On Euclidean Embeddings and Bandwidth Minimization

John Dunagan and Santosh Vempala

Department of Mathematics, MIT, Cambridge MA, 02139.
{jdunagan, vempala}@math.mit.edu
Supported in part by NSF Career Award CCR-9875024.

Abstract. We study Euclidean embeddings of Euclidean metrics and present the following four results: (1) an $O(\log^3 n\sqrt{\log\log n})$ approximation for minimum bandwidth in conjunction with a semi-definite relaxation, (2) an $O(\log^3 n)$ approximation in $O(n^{\log n})$ time using a new constraint set, (3) a lower bound of $\Theta(\sqrt{\log n})$ on the least possible volume distortion for Euclidean metrics, (4) a new embedding with $O(\sqrt{\log n})$ distortion of point-to-subset distances.

1 Introduction

The minimum bandwidth problem asks for a permutation of the vertices of an undirected graph that minimizes the maximum difference between the endpoints of its edges. This maximum difference is called the bandwidth. Minimizing the bandwidth is NP-hard [5].

The question of finding good *approximations* to the minimum bandwidth has led to two different Euclidean embeddings of graphs. One of them is obtained as a solution to a semi-definite relaxation of the problem [1]. The other is an embedding that preserves *tree volumes* of subsets of a given metric [3, 4]. The tree volume of a subset is the product of the edge lengths of a minimum spanning tree of the subset. These two embeddings were used in separate approximation algorithms.

In this paper we combine the two embeddings to obtain an improved approximation. The quality of the approximation is $O(\rho\log^{2.5} n)$ where ρ is the best possible volume distortion (defined in section 2) of a Euclidean metric. Using Rao's upper bound [6] of $O(\sqrt{\log n\log\log n})$ on ρ we obtain an approximation guarantee of $O(\log^3 n\sqrt{\log\log n})$ which improves on [4] by a factor of $\Theta(\sqrt{\log n})$.

Our approach immediately leads to the question of whether a better upper bound is possible. In section 5, we show a lower bound of $\Omega(\sqrt{\log n})$ on the volume distortion even for the path graph. Thus further improvements to bandwidth approximation will have to come from other avenues.

We then turn to the general question of embedding metrics in Euclidean space. Finding an embedding of a metric that "preserves" properties of the original metric is a classical problem. A natural property to consider in this regard is the original distance function itself. J. Bourgain [2] gave an embedding that achieves a *distortion* of $O(\log n)$ for any metric on n points, i.e. the distance between points in the embedding is within a factor of $O(\log n)$ of their distance in the metric. In other words, it "preserves" distances any point and any another point. A natural generalization would be an embedding that preserves the distance between any point and any subset of points. The distance of a point u to a subset of points S, is simply the distance of u to the closest point in S. For a Euclidean embedding, the distance of a point u to a subset S can be defined as the Euclidean distance from u to the *convex hull* of S. Thus point-to-subset distance is a direct generalization of point-to-point distance for metrics as well as for points in Euclidean space. In section 6, we give an embedding whose point-to-subset distortion is $O(\sqrt{\log n})$ for any Euclidean metric where the shortest distance is within a $poly(n)$ factor of the longest distance.

Replacing "convex" in the definition above by "affine" leads to another interesting property. In section 6 we observe that for any Euclidean embedding, the distortion of affine point-to-subset distances is also an upper bound on its volume distortion. In section 7, we formulate a new system of constraints that are separable in $O(n^{\log n})$ time, and which result in an $O(\log^3 n)$ approximation to the minimum bandwidth using the results of section 6. We conclude with the conjecture that our embedding (section 6.1) achieves the optimal volume distortion for Euclidean metrics.

2 Euclidean Embeddings of Metrics

Let $G = (V, E)$ be a finite metric with distance function $d(u, v)$. We restrict our attention throughout the paper to Euclidean embeddings ϕ of G that are contractions, i.e. the distances between embedded points are at most the original distances. As mentioned in the introduction, the *distortion* of a contraction embedding, $\phi(G)$, is

$$\max_{u,v \in V} \frac{d(u,v)}{|\phi(u) - \phi(v)|}$$

where $|\cdot|$ is the Euclidean distance (L_2 norm). A Euclidean metric on n points is a metric that is exactly realizable as the distances between n points in Euclidean space.

The *Tree Volume (Tvol)* of a metric is the product of the edge lengths of the minimum spanning tree. A subset S of a metric also induces a metric, and its tree volume, $Tvol(S)$ is the product of the edges of the minimum spanning tree of the metric induced by S.

The *Euclidean Volume (Evol)* of a subset of points $\{x_1, \ldots, x_k\}$ in some Euclidean space is the volume of the $(k-1)$-dimensional simplex spanned by the points.

Definition 1. *The* k**-volume distortion** *of a contraction embedding* ϕ *is defined as*

$$\max_{S \subseteq V, |S|=k} \left(\frac{Tvol(S)}{(k-1)!Evol(\phi(S))} \right)^{\frac{1}{k-1}}$$

Remark. The factor $(k-1)!$ in the denominator is a normalization that connects volume of a simplex to volume of a parallelepiped. The distortion as defined above is within a factor of 2 of the distortion as defined in [3, 4]. We find it unnecessary to go through the notion of "best possible volume" (Vol) used there.

The following theorem, due to Rao [6], connects the tree volume with Euclidean volume for the special case of Euclidean metrics.

Theorem 1. *For any Euclidean metric G, there exists a Euclidean embedding $\phi(G)$ whose k-volume distortion is $O(\sqrt{\log n \log\log n})$ for all k up to $\log n$.*

3 A Semi-Definite Relaxation

To arrive at the semi-definite relaxation of [1], we can start by imagining that the points of the graph are arranged along a great circle of the sphere at regular intervals spanning an arc of 90 degrees as in Figure 1. This is our approximation of laying out all the points on a line, and it is good to within a factor of 2. We relax this to allow the points to wander around the sphere, but maintaining that no two lie more than 90 degrees apart, and that they satisfy the "spreading" constraints. The objective function for our relaxation is to minimize the maximum distance between any pair of points connected by an edge in the original graph. We now give the SDP explicitly, where $G = (V, E)$ is our original graph. Note that G does not neccesarily induce a Euclidean metric, but the solution to the SDP below, where the vectors correspond to vertices of G, does induce a Euclidean metric. It is shown in [1] that this is a relaxation and that it can be solved in polytime.

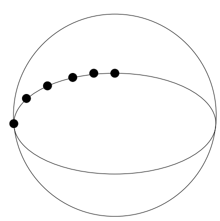

Fig. 1. Not quite points on a line, but close.

$$\begin{aligned} \min\ & b \\ u_i \cdot u_j \geq 0 & \quad \forall i, j \in V \\ |u_i| = n & \quad \forall i \in V \\ |u_i - u_j| \leq b & \quad \forall (i, j) \in E \\ \sum_{j \in S} (u_i - u_j)^2 \geq \frac{1}{12}|S|^3 & \quad \forall S \subseteq V, \forall i \in V \end{aligned}$$

4 A Rounding Algorithm

Let the Euclidean embedding obtained by solving the relaxation be $U = \{u_1, \ldots, u_n\}$. The algorithm below rounds this solution to an ordering of the vertices of G.

1. Find a volume respecting embedding of U, $\phi(U) = \{v_1, \ldots, v_n\}$, using Rao's algorithm [6] with $k = \log n$.
2. Pick a random line ℓ passing through the origin.
3. Project the points of the embedding $\phi(U)$ to ℓ and output the ordering obtained.

Denote the dimension of the embedding $\phi(U)$ by d. Upon random projection, edge lengths shrink by a factor of $\frac{1}{\sqrt{d}}$ in expectation. To analyze the quality of the approximation we obtain, we show that every edge shrinks by at least a factor of $\frac{\sqrt{\log n}}{\sqrt{d}}$, and that not too many points fall in any interval of length $\frac{1}{\sqrt{d}}$. To show that no more than m points fall in an interval, we show that no more than $\binom{m}{k}$ sets of k points fall in the interval.

We will use the following lemmas. Lemma 1 is from [4] and lemmas 2 and 3 are from [7].

Lemma 1.

$$\sum_{S \subset U, |S|=k} \frac{1}{Tvol(S)} \leq n(\log n)^{k-1}$$

Lemma 2. *Let $v \in \mathbb{R}^d$. For a random unit vector ℓ,*

$$\Pr\left[|v \cdot \ell| \leq \frac{c}{\sqrt{d}}|v|\right] \geq 1 - e^{-c^2/4}.$$

Lemma 3. *Let S be a set of vectors $v_1, \ldots, v_k \in \mathbb{R}^d$. For a random unit vector ℓ*

$$\Pr\left[max_i\{v_i \cdot \ell\} - min_i\{v_i \cdot \ell\} \leq W\right] = O\left(\frac{W^{k-1} d^{\frac{k-1}{2}}}{(k-1)! Evol(S)}\right)$$

Lemma 4. *After random projection, the number of vertices that fall in any interval of length $\frac{1}{\sqrt{d}}$ is $O(\rho \log^2 n)$, where ρ is the k-volume distortion of the embedding.*

Proof. Consider an interval of length $W = \frac{1}{\sqrt{d}}$. For a subset S of V, let X_S be a random variable that is 1 if all the vectors in S fall in the interval. Let us estimate the total number of sets S of size k that fall in the interval.

$$\mathbf{E}\left[\sum_{|S|=k} X_S\right] = \sum_{|S|=k} \mathbf{E}[X_S] \tag{1}$$

$$= \sum_{|S|=k} \Pr(X_S = 1)$$

$$\leq \sum_{|S|=k} \frac{W^k d^{\frac{k}{2}}}{(k-1)!Evol(S)} \tag{2}$$
$$= \sum_{|S|=k} \frac{1}{(k-1)!Evol(S)}$$
$$\leq \sum_{|S|=k} \frac{(\rho)^{k-1}}{Tvol(S)} \tag{3}$$
$$\leq (\rho)^k \sum_{|S|=k} \frac{1}{Tvol(S)}$$
$$\leq (\rho)^k n(\log n)^k \tag{4}$$
$$\leq (2\rho \log n)^k \tag{5}$$

Step 2 is from lemma 3, step 3 is an application of theorem 1, step 4 is from lemma 1, and step 5 follows from $k = \log n$.

We need to consider only $O(n\sqrt{\log n})$ intervals of length $\frac{1}{\sqrt{d}}$ (the longest distance is $O(n)$ originally, and by lemma 2, it maps to a distance of at most $O(n\sqrt{\log n})$ with high probability). Now by Markov's inequality, with high probability, the number of k subsets that fall in any interval of length $\frac{1}{\sqrt{d}}$ is at most $n^2(2\rho \log n)^k \leq (8\rho \log n)^k$. Thus if the number of points in such an interval is m then $\binom{m}{k} \leq (8\rho \log n)^k$ which implies that $m = O(\rho \log^2 n)$ (using $k = \log n$). □

Theorem 2. *The algorithm finds an* $O(\rho \log^{2.5} n) = O(\log^3 n\sqrt{\log \log n})$ *approximation with high probability.*

Proof. Consider an edge (i, j) in the original graph that is mapped to vectors v_i and v_j after the volume-preserving embedding. Then $\max_{(i,j)\in E} |v_i - v_j|$ is a lower bound on the bandwidth of the graph (the distance between the solution vectors of the SDP is a lower bound and this distance is only contracted during the volume-preserving embedding).

After the last step, with high probability the distance between the projections of v_i and v_j is at most $O(\frac{\sqrt{\log n}}{\sqrt{d}}|v_i - v_j|)$ for *every* pair (i, j) (lemma 2).

Thus the maximum number of intervals of length $\frac{1}{\sqrt{d}}$ any edge (i, j) can span along the random line is $O(\sqrt{\log n} \cdot |v_i - v_j|)$. Along with lemma 4 this implies that the bandwidth of the final ordering is $O(\rho \log^{2.5} n)$ times the optimum with high probability. □

5 A Lower Bound on Volume Distortion

Our bandwidth algorithm and its analysis motivate the question of whether there are embeddings with better volume distortion. In this section we show that even for a path on n vertices, the best possible volume distortion is $\Omega(\sqrt{\log n})$. Thus a further improvement in bandwidth approximation will have to come from other sources.

Theorem 3. *Let G be a path on n vertices. Then for any Euclidean embedding of G, the distortion for subsets of size up to k is* $\Omega((\log n)^{1/2-1/k})$*. For* $k = \Omega(\log\log n)$*, the distortion is* $\Omega((\log n)^{1/2})$*.*

We begin by proving that the distortion is $\Omega((\frac{\log n}{\log\log n})^{1/4})$ for subsets of size 3.

Proof of weaker bound. Let our embedding map $\{u_1, ...u_n\}$ to $\{\phi(u_1), ..., \phi(u_n)\}$ and let $P_1 = \phi(u_1), P_2 = \phi(u_2), P_3 = \phi(u_3)$. We will show a tradeoff between the area of $\{P_1, P_2, P_3\}$ and the length of P_1P_3. Applying this recursively will yield the claimed bound.

Let $|P_1P_2| = |P_2P_3| = 1$ and let the perpendicular distance from P_2 to P_1P_3 be d. Let $|P_1P_3| = 2c$. The area of the triangle is dc and the Pythagorean identity yields $1 = d^2 + c^2$. Assume that β is an upper bound on the 3-volume distortion of any subset of three points in our embedding. Then

$$\left(\frac{1}{2dc}\right)^{1/2} = \left(\frac{Tvol(S)}{(k-1)!Evol(\phi(S))}\right)^{\frac{1}{k-1}} \leq \beta$$

and since $c \leq 1$, we find $d \geq \frac{1}{2\beta^2}$.

Using the Pythagorean identity, this implies $c = \sqrt{1-d^2} \approx 1 - d^2/2 \leq 1 - \frac{1}{8\beta^4}$. Thus every distance between two points $\phi(u_i), \phi(u_{i+2})$ is at most $2\cdot(1-\frac{1}{8\beta^4})$. Now we apply the same argument to subsets of three points at distance 2 apart, $\{\phi(u_i), \phi(u_{i+2}), \phi(u_{i+4})\}$. We obtain that the distance between $\phi(u_i)$ and $\phi(u_{i+4})$ is at most $4 \cdot (1 - \frac{1}{8\beta^4})^2$. Continuing this analysis, we find that the distance $|\phi(u_1) - \phi(u_n)|$ is at most $n \cdot (1 - \frac{1}{8\beta^4})^{\log n}$.

However, our assumption that we have distortion at most β implies $|\phi(u_1) - \phi(u_n)| \geq n/\beta$. Thus we have

$$n \cdot (1 - \frac{1}{8\beta^4})^{\log n} \geq n/\beta$$

implying $\beta \geq \left(\frac{\log n}{\log\log n}\right)^{1/4}$. ◻

Proof of Stronger Bound. Consider the volume of $\{P_1, ...P_k\}$, and assume without loss of generality that $\forall i, |P_iP_{i+1}| = 1$. Now let $c_i = \frac{1}{2}|P_iP_{i+2}|$, and $d_i =$ orthogonal distance from P_{i+1} to P_iP_{i+2}. We first claim that $Evol(P_1, ...P_k) \leq \frac{\prod_{i=1}^{k-2}(2d_i)}{(k-1)!}$. The proof is by induction. Our base case is $Evol(P_1, P_2) \leq 1$, which is clear. Assume that $Evol(P_1, ...P_j) \leq \prod_{i=1}^{j-2}(2d_i)/(j-1)!$ and consider P_{j+1}. We have that the midpoint of $P_{j-1}P_{j+1}$ is d_{j-1} away from P_j. This implies that P_{j+1} is no more than $2d_{j-1}$ away from the subspace spanned by $\{P_1, ...P_j\}$. The claim follows. Our new bound on the $\{d_i\}$ follows from

$$\left(\frac{1}{\prod_{i=1}^{k-2} d_i}\right)^{\frac{1}{k-1}} \leq \left(\frac{Tvol(S)}{(k-1)!Evol(\phi(S))}\right)^{\frac{1}{k-1}} \leq \beta$$

and the bound is $\frac{\sum_{i=1}^{k-2} 2d_i^2}{k-2} \geq (\prod_{i=1}^{k-2} 2d_i^2)^{\frac{1}{k-2}} \geq \frac{1}{2}\beta^{-2(\frac{k-1}{k-2})}$ where the first inequality follows from the arithmetic mean-geometric mean inequality. As before, we have $c_i \leq 1 - d_i^2/2$. Since $|P_1 P_k| \leq 2(c_1 + c_3 + ...c_{k-3} + 1)$ and $|P_1 P_k| \leq 2(1 + c_2 + c_4 + ...c_{k-2})$, we find that $|P_1 P_k| \leq 2 + \sum_{i=1}^{k-2} c_i = (k-1)(\frac{k}{k-1} - \frac{k-2}{k-1}\frac{\sum_{i=1}^{k-2} d_i^2}{2(k-2)}) \approx (k-1)(1 - \frac{\sum_{i=1}^{k-2} d_i^2}{2(k-2)})$. Our bound on the length of $P_1 P_k$ becomes $|P_1 P_k| \leq (k-1)(\frac{k}{k-1} - \frac{k-2}{k-1}\frac{1}{8}\beta^{-2\frac{k-1}{k-2}})$. Now we apply our recursive construction again, this time on sets of size k at a time. Since we are no longer just doubling each time, we can apply our analysis only $\log_k n$ times. Plugging this in yields the bound

$$\left(\frac{k}{k-1} - \frac{k-2}{k-1}\frac{1}{8\beta^{2(\frac{k-1}{k-2})}}\right)^{\frac{\log n}{\log k}} \geq \frac{1}{\beta}$$

which simplifies to $\log n \leq 16(\log k)\beta^{2(1+\frac{1}{k-2})}\log\beta$, implying $\beta \geq (\log n)^{(1/2-1/k)}$.

□

6 Embeddings Preserving Point-To-Subset Distances

The distance of a point (or vertex) u of G to a subset of points S is simply $d(u, S) = \min_{v \in S} d(u, v)$. For points in Euclidean space, let us define the distance of a point u to a set of points S as the minimum distance from u to the *convex hull* of S, which we denote with the natural extension of $|\cdot|$. We denote the convex hull of a set of points S by $\text{conv}(S)$, and the affine hull by $\text{aff}(S)$.

Definition 2. *The* **point-to-subset distortion** *of an embedding $\phi(G)$ is*

$$\max_{u \in V, S \subset V} \frac{d(u, S)}{|\phi(u) - \text{conv}(\phi(S))|}$$

In this section we investigate the question of the best possible point-to-subset distortion of a Euclidean metric. Besides its geometric appeal, the question has the following motivation. Suppose we replaced "convex" in the definition above by "affine" and called the related distortion the *affine point-to-subset* distortion. Then we would have the following connection with volume distortion.

Lemma 5. *Let $\phi(G)$ be a contraction embedding of a metric G. Then the affine point-to-subset distortion is an upper bound on the k-volume distortion, for all $2 \leq k \leq n$.*

Proof of lemma 5. Consider a set S of vertices in G, and a mimimum spanning tree T of S. Consider any leaf u of T. If the point-to-subset distortion of our embedding ϕ is β, then

$$d(u, S \setminus \{u\}) \leq \beta|\phi(u) - \text{aff}(\phi(S \setminus \{u\}))|$$

Proceeding inductively, we find that

$$\begin{aligned} Tvol(S) &\le \beta^{k-1}(\text{volume of parallelepiped defined by } \phi(S)) \\ &\le \beta^{k-1}(k-1)!Evol(\phi(S)) \end{aligned}$$

□

We now state our main theorem on point-to-subset distortion. In the next two subsections, we define the embedding, and then prove that the embedding satisfies the theorem.

Theorem 4. *For any Euclidean metric G where the shortest distance is within a poly(n) factor of the longest distance, there exists a Euclidean embedding $\phi(G)$ whose point-to-subset distortion is $O(\sqrt{\log n})$.*

6.1 The Embedding

Let $G = (V, E)$ be a Euclidean metric with distances (edge lengths) $d(u, v)$ for all pairs of vertices $u, v \in V$. We assume that all the distances lie between 8 and $8n$. (Any polynomial upper bound on the ratio of the shortest to the longest distance would suffice). Since G is Euclidean, we can assume without loss of generality that the vertices are points in some d-dimensional Euclidean space. Given only the distances, it is trivial to find points realizing the distances by solving an SDP. The embedding we now describe was inspired by the work of Rao [6].

Before defining the embedding in general, let us consider the following illustrative example. Suppose that $d = 1$, i.e., all the points lie on a line. In this case, we could proceed by generating coordinates according to the following random process: for each R in the set $\{1, 2, 2^2, \ldots, 2^{\lfloor \log n \rfloor}\}$, we repeat the following procedure N times: choose each point from the subset $\{1, \ldots, 8n\}$ with probability $1/R$ for inclusion in a set S, and come up with a coordinate $\phi_S(v)$ for every $v \in V$. The coordinate $\phi_S(v)$ is defined to be $\min_{w \in S} |v - w|$, and then $\phi(v)$ is the vector given by the set of coordinates for v. This will yield $N \log n$ coordinates.

We now explain why this yields a $\sqrt{\log n}$ affine point-to-subset distortion. This is a stronger property than $\sqrt{\log n}$ point-to-subset distortion, and it will only be proved for $d = 1$. Consider a set $U \subset V$ and a point u, with distance on the line $d(u, U)$. For every S, we have that $|\phi_S(u) - \text{aff}(\phi_S(U))| \le d(u, U)$, and it is a simple application of Cauchy-Shwarz to get that $|\phi(u) - \text{aff}(\phi(U))| \le d(u, U)\sqrt{N \log n}$. To obtain a lower bound of $\Omega(d(u, U)\sqrt{N})$, we consider the largest R such that $R \le d(u, U)$. Denote this value of R by r; we now show that with constant probability, a set S chosen by including points in S with probability $\frac{1}{r}$ yields $|\phi_S(u) - \text{aff}(\phi_S(U))| = \Omega(d(u, U))$. We get this from the following view of the random process: fix some particular affine combination aff_0, pick points for inclusion in S at distances in $(d(u, U)/2, d(u, U))$ to the left and to the right of u, pick the rest of the points not near u with probability $\frac{1}{r}$, and we still have constant probabilty of picking another point within the two points bracketing u; the variation in $|\phi_S(u) - \text{aff}_0(\phi_S(U))|$ due to the distinct choices for this last point included in S is $\Omega(d(u,U))$ with constant probability. Feige proves [4] that the number of "distinct" affine combinations is not too great, and thus taking N sufficiently large, but still polynomial, yields that this occurs with high probability simultaneously for all "distinct" affine combinations, and thus over the uncountable set of all affine combinations. Taking N a little bit larger still (but still polynomial) then yields that this is simultaneously true for all point-subset pairs.

For the general case ($d \geq 1$), our algorithm chooses a random line, projects all the points to this random line, and then computes the coordinates as above. In detail, for each R in the set $\{1, 2, 2^2, \ldots, 2^{\lfloor \log n \rfloor}\}$, we repeat the following procedure N times:

1. Pick a random line ℓ through the origin.
2. Project all the points to ℓ, and scale up by a factor of $\sqrt{d}$. Let the projection of u be u^ℓ.
3. Place points along ℓ at unit intervals. Pick a random subset S of these points, by choosing each point with probability $\frac{1}{R}$, independently.
4. The coordinate for each vertex u along the axis corresponding to the S and ℓ pair is $\phi_S(u) = d(u^\ell, S) = \min_{w \in S} |u^\ell - w|$.

Thus the total number of dimensions is $O(N \log n)$. For the same reasons as cited above, $N = \mathit{poly}(n)$ will suffice. Thus the dimension of the final embedding is polynomial in n.

6.2 The Proof

We first upper bound the point-to-subset distances in our final embedding. Consider any point u and subset U. It is enough to consider the point $v \in U$ minimizing $d(u,v)$ by the following lemma.

Lemma 6. *For every pair $u, v \in V$,*

$$|\phi(u) - \phi(v)| \leq 2d(u,v)\sqrt{N \log n}$$

Proof. After scaling up by a factor of $\sqrt{d}$, we have that for any pair $u, v \in V$,

$$\begin{aligned}
|\phi(u) - \phi(v)|^2 &= \sum_{(S,\ell)} |\phi_S(u) - \phi_S(v)|^2 \\
&= \sum_{(S,\ell)} |d(u^\ell, S) - d(v^\ell, S)|^2 \\
&\leq \sum_{(S,\ell)} d(u^\ell, v^\ell)^2 \\
&\leq \sum_{(S,\ell)} 2d(u,v)^2 \qquad (6) \\
&= d(u,v)^2 N \log n
\end{aligned}$$

where step 6 is true with constant probability for a single random line, and with very high probability when summing over all the random lines. □

Since $|\phi(u) - \text{conv}(\phi(U))| \leq |\phi(u) - \phi(v)|, v \in U$, we have our upper bound.

Now we lower bound the point-to-subset distances. Consider again a particular point u and subset $U = \{u_i\}$, and some fixed convex combination $\{\lambda_i\}$ such that $\sum \lambda_i = 1$ and $\forall i, \lambda_i \geq 0$. Let r be the highest power of 2 less than $d(u, U)$. For any coordinate corresponding to a subset S generated using $R = r$, we show that $|\phi(u) - \sum_i \lambda_i \phi(u_i)| = \Omega(d(u, U))$ with constant probability.

Towards this goal, we claim there is a constant probability that the following two events both happen.

(i) $\sum_i \lambda_i \phi_S(u_i) \geq r/16$
(ii) $\phi_S(u) \leq r/32$

First we condition on some point within $r/32$ of u^ℓ being chosen for inclusion in S. This happens with constant probability (over choice of S). We have at least a constant probability of the λ_i's corresponding to u_i's at least $r/4$ away from u adding up to at least 2/3 (over choice of ℓ). Condition on this as well. Then we lower bound the expected value of $\sum_i \lambda_i \phi_S(u_i)$ by

$$E[\sum_i \lambda_i \phi_S(u_i)] \geq \sum_{i:|u_i^\ell - u_i| \geq r/4} \lambda_i E[\phi_S(u_i)]$$

Since $E[\phi_S(u_i)] \geq r/8$, we have that the expectation is at least $(2/3)(r/8) = r/12$. By Markov's inequality, the value of $\sum_i \lambda_i \phi_S(u_i)$ is at least $r/16$ with constant probability.

Since this happens for all coordinates with $R = r$, i.e. N of the coordinates, we obtain the lower bound. As before, a polynomially large N suffices to make the statement true with high probability for every point, every subset, and every convex combination simultaneously.

7 Convexity of k^{th} Moments

In section 6.1, we proved that our embedding did preserve all *affine* point-to-subset distances to within $O(\sqrt{\log n})$ for the case $d = 1$. Since the optimal solution to the bandwidth problem is an arrangement of points on a line, it is the case that an embedding realizing this distortion of the optimal solution exists. This implies that the constraint

$$\sum_{|S|=k} \frac{1}{Evol(S)} \leq \sum_{|S|=k} \frac{(k-1)!\rho^k}{Tvol(S)} \leq (2\rho k \log n)^k$$

with $\rho = \sqrt{\log n}$ is satisfied by a Euclidean embedding of the optimal solution. We call this constraint the k^{th} *moment constraint*. We show in this section that the above constraint is convex, and thus we can impose it explicitly in our SDP (replacing the spreading constraint), separate over it, and then apply the machinery of section 4 with $\rho = \sqrt{\log n}$ to obtain an $O(\log^3 n)$ approximation to the optimal bandwidth. The only caveat is that the constraint has $\binom{n}{k} = O(n^{\log n})$ terms, so this is not quite a polynomial time algorithm. We proceed with

Lemma 7. *Let c be fixed. The following is a convex constraint over the set of Postive Semi-Definite (PSD) matrices X.*

$$\sum_{|S|=k} \frac{1}{Evol(X_S)} \leq c$$

Proof. We analyse the constraint given in the lemma on a term by term basis. Suppose X and Y are PSD matrices, and $(X+Y)/2$ is their convex combination. Then it suffices to show that

$$\frac{1}{Evol((X+Y)/2)} \leq \frac{1}{2}\left(\frac{1}{Evol(X)} + \frac{1}{Evol(Y)}\right)$$

because the constraint in the lemma statement is just a sum over many submatrices. We actually prove the stronger statement that

$$\frac{1}{Evol((X+Y)/2)} \leq \sqrt{\frac{1}{Evol(X)}\frac{1}{Evol(Y)}}$$

which implies the former statement by the arithmetic mean-geometric mean inequality (GM $\leq$ AM). This last statement is equivalent to (clearing denominators and squaring twice)

$$Det(XY) \leq Det^2((X+Y)/2)$$

which is equivalent to

$$\begin{aligned}
1 &\leq \frac{Det^2((X+Y)/2)}{Det(XY)} \\
&= Det(\frac{1}{4}(X+Y))Det(X^{-1})Det(X+Y)Det(Y^{-1}) \\
&= Det(\frac{1}{4}(X+Y)(X^{-1})(X+Y)(Y^{-1})) \\
&= Det(\frac{1}{4}(I+YX^{-1})(XY^{-1}+I)) \\
&= Det(\frac{1}{4}(YX^{-1}+2I+XY^{-1})) \\
&= Det(\frac{1}{4}(A+2I+A^{-1}))
\end{aligned}$$

where we let $A = YX^{-1}$ at the very end. Also let $B = \frac{A+2I+A^{-1}}{4}$. We have reduced our original claim to showing that $Det(B) \geq 1$. We will show the stronger property that every eigenvalue of B is at least 1. Consider an arbitrary (eigenvector, eigenvalue)-pair of A , given by (e, λ). Then

$$Be = \frac{1}{4}(\lambda + 2 + \frac{1}{\lambda})e$$

Since $\frac{1}{4}(\lambda + 2 + \frac{1}{\lambda}) \geq 1$, we have that e is an eigenvector of eigenvalue at least 1 for B (this used that $\lambda \geq 0$, which is true since A is PSD). Since the eigenvectors of A form an orthonormal basis of the whole space, all of B's eigenvectors are also eigenvectors of A. □

8 Conclusion

We conjecture that the embedding described in section 6.1 has $O(\sqrt{\log n})$ *affine* point-to-subset distortion as well. This would directly imply that our algorithm achieves an $O(\log^3 n)$ approximation for the minimum bandwidth in polynomial time.

References

1. A. Blum, G. Konjevod, R. Ravi, and S. Vempala, "Semi-Definite Relaxations for Minimum Bandwidth and other Vertex-Ordering Problems," Proc. 30th ACM Symposium on the Theory of Computing, 1998.
2. J. Bourgain, "On Lipshitz embedding of finite metric spaces in Hilbert space," Israel J. Math. 52 (1985) 46-52.
3. U. Feige, "Approximating the bandwidth via volume respecting embeddings," in Proc. 30th ACM Symposium on the Theory of Computing, 1998.
4. U. Feige, "Improved analysis of the volume distortion of the random subsets embedding," Manuscript.
5. C. H. Papadimitriou, The NP-completeness of the bandwidth minimization problem, *Computing*, 16: 263-270, 1976.
6. S. Rao, "Small distortion and volume preserving embeddings for planar and Euclidean metrics," Proc. of Symposium on Computational Geometry, 1999.
7. S. Vempala, "Random Projection: A new approach to VLSI layout," Proc. of FOCS 1998.