Bandwidth and Low Dimensional Embedding

Yair Bartal[1] ⋆, Douglas E. Carroll[2], Adam Meyerson[3], and Ofer Neiman[4]

[1] Institute of Computer Science, Hebrew University of Jerusalem, Israel. Email: yair@cs.huji.ac.il.
[2] Department of Computer Science, University of California, Los Angeles. Email: dcarroll@cs.ucla.edu.
[3] Department of Computer Science, University of California, Los Angeles. Email: awm@cs.ucla.edu.
[4] Princeton University and Center for Computational Intractability. Email: oneiman@princeton.edu.

Abstract. We design an algorithm to embed graph metrics into ℓ_p with dimension and distortion both dependent only upon the bandwidth of the graph. In particular we show that any graph of bandwidth k embeds with distortion polynomial in k into $O(\log k)$ dimensional ℓ_p, $1 \le p \le \infty$. Prior to our result the only known embedding with distortion independent of n was into high dimensional ℓ_1 and had distortion exponential in k. Our low dimensional embedding is based on a general method for reducing dimension in an ℓ_p embedding, satisfying certain conditions, to the intrinsic dimension of the point set, without substantially increasing the distortion. As we observe that the family of graphs with bounded bandwidth are doubling, our result can be viewed as a positive answer to a conjecture of Assouad [2], limited to this family. We also study an extension to graphs of bounded tree-bandwidth.

⋆ The work was done in part while the author was at the Center for the Mathematics of Information, Caltech, and the Institute for Pure and Applied Mathematics, UCLA.

1 Introduction

The problem of embedding graph metrics into normed spaces with low dimension and distortion has attracted much research attention (cf. [14]). In this paper we study the family of graphs with bounded bandwidth. The bandwidth of an unweighted graph $G = (V, E)$ is the minimal k such that there exists an ordering of the vertices in which the end points of every edge are at most k apart. Let d_G be the shortest path metric on the graph G. Let (Y, ρ) be a metric space, we say that an embedding $f : V \to Y$ has distortion $D \geq 1$ if there exists a constant $c > 0$ such that for all $x, y \in V$,

$$d_G(x, y) \leq c\rho(f(x), f(y)) \leq Dd_G(x, y) \ .$$

Our main result is the following.

Theorem 1. *For any integer $k \geq 1$ there exist $d = d(k)$ and $D = D(k)$ with the following property. For every $p \geq 1$ and graph $G = (V, E)$ with bandwidth at most k, there exists an embedding of (V, d_G) into ℓ_p space of dimension d with distortion D. In particular we have: $D(k) = O(k^2)$ and $d(k) = O(\log^2 k)$. Alternatively we also get: $D(k) = O(k^{2.001})$ and $d(k) = O(\log k)$.*

Our work is related to a conjecture of Assouad [2]. The doubling constant of a metric space is the minimal α such that any ball of radius r can be covered by α balls of half the radius, then the doubling dimension of V is defined as $\log_2 \alpha$. Assouad proved that for any metric (V, d), the "snow-flake" metric $(V, d^{1-\epsilon})$ embeds into Euclidean space with distortion and dimension depending only on on the doubling constant of (V, d) and on ϵ. Assouad conjectured that this is possible even when $\epsilon = 0$, but this was disproved by [19] (a quantitative bound was given by [12]). It is also shown in [12] that Assouad's conjecture holds for the family of doubling tree metrics. As the doubling constant of graphs with bandwidth k can be bounded by $O(k)$, one can view our result as providing a different family of doubling metrics for which Assouad's conjecture holds.

Graphs with low bandwidth play an important role in fast manipulation of matrices, in particular computing Gauss elimination and multiplication [10]. In his seminal paper Feige [11] showed an approximation algorithm for computing the bandwidth with poly-logarithmic guarantee. The bandwidth of a graph also plays a role in certain biological settings, such as gene clustering problems [20].

There has been a great deal of previous work on embedding families of graphs into ℓ_p with bounded distortion (for example [9,12,13,18,8]). The problem of embedding graphs of bounded bandwidth has been first tackled by [7]. They show that this family of graphs includes interesting instances which do not fall within any of the cases for which constant distortion embeddings are known. In their paper they show that bounded bandwidth graphs can be embedded into ℓ_1 [7] with distortion independent of the number of vertices n. However, the distortion of their embedding was exponential in the bandwidth k. Also, the dimension of that embedding was dependent on the number of vertices (in fact polynomial in n). We improve the result of [7] for graphs of bandwidth k

in several ways: First, our embedding works for any ℓ_p space ($1 \le p \le \infty$) as a target space, not just ℓ_1. Second, the distortion obtained is polynomial in k; specifically: $O(k^{2+\theta})$. Finally, we show that the dimension can be independent of n as well, and as low as $O((\log k)/\theta)$ (for any $0 < \theta < 1$).

Note that the fact that a graph has bandwidth k can be viewed as providing an embedding into 1 dimension with expansion bounded by k, but without any control on the contraction. Our result means that by increasing the dimension to $O(\log k)$, one can get a bound not only on the expansion but also on the contraction of the embedding.

The low dimensionality of our embedding follows from a generalization we give for a result of [1], who study embedding metric spaces in their intrinsic dimension. In [1] it is shown that for any n point metric space, with doubling constant α, there exists an embedding into ℓ_p space with distortion $O(\log^{1+\theta} n)$ and dimension $O((\log \alpha)/\theta)$ (for any $0 < \theta < 1$). Here, we extend their method in a way that may be applicable for reducing the dimension of embeddings in other settings. We show sufficient conditions on an embedding of any metric space (V, d) into ℓ_p (possibly high dimensional) with distortion γ, allowing to reduce the dimension to $O((\log \alpha)/\theta)$ with distortion only $O(\gamma^{1+\theta})$.

Our embedding for graphs of bandwidth k is obtained as follows: we first provide an embedding with distortion $O(k^2)$ which satisfies the conditions of the dimension reduction theorem. Our final embedding follows from the fact that the doubling constant of graphs of bandwidth k is $O(k)$.

It is worth noting that our embedding provides bounds independent of n for all $1 \le p \le \infty$. This is unusual: most previous non-trivial results for embedding infinite graph classes into normed spaces with constant distortion (independent of n) have ℓ_1 as a target metric [9,13] (and require high dimension). This is because of strong lower bounds indicating that trees have a distortion of $\Omega(\sqrt{\log\log n})$ [15] and tree-width two graphs have a distortion of $\Omega(\sqrt{\log n})$ when embedded into ℓ_2 [17]. Since bandwidth k graphs do not include all trees, these lower bounds will not apply and we are able to embed into ℓ_2 with constant (independent of n) distortion. We observe that ℓ_2 is potentially a more natural and useful target metric.

We extend our study to graphs of bounded tree-bandwidth [7] (see Definition 7 for precise definition). While this family of graphs includes all trees and thus requires distortion at least $\Omega(\sqrt{\log\log n})$ when embedded into ℓ_2, we are still able to apply our techniques with an additional overhead related to the embedability of trees. We provide an embedding of tree-bandwidth k graphs into ℓ_2 with distortion $O(\text{poly}(k)\sqrt{\log\log n})$ and into ℓ_1 with distortion polynomial in k. Moreover, when the graph has bounded doubling dimension we can apply our dimension reduction technique to achieve distortion and dimension depending solely on the doubling dimension and on k, utilizing the embedding of [12] for trees with bounded doubling dimension.

In general there has been a great deal of work on finding low-distortion embeddings. These embeddings have a wide range of applications in approximation algorithms, and in most cases low dimension is also desirable (for example improving the running time). Our work makes further progress towards achieving low-distortion results with dimension reduced to the intrinsic dimension. In particular, our embeddings imply better bounds in applications such as nearest neighbor search, distance labeling and clustering.

1.1 Summary of Techniques

The result of [1] includes the design of a specific embedding technique (locally padded probabilistic partitions), combined with the careful application of the Lovasz Local Lemma to show that it is possible to randomly merge the coordinates of this embedding in such a way that there is a non-zero probability that no distance is contracted. This can then be combined with constructive versions of the local lemma [3,16] to deterministically produce a low-dimensional embedding with no contraction.

We decouple the embedding technique of [1] from the local lemma, showing that any embedding technique which satisfies certain locality properties as well as having a single coordinate which lower bounds each particular distance can be applied in this way. Given any metric space (V, d) with doubling constant α, we give sufficient conditions to reduce the dimension of an embedding of V into ℓ_p with distortion γ to have dimension $O((\log\alpha\log\gamma)/\log(1/\epsilon))$ and distortion $O(\gamma/\epsilon)$ where $\gamma^{-1} < \epsilon < 1$. This approach allows some modularity in defining an embedding – if we are given a low distortion embedding (potentially much lower distortion than $\log n$ for some source metrics) which satisfies the locality properties then we can maintain the low distortion while obtaining low dimension as well.

In order to demonstrate the power of this approach, we apply it to the problem of embedding bounded bandwidth graphs into ℓ_p. We first need to define a low distortion embedding. The embedding of [7] is not useful for our purpose as it does not satisfy the necessary properties (in particular the single coordinate lower bound on distances fails) and because its distortion is undesirably high (exponential in bandwidth). Instead, we define a new embedding. The basis for our embedding is the standard scale based approach [18] using probabilistic partitions of [12,1] as a black box. The problem is that when using this approach we obtain an expansion factor of 1 at each scale of the embedding. The number of scales is logarithmic in the graph diameter, giving us a total expansion of $\Theta(\log n)$. The key innovation of our bandwidth embedding is showing that the number of scales can be reduced to $O(k)$.

Of course, for any scale there may be some point pair whose distance is at that scale (there are n^2 point pairs and only $\log n$ scales after all). We cannot simply remove some scales and expect our distortion to be reasonable. Instead, we compute a set of *active scales* for each graph vertex; these are the scales that represent distances to other points which are nearby in the optimum bandwidth ordering of the graph. We will reduce coordinate values to zero for vertices which do not consider the coordinate's scale to be active. Each vertex has only $O(k)$ active scales; the issue now is that different vertices have different scales and if two adjacent vertices have different active scales we might potentially introduce large expansion. In addition, we need to show that the critical coordinates which maintain the lower bound of $d(x, y)$ (thus preventing contraction) are active at one of the two points (x or y). Instead of applying active coordinates directly, we allow coordinates to decline gracefully by upper-bounding them by the distance to the nearest point where they are inactive, then use the bandwidth ordering to prove that the critical coordinates for preventing contraction are not only active where they need to be, but have not declined by too much to be useful.

A careful analysis of this construction shows that we can obtain distortion of $O(k^2)$. We also show that our modified embedding still posses the locality properties. Thus we can apply our dimension reduction technique to get dimension $O((\log k)/\theta)$ while maintaining the distortion bound up to a factor $O(k^\theta)$.

2 Embedding in the Doubling Dimension

2.1 Preliminaries and Definitions

Definition 1. *The* **doubling constant** *of a metric space* (V, d) *is the minimal integer* α *such that for any* $r > 0$ *and* $x \in V$*, the ball* $B(x, 2r)$ *can be covered by* α *balls of radius* r*. The* **doubling dimension** *or* **intrinsic dimension***, denoted by* $\dim(V)$*, is defined as* $\log \alpha$*.*

Suppose we are given a metric space (V, d) along with a randomly selected mapping $\phi : V \to \Re^D$ for some dimension D. For $1 \leq c \leq D$ we denote by $\phi_c(x)$ the c'th coordinate of $\phi(x)$ and thus we have $\phi_c : V \to \Re$. We may assume w.l.o.g that all coordinates of all points in the range of this mapping are non-negative.

Definition 2. *The mapping* ϕ *is* **single-coordinate** (ϵ, β) **lower-bounded** *if for every pair of points* $x, y \in V$ *there is some coordinate* c *such that* $|\phi_c(x) - \phi_c(y)| \geq \beta d(x, y)$ *with probability at least* $1 - \epsilon$*.*

In the metric embedding literature, we often speak of an embedding having contraction β. For ℓ_1 embedding, this means there is a set of coordinates whose sum is lower-bounded by $\beta d(x, y)$. The single-coordinate (ϵ, β) lower-bounded condition is stronger than contraction β, although for ℓ_p norms with large values of p (i.e. as p tends towards infinity) it becomes equivalent.

Definition 3. *Given a mapping* ϕ*, the* ℓ_1 **expansion** *of* ϕ *is* $\delta = \max_{x,y} \frac{\|\phi(x)-\phi(y)\|_1}{d(x,y)}$.

We observe that the expansion of ϕ when viewed as an ℓ_p embedding for $p > 1$ will be at most the ℓ_1 expansion. On the other hand, the single-coordinate (ϵ, β) lower-bound condition will still imply that the embedding has contraction β (for any pair of points with $1 - \epsilon$ probability).

Definition 4. *A mapping* ϕ *has the* **local property** *if for every coordinate* c *we can assign a scale* s_c *which is a power of two such that the following conditions hold:*

1. *For every* $x, y \in V$ *with* $d(x, y) > s_c$ *we have either* $\phi_c(x) = 0$ *or* $\phi_c(y) = 0$*.*

2. *For every $x, y \in V$, if there is a single-coordinate lower-bound for x, y, it has scale $\Omega(d(x,y)) < s_c < d(x,y)$.*

We observe that a mapping ϕ can be viewed as an embedding of (V,d) into normed space. Provided that the mapping is single-coordinate (ϵ, β) lower-bounded, we can eliminate contraction by repeatedly (and independently) selecting such mappings many times over and weighting the results by the number of selections, then multiplying all coordinates by $\frac{1}{\beta}$. This provides an embedding into ℓ_1 with distortion upper-bounded by $\frac{\delta}{\beta}$; note that this embedding can also be viewed as into ℓ_p for any $p > 1$ and in fact will have only lower distortion (the single-coordinate lower-bound condition still guarantees non-contraction).

2.2 Low Dimensional Embedding

An embedding ϕ maps (V,d) to potentially high dimensional space, and we are interested in *reducing the dimension* of such an embedding to resemble the doubling dimension of (V,d) without increasing the distortion. While for general ϕ such a result would imply dimension reduction for ℓ_1 (which is impossible in general [6]), the additional constraints that ϕ be single-coordinate lower-bounded and local will enable us to reduce the dimension. In the full version of the paper we prove the following generalization of [1].

Theorem 2. *Suppose we are given a metric space (V,d) with doubling constant α and a mapping $\phi : (V,d) \to \Re^D$ where ϕ is single-coordinate (ϵ,β) lower-bounded, local, and has ℓ_1 expansion at most δ for some $\beta/\delta \le \epsilon \le 1/8$. Then for any $1 \le p \le \infty$ we can produce in polynomial time an embedding $\tilde{\phi} : (V,d) \to \ell_p^m$ with distortion at most $O(\delta/(\epsilon\beta))$, where $m = O\left(\frac{\log \alpha \log(\delta/\beta)}{\log(1/\epsilon)}\right)$.*

Next we construct an embedding with the local property, which will serve as a basis embedding in Section 3. Recall that a partition $P = \{C_1, \ldots, C_n\}$ of an n-point metric space (V,d) is a pairwise disjoint collection of clusters (possibly some clusters are empty) which covers V, and $P(x)$ denotes the cluster containing $x \in V$. W.l.o.g we may assume that $\min_{x \ne y \in X}\{d(x,y)\} \ge 1$. The following lemma is a generalization of a lemma of [12] and was proven in [1].

Lemma 1. *For any metric space (V,d) with doubling constant α, any $0 < \Lambda < \mathrm{diam}(V)$ and $0 < \epsilon \le 1/2$ there exists a distribution $\hat{\mathcal{P}}$ over a set of partitions $\mathcal{P}$ such that the following conditions hold.*

- *For any $1 \le j \le n$, $\mathrm{diam}(C_j) \le \Lambda$.*
- *For any $x \in V$, $\Pr_{P \sim \hat{\mathcal{P}}}[B(x, \epsilon\Lambda/(64 \log \alpha)) \nsubseteq P(x)] \le \epsilon$.*

For every scale $s \in I = \{2^i \mid -1 \le i < \log(\mathrm{diam}(V)), i \in \mathbb{Z}\}$ let $P_s = \{C_1(s), \ldots, C_n(s)\}$ be a random partition sampled from $\hat{\mathcal{P}}$ with $\Lambda = s$, and let $c_1(s), \ldots c_n(s)$ be n coordinates that are assigned to the scale s. The random mapping is defined as

$$\phi_{c_j(s)}(x) = \begin{cases} d(x, V \setminus C_j(s)) & x \in C_j(s) \\ 0 & \textit{otherwise} \end{cases} \qquad (1)$$

and

$$\phi = \bigoplus_{s \in I, 1 \le j \le n} \phi_{c_j(s)} \tag{2}$$

Proposition 1. *For any* $0 < \epsilon \le 1/2$ *the mapping* ϕ *is single-coordinate* $(\epsilon, \epsilon/(128 \log \alpha))$ *lower-bounded, and its* ℓ_1 *expansion is at most* $O(\log(\mathrm{diam}(V)))$.

Proof. For any $x, y \in V$ let s be a power of two such that $s < d(x,y) \le 2s$, then in the coordinates assigned to scale s, the first property of Lemma 1 suggests that it must be that x, y fall into different clusters of the partition associated with the coordinates. Let j be such that $x \in C_j$, it follows that with probability $1 - \epsilon$, $\phi_{c_j(s)}(x) \ge \epsilon s/(64 \log \alpha) \ge \epsilon d(x,y)/(128 \log \alpha)$ and that with probability 1, $\phi_{c_j(s)}(y) = 0$.

To see that the ℓ_1 expansion is at most $2(\log(\mathrm{diam}(V)) + 2)$, note that the triangle inequality implies that $|\phi_{c_j(s)}(x) - \phi_{c_j(s)}(y)| \le d(x,y)$ for any $x, y \in V$ and $j \in [n]$, and since $\phi_{c_j(s)}(x)$ is non-zero for a single $j \in [n]$ it follows that for any $s \in I$

$$\Sigma_{1 \le j \le n} |\phi_{c_j(s)}(x) - \phi_{c_j(s)}(y)| \le 2d(x,y), \tag{3}$$

and hence

$$\Sigma_{s \in I, 1 \le j \le n} |\phi_{c_j(s)}(x) - \phi_{c_j(s)}(y)| \le \Sigma_{s \in I} 2d(x,y) = 2(\log(\mathrm{diam}(V)) + 2) d(x,y) \ .$$

Proposition 2. *The mapping* ϕ *has the local property.*

Proof. The first local property is immediate by the first property of Lemma 1 and by (1). The second local property follows from the proof of Proposition 1.

3 Low Distortion ℓ_p-embeddings of Low Bandwidth Graphs

3.1 Preliminaries and Definitions

Definition 5. *Given graph* $G = (V, E)$ *and linear ordering* $f : V \to \{1, 2, ..., |V|\}$ *the* bandwidth *of* f *is* $max\{|f(v) - f(w)| \mid (v, w) \in E\}$. *The* bandwidth *of* G *is the minimum bandwidth over all linear orderings* f. *Given an optimal bandwidth ordering* f, *the index of* u *is simply* $f(u)$.

Definition 6. *Define* $\lambda(x, y) = |f(x) - f(y)|$ *which is the distance between* x, y *in the bandwidth ordering* f *of* G.

In what follows we are given a graph G of bandwidth k, the metric space associated with G is the usual shortest-path metric, and we assume we are given the optimal ordering f obtaining this bandwidth. This ordering is computable in time exponential in k, and since our embedding only improves upon previous work (for example Bourgain [4]) when k is quite small, it may be reasonable to assume that the ordering is given. In general computing the best bandwidth ordering is NP-Hard, and the best approximations are poly-logarithmic in n [11].

Proposition 3. *Let G be a graph of bandwidth k. Then there exists an ordering where for any $x, y \in G$, $\lambda(x, y) \leq k \cdot d(x, y)$.*

Proof. Assume $d(x, y) = r$, and let $P_{xy} = (x = v_0, v_1, \ldots, v_r = y)$ be a shortest path in G connecting x and y, then by the triangle inequality $\lambda(x, y) \leq \Sigma_{i=1}^{r} \lambda(v_{i-1}, v_i)$. By the definition of bandwidth for all $1 \leq i \leq r$, $\lambda(v_{i-1}, v_i) \leq k$, hence the proposition follows.

Proposition 4. *If $G = (V, E)$ has bandwidth k, then the doubling constant α of G is at most $4k + 1$.*[5]

Proof. Consider the ball of radius $2r$ about some point $x \in V$. We must show that this ball can be covered by at most $4k + 1$ balls of radius r.

Consider any integer $0 < a \leq r$. Let Y_a be the set of points y such that $d(x, y) = a$; similarly let Y_{a+r} be the set of points y such that $d(x, y) = a + r$. We claim that the set of balls of radius r centered at points in $\{x\} \cup Y_a \cup Y_{a+r}$ covers the ball of radius $2r$ around x. In particular, consider any point z in this ball. If $d(x, z) \leq r$ then $z \in B(x, r)$. If $a \leq d(x, z) < a + r$ then there is some shortest path from x to z of length $d(x, z)$ which must include a point y with $d(x, y) = a$ and $d(y, z) = d(x, z) - a < r$. It follows that $z \in B(y, r)$ and that $y \in Y_a$. If $a + r \leq d(x, z) < 2r$ then again there is a shortest path from x to z of length $d(x, z)$ which must include a point y with $d(x, y) = a + r$ and $d(y, z) = d(x, z) - a - r < r$. It follows that $z \in B(y, r)$ and $y \in Y_{a+r}$.

Now consider the various sets $\{x\} \cup Y_a \cup Y_{a+r}$ as we allow a to range from 1 to r. With the exception of x, these sets are disjoint for distinct values of r. So every point in $B(x, 2r)$ other than x appears exactly once. It follows that there must be some choice of a such that the size of this set is only $1 + \frac{1}{r}|B(x, 2r)|$. Since the graph G has bandwidth k, it follows that any pair of adjacent nodes are within k of each other in the bandwidth ordering. So all points in $B(x, 2r)$ are within $2rk$ of x in the ordering, and thus there are at most $4rk$ such points. From this it follows that we need only $4k + 1$ balls to cover $B(x, 2r)$.

The remainder of this section will be devoted to proving our main theorem, that graphs of bounded bandwidth embed into ℓ_p with low dimension and distortion.

Theorem 3. *Let G be a graph with bandwidth k and let $0 < \theta < 1$, then for any $p \geq 1$, there exists an embedding of G into ℓ_p with distortion $O(k^{2+\theta})$ and dimension $O((\log k)/\theta)$.*

3.2 Proof of Theorem 3

Consider the mapping ϕ defined in (2). By Proposition 1 combined with Theorem 2 (noting that for unweighted graphs we get $O(\log n)$ ℓ_1 expansion) we can transform it into an embedding of a graph with bandwidth k into any ℓ_p space of

[5] A somewhat simpler argument could be applied to give an $O(k)$ bound on the doubling constant, which would suffice for our application. The argument presented here seems to give a better estimate on the constant.

dimension $O((\log k)/\theta)$ with distortion $O(\log^{1+\theta} n)$ for any $0 < \theta \leq 1$. Our main innovation is to reduce the number of scales effecting each of the points, thereby reducing the overall distortion to $O(k^2)$.

Let G be a graph with bandwidth k and f be the optimal ordering obtaining this bandwidth. Let $\alpha \leq 4k+1$ be the doubling constant of G. For each scale s, we will say that scale s is *active* at point x if there exists a y such that $\lambda(x, y) \leq k$ and $s/8 \leq d(x, y) \leq 4s$. We define $h_s(x)$ to be the distance from x to the nearest point z for which s is not active (note that $h_s(x) = 0$ if s is not active at x). We then define a mapping $\hat{\phi}$ as follows (recall the definition of s_c in Definition 4):

$$\hat{\phi}_c(x) = \min(\phi_c(x), h_{s_c}(x))$$

We will claim that for suitable values of ϵ, this $\hat{\phi}$ is single-coordinate $(\epsilon, \frac{1}{k})$ lower-bounded for all point pairs x, y with $|f(x) - f(y)| \leq \frac{1}{4}d(x, y)$, that it is local, and that it has ℓ_1 expansion bounded by $O(k)$. The final embedding will be $\hat{\phi}$ concatenated with an extra coordinate f, which is the location of the points in the bandwidth ordering. This will allow us to apply Theorem 2 without the f coordinate, then add in the f coordinate to get our final embedding.

Lemma 2. *The mapping $\hat{\phi}$ has ℓ_1 expansion at most $O(k)$.*

Proof. Consider any pair of points x, y. We observe that the total number of scales which are *active* for these two points is at most $O(k)$, this is because for x there are at most $2k - 1$ other points z satisfying $\lambda(x, z) \leq k$, and each of these points may activate at most 6 different scales. We conclude that there are at most $O(k)$ non-zero coordinates for these two points. So the ℓ_1 expansion expression has only $O(k)$ non-zero terms. Let c be a non-zero coordinate. The triangle inequality suggests that each coordinate produces expansion of at most 1 in ϕ, that is

$$\phi_c(x) - \phi_c(y) \leq d(x, y)$$

If $\hat{\phi}_c(y) = \phi_c(y)$ then since $\hat{\phi}_c(x) \leq \phi_c(x)$, we can write:

$$\hat{\phi}_c(x) - \hat{\phi}_c(y) \leq \phi_c(x) - \phi_c(y) \leq d(x, y)$$

On the other hand, suppose that $\hat{\phi}_c(y) = h_s(y)$ where $s = s_c$. Then there is some z where scale s is inactive, such that $h_s(y) = d(y, z)$. Now

$$\hat{\phi}_c(x) - \hat{\phi}_c(y) \leq h_s(x) - h_s(y) \leq d(x, z) - d(y, z) \leq d(x, y)$$

From this we conclude that each non-zero coordinate produces expansion at most 1, and when we total this over $O(k)$ non-zero coordinates we get total ℓ_1 expansion at most $O(k)$.

We note that adding the coordinate f does not increase the expansion by much. In particular, for any point pair x, y we have $|f(x) - f(y)| \leq kd(x, y)$

by Proposition 3. So the extra coordinate increases expansion by at most an additive k.

The tricky part is proving that the mapping is single-coordinate $(\epsilon, \frac{1}{k})$ lower-bounded. Given some pair of points x, y, one might imagine that the critical coordinates were deemed *inactive* for x and y, and thus the single-coordinate lower-bound will no longer hold. We will prove that this is not the case.

Lemma 3. *For any $k^{-1/2} \leq \epsilon \leq 1/2$ the mapping $\hat{\phi}$ is single-coordinate $(\epsilon, \Omega(\frac{1}{k}))$ lower-bounded for any pair of points x, y with $|f(x) - f(y)| \leq \frac{1}{4}d(x, y)$.*

Proof. Consider any pair of points x, y with $|f(x) - f(y)| \leq \frac{1}{4}d(x, y)$. Let $x' \in B(x, \frac{d(x,y)}{8k})$ and $y' \in B(y, \frac{d(x,y)}{8k})$. Let s be the scale such that $d(x, y)/2 \leq s < d(x, y)$. We will show that scale s must be *active* at x' or at y'. But this holds for *any pair* of points x', y' from the appropriate balls around x, y. It follows that for one of these two balls it must be the case that scale s is active at *all points in the ball*. Suppose without loss of generality that this is $B(x, \frac{d(x,y)}{8k})$. Then since all points in this ball have scale s active, we conclude that $h_s(x) \geq \frac{d(x,y)}{8k}$. By Proposition 1 and the local property of ϕ, there is a coordinate c assigned to scale s, which with $1 - \epsilon$ probability, has $\phi_c(x) \geq \Omega(\frac{\epsilon}{\log \alpha})d(x, y)$ and by the first local property of ϕ also $\phi_c(y) = 0$. If this event occurs, then since $\Omega\left(\frac{\epsilon}{\log \alpha}\right) \geq \Omega\left(\frac{k^{-1/2}}{2 \log k}\right) \geq \Omega(1/k)$, we get that $\hat{\phi}_c(x) \geq d(x, y) \min(\Omega(\frac{\epsilon}{\log \alpha}), \frac{1}{8k}) \geq \Omega(\frac{d(x,y)}{k})$, and of course $\hat{\phi}_c(y) = 0$. We conclude that x, y are $(\epsilon, \Omega(1/k))$ lower bounded.

In the remainder of proof we show that indeed scale s must be active at either x' or y'. Since $d(x, x') \leq d(x, y)/(8k)$ and $d(y, y') \leq d(x, y)/(8k)$ it follows that $d(x', y') \geq d(x, y)(1 - \frac{1}{4k}) \geq \frac{3}{4}d(x, y)$. On the other hand, $|f(x) - f(x')| \leq \frac{d(x,y)}{8}$ and similarly for $|f(y) - f(y')|$ from which we can conclude that $|f(x') - f(y')| \leq \frac{1}{2}d(x, y)$. Now consider a fixed shortest path from x' to y'. Assume without loss of generality that $f(x') < f(y')$. We define two special points along this path as follows:

- $\tilde{x}$ is the first point on the path from x' to y' such that for all points z subsequent to or equal to $\tilde{x}$ on the path, we have $f(z) \geq f(x')$.
- $\tilde{y}$ is the first point on the path from $\tilde{x}$ to y' with $f(\tilde{y}) \geq f(y')$

These points will be auxiliary points showing that scale s is active at either x' or y'. For instance to show that scale s is active at x' it is enough to show that $\lambda(x', \tilde{x}) \leq k$ and that $s/8 \leq d(x', \tilde{x}) \leq 4s$. We observe that because any pair of consecutive vertices on a path are at most k apart in the bandwidth ordering, it must be that $|f(x') - f(\tilde{x})| \leq k$ and $|f(y') - f(\tilde{y})| \leq k$. Note that for every point z on the path from $\tilde{x}$ to $\tilde{y}$, the value of $f(z)$ is a unique point between $f(x')$ and $f(y')$. We conclude that $d(\tilde{x}, \tilde{y}) \leq |f(x') - f(y')| \leq \frac{1}{2}d(x, y)$. Since these points are on the shortest path, we know that $d(x', y') = d(x', \tilde{x}) + d(\tilde{x}, \tilde{y}) + d(\tilde{y}, y')$. It follows that either $d(x', \tilde{x}) \geq \frac{1}{8}d(x, y)$ or $d(\tilde{y}, y') \geq \frac{1}{8}d(x, y)$.

On the other hand, it is not hard to see that $d(x', \tilde{x}) \leq d(x', y') \leq d(x, y) + \frac{1}{4k}d(x, y) \leq 2d(x, y)$ and similarly for $d(y', \tilde{y})$. We conclude that indeed scale s must be active for one of x', y'.

Lemma 4. *The mapping $\hat{\phi}$ is local.*

Proof. The first condition follow immediately from the fact that ϕ is local and $\hat{\phi}_c(x) \leq \phi_c(x)$ for all c and x. The second condition follow from Lemma 3.

We now combine the lemmas and apply Theorem 2 to $\hat{\phi}$. This guarantees bounded contraction for point pairs with $|f(x)-f(y)| \leq \frac{1}{4}d(x,y)$. We add the single additional coordinate f, and this guarantees bounded contraction for points with $|f(x)-f(y)| \geq \frac{1}{4}d(x,y)$. Choosing for any $0 < \theta < 1$, $\epsilon = k^{-\theta}$ will give distortion $O(k^{2+\theta})$ and dimension $O((\log k)/\theta)$.

4 Tree Bandwidth

We will give an embedding of a graph of low tree-bandwidth [7] into ℓ_p. The distortion will be polynomial in k, with a multiplicative $O(\sqrt{\log\log n})$ term for $p > 1$ [5]. This improves upon the result of [7] by reducing the distortion and extending to ℓ_p. All the proofs appear in the full version of the paper.

Definition 7. *[[7]] Given a graph $G = (V, E)$, we say that it has* **tree-bandwidth** *k if there is a rooted tree $T = (I, F)$ and a collection of sets $\{X_i \subset V | i \in I\}$ such that: $\forall i, |X_i| \leq k$, $V = \bigcup X_i$, the X_i are disjoint, $\forall (u,v) \in E$, u and v lie in the same set X_i or $u \in X_i$ and $v \in X_j$ and $(i,j) \in F$, and if i has parent $p(i)$ in T, then $\forall v \in X_i, \exists u \in X_{p(i)}$ such that $d(u,v) \leq k$. T is called the* decomposition tree *of G.*

Theorem 4. *There is a randomized algorithm to embed tree-bandwidth k graphs into ℓ_p with expected distortion $O(k^3 \log k + k\rho)$ where ρ is the distortion for embedding the decomposition tree into ℓ_p.*

In the case of ℓ_1, there is a simple embedding of a tree with $\rho = 1$. For ℓ_2, the bound of [5] ensures $\rho = O(\sqrt{\log\log n})$.

We can also apply Theorem 2 to reduce the dimension the embedding of Theorem 4. To do this we need to bound the dimension in which the tree can be embedded. We have the following lemma,

Lemma 5. *Let G be a graph with tree bandwidth k, and let α be the doubling constant of G, then the doubling dimension of the decomposition tree T for G is $\log \alpha_T = O((\log \alpha)(\log k))$.*

It follows that we can use an embedding for the decomposition tree T of G where the distortion and dimension are functions of the doubling dimension of T, as shown in [12], and therefore are a function of α and k alone.

Theorem 5. *Suppose that we are given a tree-bandwidth k graph along with its tree decomposition. Let the doubling constant of this graph be α. Let α_T be the doubling constant of T, given by Lemma 5. Further, suppose that there exists an embedding of the tree decomposition into $d(\alpha_T)$ dimensional ℓ_p with distortion $\rho(\alpha_T)$. Then for any $0 < \theta < 1$ there is an embedding of the graph into ℓ_p with expected distortion $O(k^{3+\theta}\log k + k\rho(\alpha_T))$ and dimension $O((\log \alpha)/\theta + d(\alpha_T))$.*

References

1. I. Abraham, Y. Bartal, and O. Neiman. Embedding metric spaces in their intrinsic dimension. In *SODA '08: Proceedings of the 18th ann. ACM-SIAM sym. on Discrete algorithms*, 2008.
2. P. Assouad. Plongements lipschitziens dans $\mathbb{R}^n$. *Bull. Soc. Math. France*, 111(4):429–448, 1983.
3. J. Beck. An algorithmic approach to the lovasz local lemma. *Random Struct. Algorithms*, 2:343–365, 1991.
4. J. Bourgain. On Lipschitz embedding of finite metric spaces in Hilbert space. *Israel J. Math.*, 52(1-2):46–52, 1985.
5. J. Bourgain. The metrical interpretation of superreflexivity in Banach spaces. *Israel J. Math.*, 56(2):222–230, 1986.
6. B. Brinkman and M. Charikar. On the impossibility of dimension reduction in l_1. In *FOCS*, pages 514–523, 2003.
7. D. E. Carroll, A. Goel, and A. Meyerson. Embedding bounded bandwidth graphs into ℓ_1. *Proceedings of the 33rd Ann. ICALP*, pages 27–37, 2006.
8. A. Chakrabarti, A. Jaffe, J. R. Lee, and J. Vincent. Embeddings of topological graphs: Lossy invariants, linearization, and 2-sums. In *Proceedings of the 49th Ann. IEEE Sym. on Foundations of Computer Science*, pages 761–770, 2008.
9. C. Chekuri, A. Gupta, I. Newman, Y. Rabinovich, and A. Sinclair. Embedding k-outerplanar graphs into ℓ_1. In *SODA '03: Proceedings of the 14th ann. ACM-SIAM sym. on Discrete algorithms*, pages 527–536, 2003.
10. P. Z. Chinn, J. Chvtalov, A. K. Dewdney, and N. E. Gibbs. The bandwidth problem for graphs and matrices - a survey. *Journal of Graph Theory*, (6):223–254, 1982.
11. U. Feige. Approximating the bandwidth via volume respecting embeddings. In *Proceedings of the 30th ann. ACM sym. on Theory of computing*, STOC '98, pages 90–99, New York, NY, USA, 1998. ACM.
12. A. Gupta, R. Krauthgamer, and J.s R. Lee. Bounded geometries, fractals, and low-distortion embeddings. In *Proceedings of the 44th Ann. IEEE Sym. on Foundations of Computer Science*, page 534, 2003.
13. A. Gupta, I. Newman, Y. Rabinovich, and A. Sinclair. Cuts, trees and ℓ_1-embeddings of graphs. In *Proceedings of the 40th Ann. Sym. on Foundations of Computer Science*, pages 399–409, 1999.
14. N. Linial, E. London, and Y. Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995.
15. J. Matousek. On embedding trees into uniformly convex banach spaces. *Israel Journal of Mathematics*, 114:221–237, 1999.
16. R. A. Moser and G. Tardos. A constructive proof of the general lovász local lemma. *J. ACM*, 57(2):1–15, 2010.
17. Ilan Newman and Yuri Rabinovich. A lower bound on the distortion of embedding planar metrics into euclidean space. In *SCG '02: Proceedings of the 18th ann. sym. on Computational geometry*, pages 94–96. ACM Press, 2002.
18. S. Rao. Small distortion and volume preserving embeddings for planar and Euclidean metrics. In *Proceedings of the 15th Ann. Sym. on Computational Geometry*, pages 300–306, New York, 1999. ACM.
19. S. Semmes. On the nonexistence of bilipschitz parameterizations and geometric problems about a_∞ weights. *Revista Matemática Iberoamericana*, 12:337–410, 1996.
20. Q. Zhu, Z. Adam, V. Choi, and D. Sankoff. Generalized gene adjacencies, graph bandwidth, and clusters in yeast evolution. *IEEE/ACM Trans. Comput. Biol. Bioinformatics*, 6(2):213–220, 2009.