i-1}|x_i)_b}\right\} \qquad \text{for } x, y \in \partial\mathcal{X} \quad (6)$$

where the infimum is taken over all finite sequences $x = x_0, x_1, \ldots, x_m = y$ in $\partial\mathcal{X}$, and where $e^{-\infty} = 0$. Note that $d_{\partial\mathcal{X}}$ depends only on Gromov products in $\partial\mathcal{X}$, and hence extends to the set of leaves $L(G)$ of a (finite) graph metric G, as described above. The following important property was shown in [5, p. 285].

Lemma 5.1. *There exists a constant ϵ_0 such that, if $\mathcal{X}$ is δ-hyperbolic and $\epsilon\delta \le \epsilon_0$, then*

$$\frac{1}{2}e^{-\epsilon(x|y)_b} \le d_{\partial\mathcal{X}}(x, y) \le e^{-\epsilon(x|y)_b} \qquad \forall x, y \in \partial\mathcal{X}$$

Throughout the rest of the paper, we simply write $d_{\partial\mathcal{X}}$ for the boundary metric, implicitly assuming that our choice of ϵ satisfies Lemma 5.1, since all other relevant properties of $d_{\partial\mathcal{X}}$ do not depend on the choice of b or ϵ. The following property of the boundary is crucial.

Lemma 5.2 ([2, Proposition 6.2 & Theorem 9.2]). *If $(\mathcal{X}, d)$ is δ-hyperbolic and has bounded growth at some scale, then $(\partial\mathcal{X}, d_{\partial\mathcal{X}})$ is doubling and has a bounded diameter.*

Hyperbolic extension. One can also define a metric space that acts as the "opposite" of the boundary. Let $(\mathcal{X}, d)$ be a bounded metric space with diameter D. The *hyperbolic extension* $\mathrm{Con}(\mathcal{X})$ of $(\mathcal{X}, d)$ is the metric space $(\mathcal{X} \times (0, D], d_e)$ where

$$d_e((x, h), (x', h')) = 2\log\left(\frac{d(x, x') + \max(h, h')}{\sqrt{hh'}}\right) \qquad (7)$$
$$\text{for } (x, h), (x', h') \in \mathcal{X} \times (0, D]$$

It can easily be shown that $\mathrm{Con}(\mathcal{X})$ is δ-hyperbolic. Furthermore, there exists a good additive embedding of $\mathcal{X}$ into $\mathrm{Con}(\partial\mathcal{X})$ and vice versa (see [2, Theorem 8.2]). Finally, if there is a good *multiplicative* embedding f of $\mathcal{X}_1$ into $\mathcal{X}_2$, then we can find a good *additive* embedding $\hat{f}$ of $\mathrm{Con}(\mathcal{X}_1)$ into $\mathrm{Con}(\mathcal{X}_2)$ (see [2, Theorem 7.4]).

We are now ready to describe step (3) of our algorithm. Given a good multiplicative embedding f of ∂G into $\partial\mathbb{H}^k \approx \mathbb{R}^{k-1}$, we can compute a good additive embedding $\hat{f}$ of $\mathrm{Con}(\partial G)$ into $\mathrm{Con}(\partial\mathbb{H}^k)$. We can further compute good additive embeddings g and h from G into $\mathrm{Con}(\partial G)$ and from $\mathrm{Con}(\partial\mathbb{H}^k)$ into $\mathbb{H}^k$, respectively. The extension F of f is simply given by $F = h \circ \hat{f} \circ g$, which is a good additive embedding of G into $\mathbb{H}^k$ (see Fig. 5). Furthermore, the embedding F can be computed using f in linear time, since $\hat{f}$, g, and h are explicitly stated in [2].[5]

[4] Using the terminology of [2], our construction implies that the set of geodesic rays starting from b is cobounded in the geodesic version of the graph metric (since each node is on a geodesic ray). This avoids the method employed in [2] to deal with finite metrics, and results in a more natural approach to deal with unweighted graph metrics.

[5] These embeddings require diameters of ∂G and $f(\partial G)$, which cannot be computed in linear time, but an approximation of the diameters suffices for our purpose. In particular, a 2-approximation of the diameter D of $(\mathcal{X}, d)$ is easily computed in linear time, since $D \le 2\max_{y\in\mathcal{X}} d(x, y) \le 2D$ for all $x \in \mathcal{X}$.

5.2 Beacon-based Embedding for δ-hyperbolic Graphs

In this section, we show how to efficiently compute the embedding f of ∂G into $\mathbb{R}^{k-1}$. In order to compute f, we adapt the algorithm of Kleinberg et al. [9], where a fixed number of nodes (beacons) are used to embed a doubling metric space into a fixed-dimensional Euclidean space. Since we cannot efficiently compute a metric on ∂G, we need a modified approach. Before we describe our method, however, let us briefly review the beacon-based scheme of [9].

Given a doubling metric space $(\mathcal{X}, d)$, we pick a constant number of *beacons* $\mathcal{B} \subseteq \mathcal{X}$, and consider only the distances between pairs in $E_{\mathcal{B}} = \mathcal{B} \times \mathcal{X}$. Kleinberg et al. [9] show that, for a suitably chosen set $\mathcal{B}$, if f is a good multiplicative embedding of $\mathcal{X}$ into Euclidean space for pairs in $E_{\mathcal{B}}$, then it is also a good multiplicative embedding on most pairs of $\mathcal{X}$. This can be formalized by the concept of an (ε, β)-*base*. A set of beacons $\mathcal{B}$ is an (ε, β)-base if for all but an ε-fraction of the pairs $(u, v) \in \mathcal{X} \times \mathcal{X}$, there exists a $b \in \mathcal{B}$ such that $\min(d(u, b), d(v, b)) \leq \beta d(u, v)$. If $\mathcal{X}$ is s-doubling, then a random sample of $O(\frac{1}{\varepsilon} \log \frac{1}{\varepsilon})(\frac{2}{\beta})^{2 \log s}$ beacons is an (ε, β)-base with high probability. In addition, if $\mathcal{B}$ is an (ε, β)-base, and f is an embedding of $\mathcal{X}$ into Euclidean space with multiplicative distortion $\Delta \leq \frac{1}{4\beta}$ on $E_{\mathcal{B}}$, then f has $O(\Delta)$ multiplicative distortion on all but an ε fraction of node-pairs in $\mathcal{X}$. Furthermore, f can be computed efficiently using only the distance pairs in $E_{\mathcal{B}}$.

We now describe how to utilize this beacon approach to our setting. We pick an arbitrary node $b \in V$ of the graph G, and compute the shortest path tree rooted at b, which takes linear time using breadth-first search. We identify ∂G with the set of leaves $L(G)$, using the metric in Equation 6. We then randomly pick a constant number of beacons $\mathcal{B} \subseteq L(G)$, and compute distances from the beacons to all the nodes of V, which also takes linear time. Let $E_{\mathcal{B}} = \mathcal{B} \times L(G)$. Using the shortest paths trees from the beacon nodes, we can also compute $(u|v)_b$ for $(u, v) \in E_{\mathcal{B}}$. Let $\tilde{d}(u, v) = \frac{1}{2} e^{-\epsilon(u|v)_b}$, for ϵ small enough. By Lemma 5.1 we have $\tilde{d}(u, v) \leq d_{\partial G}(u, v) \leq 2\tilde{d}(u, v)$. Unfortunately $\tilde{d}$ may not be a metric, but we can easily fix that problem by considering the weighted graph $\tilde{G} = (L(G), E_{\mathcal{B}})$ with the weights given by $\tilde{d}$, using the shortest path metric $\bar{d}$ on $\tilde{G}$.

LEMMA 5.3. *We can compute $\bar{d}$ for $(u, v) \in E_{\mathcal{B}}$ in linear time and $\bar{d}(u, v) \leq d_{\partial G}(u, v) \leq 2\bar{d}(u, v)$.*

PROOF. First, it is clear that $\bar{d}(u, v) \leq \tilde{d}(u, v) \leq d_{\partial G}(u, v)$. Now let $u = u_0, u_1, \ldots, u_m = v$ be the shortest path from u to v in $\tilde{G}$. Then we get the following:

$$d_{\partial G}(u, v) \leq \sum_{i=1}^{m} d_{\partial G}(u_{i-1}, u_i) \leq 2 \sum_{i=1}^{m} \tilde{d}(u_{i-1}, u_i) = 2\bar{d}(u, v)$$

Next we compute $\bar{d}$ by updating the weights $w(u, v)$ of $\tilde{G}$. First we compute new weights defined by $w'(u, v) = min_{x \in L(G) \setminus \mathcal{B}} w(u, x) + w(v, x)$ for $u, v \in \mathcal{B}$. This step runs in $O(|\mathcal{B}|^2 |L(G)|)$ time. If $w'(u, v) < w(u, v)$, then we replace $w(u, v)$ by $w'(u, v)$. Next we compute shortest paths for all pairs of nodes on the subgraph of $\tilde{G}$ induced by $\mathcal{B}$, and update the weights accordingly. This step runs in $O(|\mathcal{B}|^3)$ time. Since the shortest path between $u, v \in \mathcal{B}$ in $\tilde{G}$ cannot contain two consecutive nodes of $L(G) \setminus \mathcal{B}$, the weight $w(u, v)$ now corresponds to the distance $\bar{d}(u, v)$ for $u, v \in \mathcal{B}$. Finally we can compute $\bar{d}(u, v) = \min_{x \in \mathcal{B}} w(u, x) + w(x, v)$ for $u \in L(G) \setminus \mathcal{B}$ and $v \in \mathcal{B}$. This is correct, since the first node (after u) of the shortest path in $\tilde{G}$ between u and v must be in $\mathcal{B}$ (it can be v itself). This step runs in $O(|\mathcal{B}|^2 |L(G)|)$ time. Since $\mathcal{B} = O(1)$, the algorithm runs in linear time. $\square$

The metric $\bar{d}$ is the one for which we can compute the beacon-based embedding of [9] in linear time. Because $\bar{d}(u, v) \leq d_{\partial G}(u, v) \leq 2\bar{d}(u, v)$, the embedding also achieves constant multiplicative distortion with respect to the metric $d_{\partial G}(u, v)$, for all but an ε fraction of node-pairs in ∂G. Let f be this embedding, and let F be its extension, as described in Section 5.1. The embedding f essentially maps geodesic rays (or shortest paths) in $G = (V, E)$ to geodesic rays in $\mathbb{H}^k$. For $u \in \partial G$, let $A(u)$ be the set of nodes $v \in V$ such that v is on the geodesic ray corresponding to u. Note that v can be on multiple geodesic rays; in that case we can assign v to one of those rays arbitrarily. If f has constant multiplicative distortion on a pair (u_1, u_2) for $u_1, u_2 \in dG$, then F has constant additive distortion on all pairs in $A(u_1) \times A(u_2)$. So, if $A(u_1)$ and $A(u_2)$ are both very large, and (u_1, u_2) is a pair with bad distortion, we may get many pairs with bad distortion under F. To remedy this issue, we sample beacons based on the size of $A(u)$, i.e., the probability of choosing $u \in \partial G$ as a beacon is $|A(u)|/|V|$. We now obtain the following result.

THEOREM 5.4. *Given a δ-hyperbolic graph metric G with bounded degree, and any constant $\varepsilon > 0$, we can compute in linear time an embedding F of G into $\mathbb{H}^k$ such that, with high probability, F is a good additive embedding for all but an ε fraction of all node-pairs in G.*

PROOF. We use the algorithm described in Section 5.2. The running time directly follows from Lemma 5.3 and the discussion in Section 5.2. We need to show that F achieves constant additive distortion for all but an ε fraction of distance pairs in G. Let f be the embedding of ∂G into $\mathbb{R}^{k-1}$ used to construct F. Now imagine replacing each $u \in dG$ by the set $A(u)$ placed arbitrarily close to the original u in the metric space, without changing the doubling constant (we can use a metric representing points on a line for the distances among $A(u)$). Let $(\mathcal{X}, d)$ be the resulting metric space. We can run the beacon-based algorithm on $\mathcal{X}$; let f' be the corresponding embedding. Let $\mathcal{B}'$ be a sample set of beacons of $\mathcal{X}$ (chosen uniformly to construct f'), and let $\mathcal{B}$ be the corresponding sample set of beacons of ∂G (to construct f) such that $u \in \mathcal{B}$ if $A(u) \cap \mathcal{B}' \neq \emptyset$. Note that the probability of choosing $\mathcal{B}'$ is the same as the probability of choosing $\mathcal{B}$ (assuming we can choose the same beacon multiple times), since the probability of choosing $u \in \partial G$ as a beacon is $|A(u)|/|\mathcal{X}|$.

Let $u_1, u_2 \in \partial G$ where $u_1 \neq u_2$, and let $v_1 \in A(u_1)$ and $v_2 \in A(u_2)$. We know that all but an ε fraction of distance pairs (v_1, v_2) of $\mathcal{X}$ have constant multiplicative distortion under embedding f'. Let Δ be the multiplicative distortion of f' on pairs in $\mathcal{B}' \times \mathcal{X}$. By construction, we can assume that Δ is also the multiplicative distortion of f on pairs in $\mathcal{B} \times \partial G$. Recall that (v_1, v_2) has $O(\Delta)$ distortion under f' if there exists a $b' \in \mathcal{B}'$ such that $\min(d(v_1, b'), d(v_2, b')) \leq \beta d(v_1, v_2)$, for some $\beta \leq \frac{1}{4\Delta}$. By construction, it must also hold that there exists a $b \in \mathcal{B}$ such that $\min(d_{\partial G}(u_1, b), d_{\partial G}(u_2, b)) \leq \beta d_{\partial G}(u_1, u_2)$. Hence, the pair (u_1, u_2) of ∂G must have

$O(\Delta)$ multiplicative distortion under f. As a result, the corresponding pair (v_1, v_2) of G must have constant additive distortion under F. Thus, F is a good additive embedding for all but an ε fraction of all node-pairs in G. $\square$

Remark. An astute reader may have noticed that our embedding algorithm requires knowledge of δ of the graph G, just as the beacon-based Euclidean embedding of [9] requires knowledge of the doubling constant of the input metric. In practice, the value of δ is neither known nor computable in linear time, but empirical evidence suggests that it is typically small, and so randomly choosing a sufficiently large but constant number of beacons $\mathcal{B} \subseteq \partial G$ is enough for computing an embedding F of G into $\mathbb{H}^k$. Indeed, our mathematical framework can be viewed as lending theoretical justification for beacon-based hyperbolic-space embeddings, like the one implemented in [20].

6. CONCLUSION

In this paper, we explored the problem of embedding undirected graphs into hyperbolic space with low distortion. The wide-spread use of hyperbolic space for network analysis and routing, including low-distortion embedding of Internet scale graphs [17, 18, 20] and greedy routing [8], naturally begs a theoretical investigation of these phenomenon from a worst-case perspective. Our lower bounds give a definitive *negative* answer to the question: can all graphs be embedded in hyperbolic space with small distortion (multiplicative or additive)? On the positive side, we identified a fundamental measure of graphs, quasi-cyclicity, and showed that graphs with small distortion must also have small quasi-cyclicity. The structure of many graphs, including social networks and Internet, may naturally limit the quasi-cyclicity of those graphs, offering one possible explanation for their nice embedding into hyperbolic space. We also presented a simple and fast algorithm for embedding large bounded-degree graphs into hyperbolic space so that all but an ε fraction of the nodes have a constant additive distortion as long as the input graph is δ-hyperbolic, or has constant quasi-cyclicity.

Our work suggests a number of open problems and research directions. First, is the converse of Theorem 4.1 also true, i.e., is an (α, β)-quasi-cycle of size n always strictly $\Omega(n)$-hyperbolic? This can easily be shown for $\alpha > \frac{1}{2}$, but the problem is open for other values of α (and β). On a related note, Theorem 3.5 and its counterpart Theorem 4.2 are not exactly converses. Is it possible to obtain a tight upper bound on the distortion as a function of the quasi-cyclicity of a metric space? Finally, it would be interesting to evaluate how well our embedding algorithm works in practice. Can it improve upon existing, more direct algorithms to embed graphs into hyperbolic space, like the one implemented in [20]?

7. REFERENCES

[1] P. Assouad. Plongements lipschitziens dans $\mathbb{R}^n$. *Bulletin de la Société Mathématique de France*, 111:429–448, 1983.

[2] M. Bonk and O. Schramm. Embeddings of Gromov hyperbolic spaces. *Geometric & Functional Analysis*, 10(2):266–306, 2000.

[3] V. Chepoi, F. Dragan, B. Estellon, M. Habib, and Y. Vaxès. Diameters, centers, and approximating trees of δ-hyperbolic geodesic spaces and graphs. In *Proc. 24th Annual Symposium on Computational Geometry* (SoCG '08), pp. 59–68, 2008.

[4] J. Fakcharoenphol, S. Rao, and K. Talwar. A tight bound on approximating arbitrary metrics by tree metrics. *Journal of Computer and System Sciences*, 69(3):485–497, 2004.

[5] E. Ghys and P. de la Harpe (eds.). *Sur les groupes Hyperboliques d'après Mikhael Gromov*, volume 83 of *Progress in Mathematics*. Birkhäuser, Boston, 1990.

[6] M. Gromov. Hyperbolic groups. *Mathematical Sciences Research Institute Publications*, Springer, 8:75–263, 1987.

[7] W. B. Johnson and J. Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. *Contemporary Mathematics*, 26:189–206, 1984.

[8] R. Kleinberg. Geographic routing using hyperbolic space. In *Proc. 26th INFOCOM*, pp. 1902–1909, 2007.

[9] J. Kleinberg, A. Slivkins, and T. Wexler. Triangulation and embedding using small sets of beacons. *Journal of the ACM*, 56(6):32, 2009.

[10] B. Knaster, C. Kuratowski, and S. Mazurkiewicz. Ein Beweis des Fixpunktsatzes für n-dimensionale Simplexe. *Fundamenta Mathematicae*, 14:132–137, 1929.

[11] R. Krauthgamer and J. R. Lee. Algorithms on negatively curved spaces. In *Proc. 47th Annual IEEE Symposium on Foundations of Computer Science* (FOCS), pp. 119–132, 2006.

[12] D. Krioukov, F. Papadopoulos, A. Vahdat, and M. Boguna. Curvature and temperature of complex networks. *Physical Review E*, 80(3):035101(R), 2009.

[13] J. Matoušek. *Lectures on Discrete Geometry*. Springer-Verlag New York, Inc., 2002.

[14] F. Papadopoulos, D. V. Krioukov, M. Boguñá, and A. Vahdat. Greedy forwarding in dynamic scale-free networks embedded in hyperbolic metric spaces. In *Proc. 29th INFOCOM*, pp. 2973–2981, 2010.

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

[16] R. Sarkar. Low distortion Delaunay embedding of trees in hyperbolic plane. In *Proc. 19th International Symposium on Graph Drawing*, pp. 355–366, 2012.

[17] Y. Shavitt and T. Tankel. Big-bang simulation for embedding network distances in Euclidean space. *ACM Trans. on Networking*, 12(6):993–1006, 2004.

[18] Y. Shavitt and T. Tankel. Hyperbolic embedding of internet graph for distance estimation and overlay construction. *ACM Trans. on Networking*, 16(1):25–36, 2008.

[19] X. Zhao, A. Sala, C. Wilson, H. Zheng, and B. Y. Zhao. Orion: shortest path estimation for large social graphs. In *Proc. 3rd Workshop on Online Social Networks*, pp. 9–9, 2010.

[20] X. Zhao, A. Sala, H. Zheng, and B. Y. Zhao. Efficient shortest paths on massive social graphs. In *Proc. 7th International Conference on Collaborative Computing*, pp. 77–86, 2011.

APPENDIX

A. ADDITIONAL PROOFS

Lemma 3.2. *Let $p \in \mathbb{H}^k$ and let $P\colon [0,1] \to \mathbb{H}^k$ be a curve such that $\angle P(0)pP(1) = \alpha$ and $d_{\mathbb{H}}(p, P(t)) \geq R$ for $0 \leq t \leq 1$. Then the length of P is at least $\alpha \sinh(R)$.*

Proof. As for Euclidean space, we can define polar coordinates for $\mathbb{H}^k$, using the distance and angles w.r.t. a fixed point in $\mathbb{H}^k$. We need the formula to compute the arc length of a curve $P(t) = (r(t), \phi(t))$ $(0 \leq t \leq 1)$ in polar coordinates in $\mathbb{H}^2$, which is as follows:

$$|P| = \int_0^1 \sqrt{\left(\frac{dr}{dt}\right)^2 + \sinh^2(r(t)) \left(\frac{d\phi}{dt}\right)^2}\, \mathrm{d}t \tag{8}$$

Now consider the hyperbolic plane through p, $P(0)$, and $P(1)$. Choose polar coordinates (r, ϕ) for this hyperbolic plane such that p is at the center $(r = 0)$. We can extend these polar coordinates to $\mathbb{H}^k$, but the remaining coordinates can be ignored for our bound. If we write $P(t) = (r(t), \phi(t))$, then we obtain the claimed bound using Equation 8:

$$\begin{aligned}
|P| &\geq \int_0^1 \sqrt{\left(\frac{dr}{dt}\right)^2 + \sinh^2(r(t)) \left(\frac{d\phi}{dt}\right)^2}\, \mathrm{d}t \\
&\geq \int_0^1 \sqrt{\sinh^2(R) \left(\frac{d\phi}{dt}\right)^2}\, \mathrm{d}t \\
&= \sinh(R) \int_{\phi(0)}^{\phi(1)} \mathrm{d}\phi \\
&\geq \alpha \sinh(R)
\end{aligned}$$

□

Lemma A.1. *Consider the metric space $(S_n, \lambda d_S)$, where S_n is the star graph with center v_0 and leaves $v_1, \ldots, v_n$, and λ is a scaling factor of the standard metric. If there exists an embedding f of $(S_n, \lambda d_S)$ into $\mathbb{H}^k$ with only constant multiplicative distortion, then $\lambda k = \Omega(\log n)$.*

Proof. Let $p_i = f(v_i)$ be the embedded points in $\mathbb{H}^k$. Since f is non-expansive, we get that $d_{\mathbb{H}}(p_0, p_i) \leq \lambda$ for $1 \leq i \leq n$. If the embedding f has multiplicative distortion ρ, then the balls in $\mathbb{H}^k$ with radius $\frac{\lambda}{\rho}$ centered at the points p_i $(1 \leq i \leq n)$ must be interior disjoint. Furthermore, all these balls must lie inside the ball with radius 2λ centered at p_0. The volume of a ball with radius r in $\mathbb{H}^k$ is $V_r = G(k) \int_0^r \sinh^{k-1}(s)\, \mathrm{d}s$, where G is a function depending only on k. This can also be written as $V_r = G(k) 2^{\Theta(rk)}$. Since ρ is constant and n balls with radius $\frac{\lambda}{\rho}$ must fit in a ball with radius 2λ, there must be constants $c_2 > c_1 > 0$ such that the following inequality holds:

$$\begin{aligned}
G(k) n 2^{c_1 \lambda k} &\leq G(k) 2^{c_2 \lambda k} \\
\log n + c_1 \lambda k &\leq c_2 \lambda k
\end{aligned}$$

From the above inequality we can directly conclude that $\lambda k = \Omega(\log n)$. □