Improved Algorithms for Optimal Embeddings

(This paper supersedes previous versions that appear in ACM Transactions of Algorithms, August 2008 and ECCC, August 2006. The current version corrects attributions to concurrent works on the same problem.)

Nishanth Chandran[1 ‡] Ryan Moriarty [1 ‡] Rafail Ostrovsky [§]
Omkant Pandey [1 ‡] MohammadAli Safari [2] Amit Sahai [1 ¶]

Abstract

In the last decade, the notion of metric embeddings with small distortion received wide attention in the literature, with applications in combinatorial optimization, discrete mathematics and bio-informatics. The notion of embedding is, given two metric spaces on the same number of points, to find a bijection that minimizes maximum Lipschitz and bi-Lipschitz constants. One reason for the popularity of the notion is that algorithms designed for one metric space can be applied to a different metric space, given an embedding with small distortion. The better the distortion, the better is the effectiveness of the original algorithm applied to a new metric space.

The goal that was recently studied by Kenyon, Rabani, and Sinclair [KRS04] is to consider all possible embeddings between two *finite* metric spaces and to find the best possible one, i.e., consider a single objective function over the space of all possible embeddings that minimizes the distortion. In this paper we continue this important direction. In this paper, we are able to provide an algorithm to find the optimal bijection between two line metrics provided that the optimal distortion is smaller than 13.602. Further, we show an inherent limitation of algorithms using the "forbidden pattern" based dynamic-programming approach that they cannot find optimal mapping if the optimal distortion is more than $7 + 4\sqrt{3} (\simeq 13.928)$. Thus, our results are almost *optimal* for this method. Finally, we also show that previous techniques for *general* embeddings apply to a (slightly) more general class of metrics.

[1] Department of Computer Science, University of California, Los Angeles, 90095.
Contact: `{nishanth,ryan,omkant,sahai}@cs.ucla.edu`
[2] Department of Computer Science, University of Alberta.
Contact: `msafarig@cs.ualberta.ca`

[‡] Research supported in part by grants listed below.

[§] Department of Computer Science and Mathematics, University of California, Los Angeles, 90095.
Contact: `rafail@cs.ucla.edu`; Supported in part by IBM Faculty Award, Xerox Innovation Group Award, NSF grants no. 0430254, 0716835, 0716389 and U.C. MICRO grant.

[¶] This research was supported in part by NSF ITR and Cybertrust programs (including grants 0627781, 0456717, 0716389, and 0205594), a subgrant from SRI as part of the Army Cyber-TA program, an equipment grant from Intel, an Okawa Research Award, and an Alfred P. Sloan Foundation Research Fellowship.

1 Introduction

For a bijection $\sigma : U \to V$ between two n-point metric spaces (U, d) and (V, d'), the expansion of σ is defined as

$$\text{expansion}(\sigma) = \max_{x,y \in U, x \neq y} \frac{d'(\sigma(x), \sigma(y))}{d(x, y)}$$

The distortion σ is defined as follows: $\text{dist}(\sigma) = \text{expansion}(\sigma) \times \text{expansion}(\sigma^{-1})$. The MINIMUM DISTORTION problem is to find a bijection σ between two equal-sized finite metric spaces (U, d) and (V, d') such that $\text{dist}(\sigma)$ is minimum over all possible bijections.

The MINIMUM DISTORTION problem is interesting to study for both theoretical as well as practical reasons. From complexity theoretic point of view, it has interesting connections to GRAPH ISOMORPHISM [For96]. In particular, graph isomorphism on two input graphs G and H is trivially reduced to deciding if there exists an isometric (i.e., distortion 1) bijection between $\mathcal{M}_G$ and $\mathcal{M}_H$, where $\mathcal{M}_X$ denotes the shortest path metrics of a graph X.

On the practical side, we note that applications dealing with shape matching and object recognition (e.g., signature matching, character recognition, matching facial features, pattern matching in complicated protein structures, and so on) require good measures of similarity. Distortion is an attractive measure of similarity between two point sets [AKOF03, HKW98, BMP02, CLM03]. From the point of view of aforementioned applications, good algorithms for finding minimum distortion bijection (or optimal bijection) are highly desirable.

Kenyon, Rabani, and Sinclair [KRS04] show that the MINIMUM DISTORTION problem is NP-hard even to approximate (within a factor of 2), and provide two positive results:

- A polynomial time algorithm for *exactly* finding the minimum distortion bijection between two line metrics if the optimal bijection has distortion strictly less than $3 + 2\sqrt{2}$.
- A parameterized polynomial time algorithm for *exactly* finding optimal bijection between bounded-degree tree metric and an arbitrary unweighted graph metric.

In this paper, we improve and generalize the results of Kenyon, Rabani, and Sinclair. We note that these results were independently and concurrently discovered by Kenyon, Rabani, and Sinclair [KRS09], as well.

- In particular, we first provide a polynomial time algorithm for *exactly* finding an optimal bijection between two line metrics if the optimal bijection has distortion strictly less than 13.602.

  To achieve this improvement, we take a more general approach. In particular, [KRS04] look at a single pattern (partial bijection of size 4) and its inverse. They call this pattern a forbidden pattern. The presence of such patterns guarantees high distortion $(3+2\sqrt{2})$. We generalize this approach and look for patterns of higher sizes whose presence will guarantee even higher distortion. We call these patterns, *non-separable permutations*. Absence of such permutations guarantees that the dynamic programming approach can be applied to find the optimal bijection/permutation.

- Next, based on the idea of families of non-separable permutations, we are able to design a dynamic programming algorithm which finds a minimum distortion bijection on more instances than [KRS04]. Thus our work answers a direct open question posed in [KRS04].
- We also show a limitation of the "forbidden pattern" approach, by showing that there exists arbitrarily large families of forbidden patterns with bounded distortion. This lower bound shows the extent to which this approach will be useful and indicates a new approach must be taken to pass this bound.

1.1 Related Work

The problem of embedding distance metrics into geometric spaces has been studied extensively [Kru64a, Kru64b, She62a, She62b, JL03, Lin02]. The MINIMUM DISTORTION problem is a natural variant of bi-Lipschitz embeddings questions that were initially motivated by the study of Banach spaces.

A problem closely related to MINIMUM DISTORTION is MINIMUM BANDWIDTH. MINIMUM DISTORTION can be viewed as a variation and generalization of the MINIMUM BANDWIDTH problem [CCDG82, DPS02]. Good solutions for the MINIMUM BANDWIDTH problem, however, typically incur very large contraction and hence do not seem useful for solving MINIMUM DISTORTION.

After its introduction, MINIMUM DISTORTION problem has received considerable attention in the research community. Most of the results, however, have been negative showing that the problem is hard even to approximate. Among such results are the results of Hall and Papadimitriou who show that the line embedding problem is hard to approximate even within large factors if distortion is high [HP05], and Papadimitriou and Safra [PS05] who show that the general embedding problem is hard to approximate within a factor of 3 in three-dimension.

Due to such results, some of the research work focusses on *approximating* MINIMUM DISTORTION (e.g., see the work in [BCIS05, BDG$^+$05, HP05]) under certain circumstances (e.g., consider only injections, focus on alternate definitions of distortion such as additive distortion, and so on).

We remark that after the work of [KRS04], most of the research focussed on either approximating the distortion [BDG$^+$05, BCIS05] or proving the hardness of approximating it [HP05, PS05]. Hall and Papadimitriou in [HP05], show that line embeddings are hard to approximate even within large factors when the distortion is high.

Our results were independently and concurrently discovered by Kenyon, Rabani, and Sinclair [KRS09]. Our paper combines work done by the authors in [ES, CMO$^+$, Saf07]. In an earlier version of this work, we proved a lemma on non-separable permutations that was at the heart of our analysis and construction and was sufficient to achieve our results. Subsequently, Claire Kenyon [Ken] graciously pointed to us a result of Albert and Atkinson [AA05] which subsumed (and predated) our lemma, which we therefore use instead of our original lemma. This lemma was also used by [KRS09] to obtain results similar to ours.

We also consider (in [CMO$^+$, CMO$^+$08], as well as here,) the case of embedding a bounded degree unweighted tree metric into an arbitrary unweighted graph metric. In all these versions, we prove that the algorithm of [KRS04] actually works for a larger class of graphs - unweighted bounded degree graphs with maximum cycle length three. That is, we show that their algorithm finds optimal bijection between a bounded degree graph with maximum cycle length three and an arbitrary unweighted graph metric. We recently received a preprint of [KRS09]; we note that our work [CMO$^+$, CMO$^+$08] partially addresses an open question posed by Kenyon et al. [KRS09] to consider more than unweighted tree metrics.

2 Line Embeddings

In this section, we focus on computing an optimal embedding between two fixed line metrics. A line metric is a set of points on a real one-dimensional line with the distance between any pair of points being their ℓ_1 distance (any ℓ_k distance is equivalent for one-dimensional points).

As we mentioned earlier, Kenyon et al. [KRS04] consider the problem of optimally embedding one fixed line metric into another fixed one. They propose a polynomial-time, dynamic programming based, algorithm that computes the optimal embedding if the distortion is less than $3+2\sqrt{2}$. Kenyon et al. show that any bijection that contains the bijection in Fig. 1 as a partial bijection corresponds to an embedding with distortion at least $3+2\sqrt{2}$. These bijections have a nice structure that allows finding the optimal such permutation using dynamic programming in polynomial time.

Notice that we can view any embedding as a mapping from source points to destination points or, simply, as a permutation. In this section, we improve the result of Kenyon et al. [KRS04] by considering a less restricted class of permutations called k-separable permutations. We improve the threshold value on distortion below which an optimal embedding can be found in polynomial time from $3 + 2\sqrt{2}$ to 13.602. Let us now introduce some basic definitions.

2.1 Basic Definitions

Assume the optimal embedding between U and V is the permutation π. We specify a permutation π with the notation $(\pi(1), \pi(2), \cdots, \pi(n))$.

Permutation π_n of size n *contains* permutation π_k of size k, if there exist indices $l_1 < l_2 < \cdots < l_k$ such that for all $1 \le i < j \le k$, $\pi_k(i) < \pi_k(j)$ iff $\pi_n(l_i) < \pi_n(l_j)$. In this case, we refer to π_k as a *sub-permutation* of π_n. In particular, $\pi_n^{x,y}$ is the unique permutation of size $y-x+1$ such that $\pi_n^{x,y}(i) < \pi_n^{x,y}(j)$ iff $\pi_n(i+1-x) < \pi_n(j+1-x)$.

By $[i, j]$, $i < j$, we mean the set of numbers from i to j. A *nice interval* I in π is either a singleton or is a set of at least two consecutive numbers from 1 to n such that their mapping, via π, is still a set of consecutive numbers. For example, the permutation $(4, 3, 1, 2)$ contains several nice intervals: $[1, 2]$, $[3, 4]$, $[2, 4]$ and $[1, 4]$.

If the interval $[1, n]$ can be decomposed into a constant number of sub-intervals such that each sub-interval is mapped, via π, to a sub-interval in V and this property recursively holds for all sub-intervals, then we can use dynamic programming and find the optimal embedding. More formally, an interval I is *k-separable*, with respect to π, if either it has at most k points or it can be partitioned into nice sub-intervals $I_1, I_2, \cdots, I_m$ $(1 < m \le k)$ such that each I_i is k-separable. π is k-separable iff the interval $[1, n]$ is k-separable with respect to π. The *separability* of π is the minimum $k > 1$ such that π is k-separable.

For example, the permutation $\pi = (2, 4, 3, 6, 5, 1)$ is 3-separable. $I_1 = [1, 3]$, $I_2 = [4, 5]$, $I_3 = [6]$, and it is clear that I_1, I_2, and I_3 are 3-separable as well.

Every 3-separable permutation is 2-separable, since for any three nice sub-intervals that partition a permutation, two may be merged to form a nice sub-interval. Therefore, we don't have any permutation with separability 3. It's also easy to see that for $k \ge 4$, there exist permutations of size k with separability k. These permutations could be interpreted in a simpler way: they don't have any nice interval except the interval $[1, k]$. We refer to these special k-separable permutations as *non-separable* permutations.

The distortion incurred by a permutation π, denoted by $\mathrm{dist}(\pi)$, is the minimum distortion incurred by embedding any two line metrics U and V via π. For example, $\mathrm{dist}(\pi)$ for the permutation in Fig. 1 equals $3 + 2\sqrt{2}$ and happens when $[a, b, c, x, y, z] = [1, \sqrt{2}, 1, 1, \sqrt{2}, 1]$. As we see later, Theorem 2.3 states that $\mathrm{dist}(\pi)$ equals the largest eigenvalue of a 0-1 matrix corresponding to π.

Corresponding to every permutation π of size n, there exist three permutations π^0, π^1, and π^{-1} that are similar to π and incur the same distortion. For all i's, $\pi^0(i) = n + 1 - \pi(i)$, $\pi^1(\pi(i)) = i$, and $\pi^{-1}(i) = n + 1 - \pi^1(i)$. For example, if $\pi = (2, 4, 1, 3)$, $\pi^0 = \pi^1 = (3, 1, 4, 2)$ and $\pi^{-1} = \pi$. Throughout this section, we always assume that a permutation comes with all its four symmetric forms. For example, when we say 2-separable permutations avoid $\pi = (2, 4, 1, 3)$ we mean they avoid $(3, 1, 4, 2)$ as well.

Let Π_k be the set of all non-separable permutations of size k. Let d_k be the minimum distortion over all permutations in Π_k. For example, $\Pi_4 = \{(2, 4, 1, 3)\}$, $\Pi_5 = \{(2, 4, 1, 5, 3), (2, 5, 3, 1, 4), (3, 5, 1, 4, 2)\}$, and it's not hard to see that $d_4 = d_5 = 3 + 2\sqrt{2}$. Note that by $\pi \in \Pi_k$ we implicitly mean $\pi^0, \pi^1, \pi^{-1} \in \Pi_k$ as well. So, $(3, 1, 5, 2, 4)$ is also in Π_5.

2.2 Forbidden Permutations

One commonly asked question regarding many permutation classes is whether they can be characterized by a finite forbidden set of permutations or not. For example, a permutation is 2-separable if and only if

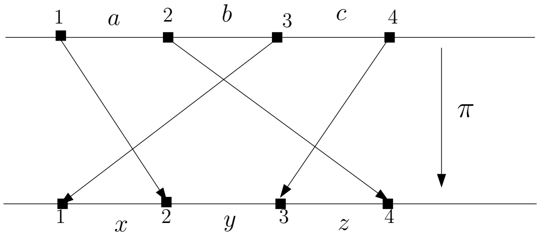


Figure 1: The 4-separable permutation $(2, 4, 1, 3)$.

it contains neither $(2, 4, 1, 3)$ nor $(3, 1, 4, 2)$[BBL98].

Interestingly one can generalize this statement for k-separable permutations.

Theorem 2.1. *A permutation is k-separable if and only if*

- *For odd k, it doesn't contain any permutation in Π_{k+1}.*
- *For even k, it contains neither a permutation in Π_{k+1} nor π^*_{k+2}.*

where π^*_{2m} is the permutation of size $2m$ in which $\pi^*_{2m}(2i) = i$ and $\pi^*_{2m}(2i-1) = i+m$.

Proof. Assume π is not k-separable. Then, it must contain a non-separable permutation π^0 of size $k_0 > k$. According to Albert and Atkinson [AA05], every non-separable permutation of size m either contains a non-separable permutation of size $m-1$ or is identical to π^*_m. As π^*_m contains π^*_{m-2} the statement of the theorem follows by repeatedly using the theorem of Albert and Atkinson [AA05]. □

Note: Albert and Atkinson [AA05] (See Theorem 4 in their paper) use the notion *simple* for non-separable and call π^*_{2m} an exceptional permutation. They obtain their result by using results from Schmerl and Trotter [ST93] on partially ordered sets. We thank Claire Kenyon [Ken] for pointing us to [AA05], which was also used by [KRS09] to obtain results similar to ours.

2.3 Embedding between two line metrics

In this section we prove the following theorem which is a generalization of Kenyon et al.'s result.

Theorem 2.2. *For any two line metrics U and V and any k either the distortion of the optimal embedding between U and V is greater than d_{k+1} or there exists an $O(k!n^{5k+2})$ time algorithm (which is a polynomial in n when k is a constant) for computing the optimal embedding.*

Recall that d_{k+1} is the minimum distortion over all permutations in Π_{k+1}. Let π be the optimal embedding permutation. If π is not k-separable then, according to Theorem 2.1, π contains either a permutation in Π_{k+1} or π^*_{k+2} (in the case that k is even). From Sections 2.5 and 2.6, we can compute $\text{dist}(\pi^*_{k+2}) = 2k + 2\sqrt{k(k-1)} - 1$ and also conclude that $\text{dist}(\pi^*_{k+2}) \geq d_{k+1}$ for all k. Hence if π is not k-separable, then $\text{dist}(\pi) \geq d_{k+1}$. Otherwise, an algorithm for finding the optimal embedding follows.

2.4 The Algorithm

2.4.1 Algorithm Intuition

Our algorithm will guarantee that we solve all inputs whose optimal bijection π is k-separable, where k is the parameter. Note that if $\text{dist}(\pi) < d_{k+1}$, then π is k-separable. Setting $k = 46$, we get an algorithm that computes the bijection when the optimal bijection π has distortion $\text{dist}(\pi) < 13.602$.

At an intuitive level our algorithm will work as follows. It looks at every possible subinterval of the points in U against every possible subinterval of the points in V starting with size 2 and working up to size n. It will break the subintervals into every possible k subsubintervals (including the empty sets). It will then try match these k subsubintervals by trying all $k!$ possible bijections of the subsubintervals. If a match is found with low enough distortion the match will be saved for future reference. How the subintervals are mapped is no longer important; the only things we need to know about the subinterval to continue the process is whether there was a bijection with distortion less than d_{k+1}, and the image and the preimages respectively of the first and last point of U and the first and last point in V. The reason we need to keep the mappings of the first and last points in U and V is because when we try to combine two subintervals we need to check the expansion and inverse expansion between them. We store this information in a table. When the subinterval is U and V, if we can map U to V by the same process with distortion less than d_{k+1} we output "yes".

Another way to think about the algorithm is that the algorithm is looking for mappings that contain a pattern size k_1 for some $k_1 \leq k$. If it finds such a pattern it now thinks of that entire set as one mapping that could be part of another pattern of size $\leq k$ and looks for such a pattern.

2.4.2 Algorithm

The algorithm gets as input, two line metrics (U, d) and (V, d'). It also gets as parameters, $\alpha = \sqrt{d_{k+1}}$ the maximum expansion and inverse expansion allowed, as well as a parameter k which is related to the bijections that the algorithm tries.

The algorithm proceeds by building a dynamic programming boolean table T which is indexed by the following parameters:

- a subinterval $I = \{u_m, u_{m+1}, \ldots, u_{m+c-1}\}$ of U and a subinterval $J = \{v_{m'}, v_{m'+1}, \ldots, v_{m'+c-1}\}$ of V of the same size $c \geq 1$;
- four elements $v, v' \in J$ and $u, u' \in I$.[1] v is the image of the first point in I. v' is the image of the last point in I. Similarly u is the preimage of the first point in J. And u' is the preimage of the last point in J.

We set the table entry $T[I, J, v, v', u, u']$ to true if there is a bijection $\sigma : I \to J$ such that $\sigma(u_m) = v, \sigma(u_{m+c-1}) = v', \sigma^{-1}(v_{m'}) = u$ and $\sigma^{-1}(v_{m'+c-1}) = u'$, and with expansion and inverse expansion at most α.

The algorithm runs from $c = 1$ to n. The base case $c = 1$ is trivial, with all entries set to true. For $c > 1$, compute every entry $T[I, J, v, v', u, u']$ with $|I| = c$ and $|J| = c$ as follows: consider all partitions of I and J into $2 \leq k_0 \leq k$ subintervals $I = \bigcup_{a=1}^{k_0} I_a$, $J = \bigcup_{b=1}^{k_0} J_b$. Try all possible combinations of pairs of I_a, J_b ($\sigma(I_a) = J_b$) over all a,b and set $T[I, J, v, v', u, u']$ to true if and only if in at least one of the combinations, the following conditions hold:

- $\forall a, b \;\; T[I_a, J_b, v_b, v'_b, u_a, u'_a]$ is true, where $\sigma(I_a) = J_b$.
- Let $J_{b_1} = \sigma(I_a), J_{b_2} = \sigma(I_{a+1}), I_{a_1} = \sigma^{-1}(J_b), I_{a_2} = \sigma^{-1}(J_{b+1})$. Then,

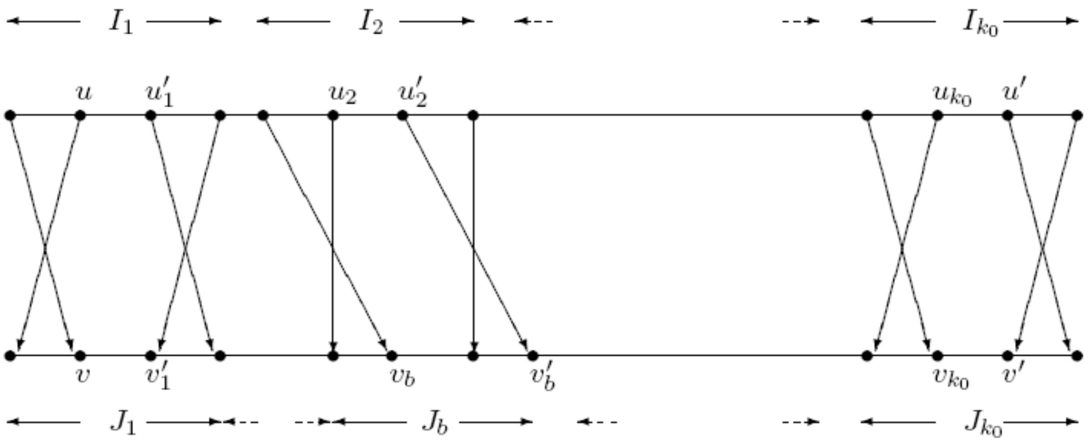
$$\begin{aligned} d(v_{b_2}, v'_{b_1}) &\leq \alpha \cdot d(\min(I_{a+1}), \max(I_a)) \\ d(u_{a_2}, u'_{a_1}) &\leq \alpha \cdot d(\min(J_{b+1}), \max(J_b)) \end{aligned}$$

  These inequalities ensure that the edges connecting the subintervals have expansion and inversion expansion at most α.[2]

---

[1] Here, v' and u' do not denote images. They are just normal points. The same will hold throughout this subsection and we will specifically mention the images.

[2] Note that we only need to consider the expansion and inverse expansion of edges [KRS04].



Filling in the table: a possible case

Once the table is prepared, the algorithm just checks that $T[U, V, v, v', u, u']$ is true for some (v, v', u, u').

2.4.3 Analysis

Correctness For the correctness of this algorithm we must show that we can solve any bijection whose optimal mapping is has distortion less than d_{k+1}. Since, the distortion of the optimal mapping is less than d_{k+1}, the optimal mapping is k-separable. Hence, the permutation π contains only nice intervals of sizes at most k. Thus the algorithm will try each of these partial mappings (on the nice intervals) and return a value of true for them.

Running Time The running time of the algorithm is easy to bound. Notice that the table size is just $O(n^7)$. Computing each entry $T[I, J, v, v', u, u']$ of the table is polynomial in n: the sets I and J can be split into $k_0 \leq k$ sets in $O(n^{k-1})$ ways and for each such possible splitting we store $4(k_0 - 2) + 2 + 2 \leq 4(k - 1)$ mappings, which can be done in $O(n^{4k-4})$; and finally there are $k!$ possible ways of mapping various I_a to various J_b. Thus computing each entry takes $O(n^{4k-4} \cdot n^{k-1} \cdot k!) = O(k!n^{5k-5})$ time. So, computing the whole table takes $O(k!n^{5k-5} \cdot n^7) = O(k!n^{5k+2})$.

This also completes the proof of Theorem 2.2.

2.5 Largest Eigenvalue

In this subsection, we provide an interesting observation that the distortion of non-separable patterns can be computed by computing the *largest eigenvalue* of the 0-1 matrix of their permutation. This observation suggests that we can find minimum distortions using a computer program.

Assume the distortion corresponding to a permutation π of $[1, n]$ is λ. That means that for any two line metrics of n points each, the distortion using π is at least λ and there exists a pair of line metrics whose distortion, using π, is exactly λ. In fact it is not hard to see that the maximum expansion and inverse expansion in embedding U to V happens for a pair of consecutive points, so we need to care only about them. Finding $\text{dist}(\pi)$ corresponds to solving a set of linear equations. For example, for the permutation in Fig. 1, the linear equations are as follows.

k	5	7	9	11	13	15	17	19
distortion	8.352	10.896	12.045	12.651	13.007	13.233	13.385	13.492

Table 1: Distortion of $\hat{\pi}_k$ for several values of k.

$$\begin{aligned} y+z &\leq \sqrt{\lambda}a \\ x+y+z &\leq \sqrt{\lambda}b \\ x+y &\leq \sqrt{\lambda}c \\ a+b &\leq \sqrt{\lambda}x \\ a+b+c &\leq \sqrt{\lambda}y \\ b+c &\leq \sqrt{\lambda}z \end{aligned}$$

or equivalently $AX \leq \sqrt{\lambda}X$, where A is the adjacency matrix corresponding to π and X is $[a,b,c,x,y,z]^T$. In general, for a permutation π of size n that corresponds to embedding between two line metrics of size n, A has $2n-2$ rows and columns where, for all $0 \leq i,j < n$, $A[i,j] = A[n+i,n+j] = 0$ and $A[i,n+j] = A[n+i,j]$ is one iff the interval $[\pi(i),\pi(i+1)]$ (or $[\pi(i+1),\pi(i)]$ if $\pi(i) > \pi(i+1)$) contains the interval $[j,j+1]$ and is zero otherwise.

We can also assume that, by scaling edge weights in U or V if necessary, the expansion and contraction both equal $\sqrt{\lambda}$. Thus, for any single edge in U and V we write an inequality to make sure that its corresponding expansion does not exceed $\sqrt{\lambda}$.

Since we are interested in minimizing λ we better make the equality $AX = \sqrt{\lambda}X$.[3] Therefore, $\sqrt{\lambda}$ is an eigenvalue of A. It is well known that when all entries of A are positive, the only eigenvalue whose corresponding eigenvector is positive is the largest eigenvalue ([HJ86], Chapter 8.2.). Thus, $\sqrt{\lambda}$ is the largest eigenvalue of A.

Theorem 2.3. *Let A_π be the 0-1 matrix corresponding to π and let its largest eigenvalue be λ. Then, the distortion of π is exactly λ^2 and is obtained when the edge lengths are taken according to the eigenvector corresponding to λ.*

2.6 Bounding d_k

Although d_k is increasing in k, it remains bounded. This is somewhat disappointing since if it was unbounded we could imagine an algorithm that finds an optimal embedding for any two line metrics, no matter how large the optimal distortion is, whose running time is a function of the distortion.

Theorem 2.4. *For any value k there exists a non-separable permutation π_k whose distortion is at most $7+4\sqrt{3}$.*

Proof. Let $\hat{\pi}_{2n}$ be the permutation on $[1,2n]$ where $\hat{\pi}_{2n}(1) = 2$, $\hat{\pi}_{2n}(2n) = 2n-1$, $\hat{\pi}_{2n}(2i) = 2i+2$, and $\hat{\pi}_{2n}(2i+1) = 2i-1$, for $i = 1,2,\cdots,n-1$. Similarly, $\hat{\pi}_{2n+1}$ is defined as follows. $\hat{\pi}_{2n+1}(i) = \hat{\pi}_{2n}(i)$, for $i = 1,2,\cdots,n-1$, $\hat{\pi}_{2n+1}(2n) = 2n+1$, and $\hat{\pi}_{2n+1}(2n+1) = 2n-1$(See Fig. 2).

Set $d_U(2i-1,2i) = 1$, $d_U(2i,2i+1) = \sqrt{3}$, $d_V(2i-1,2i) = 2+\sqrt{3}$, and $d_V(2i,2i+1) = 3+2\sqrt{3}$. The distortion corresponding to this pair of point sets is $7+4\sqrt{3}$ which means $d_k \leq 7+4\sqrt{3} \simeq 13.928$. □

---

[3] To see this, let λ be the smallest distortion and assume $AX \leq \sqrt{\lambda}X$ for some positive vector X. Let X be the one with smallest sum of elements, i.e. $X_1 + X_2 + ... + X_n$. If $AX < \sqrt{\lambda}X$ then $A_iX < \sqrt{\lambda}X_i$ (for some i) which means we can replace X_i by A_iX without violating the condition $AX \leq \sqrt{\lambda}X$. This is a contradiction because now the new sum is the one with smallest sum of elements.

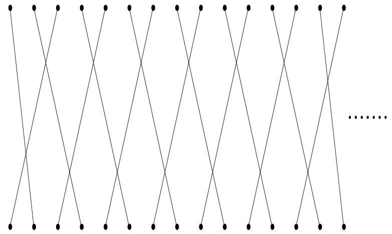


Figure 2: Illustration of permutation $\hat{\pi}_{15}$.

k	4	6	9	12	15	24
d_k	5.828	8.352	9.899	10.896	11.571	12.850
k	30	34	38	42	46	
d_k	13.131	13.316	13.443	13.534	13.602	

Table 2: d_k.

Table 1 shows the exact distortion of such permutations for small values of k. Finding d_k for small k's (by computing the eigenvalue corresponding to all permutations in Π_k and taking the minimum) suggests that d_k converges to $7 + 4\sqrt{3}$. Table 2 shows the value of d_k for different k's.

Limitation of the approach. It is easy to see that if the pattern $\hat{\pi}_{15}$ keeps extending to infinity, then its distortion is $7+4\sqrt{3}$. Using the tightness property of edges in this pattern, we get the following equations,

$$\begin{array}{l} \alpha a = 2x + 3y \;\; ; \;\; \alpha b = x + 2y \\ \alpha x = 2a + 3b \;\; ; \;\; \alpha y = a + 2b \end{array} \Rightarrow \begin{array}{c} \alpha(2b-a) = y \;\; ; \;\; \alpha(2a-3b) = x \\ \alpha^2(2a-3b) = 2a+3b \;\; ; \;\; \alpha^2(2b-a) = a+2b \end{array}$$

From which we get $\alpha^2 = 7 + 4\sqrt{3} \approx 13.928$

3 Bounded Degree Graphs with Short Cycles

Theorem 3.1. *Let (U, d) be the shortest-path metric of an unweighted graph G of maximum degree b ($b \neq 1$) . Let (V, d') be the shortest-path metric of an arbitrary unweighted graph G'. Then, the problem of finding an optimal bijection between U and V is NP-Hard.*

Proof. This proof is based on the proof that it is NP-hard to approximate the minimum distortion problem within a factor better than 2 given in [KRS04]. Let G' be an unweighted, undirected graph on n vertices. Construct a metric (V, d') by setting $d'(u, v) = 1$ if u, v is an edge of G', and $d'(u, v) = 2$ otherwise. Let the bounded degree graph G be the unweighted cycle on n vertices, C. Clearly C is of bounded degree $b = 2$ and construct the metric (U, d) in the same manner as (V, d'). It is easy to check that, if G' contains a Hamilton cycle, then an optimal bijection between (U, d) and (V, d') has distortion exactly 2. If G' does not contain a Hamilton cycle, then any bijection must have distortion at least 4. Hence the problem of finding an optimal bijection between (U, d) and (V, d') as described above is NP-Hard. Since the given instance is a particular case of the metrics in the lemma, the lemma is true. □

In this section, we prove the following in a very similar manner to the algorithm presented in [KRS04].

Theorem 3.2. *Let (U, d) be the shortest-path metric of an unweighted graph G of maximum degree b and largest cycle length 3. Let (V, d') be the shortest-path metric of an arbitrary unweighted graph G'. Then, for any fixed constants b and α, there is an $O(n^2)$ algorithm that decides whether there exists a bijection between U and V with expansion and inverse expansion at most α.*

3.1 Structural properties

We begin with a few definitions. For a subset of vertices $A \subseteq G$, let $\Gamma(A)$ denote the set of neighbors of A that lie outside A. We also use $\Gamma(v)$ to denote the set of neighbors of a vertex $v \in G$.

Definition 1. *We say that a graph G is* graphrooted *at vertex r_0 by assigning every vertex $v \in G$ a value $l(v)$ that is equal to the length of the shortest path from v to r_0 in G (with $l(r_0) = 0$). By level(i), we denote the set of all vertices v in G such that $l(v) = i$.*

Definition 2. *G_r is the subgraph rooted at vertex r according to the following definition:*

1. *r is in G_r.*
2. *If there exists a path from r to a vertex v in G such that for all vertices v' along this path (including v), $l(v') > l(r)$, then $v \in G_r$.*
3. *If (v_1, v_2) is an edge in G, and both v_1 and v_2 are $\in G_r$, then the edge (v_1, v_2) is an edge in G_r*

We now prove the following lemma (based on the proof of [KRS04] in the case where (U, d) is the shortest-path metric of an unweighted tree T of maximum degree b). Let $B(u, l)$ (resp., $B'(u, l)$) denote the closed ball of radius l around any vertex u in G (resp., in G'). For a subset of vertices $A \subseteq G$ (resp., in G'), let $\Gamma(A)$ (resp., $\Gamma'(A)$) denote the set of neighbors of A that lie outside A. Assume that G is graphrooted at an arbitrary vertex r_0. The subgraph rooted at any vertex r of G (as defined earlier) is denoted by G_r.

Lemma 3.3. *Let $\sigma : U \to V$ be a bijection with expansion and inverse expansion at most α. Then*

1. *G' has maximum vertex degree at most b^α.*
2. *For any vertex $r \in G$, each connected component of $G' \backslash B'(\sigma(r), \alpha^2)$ lies either entirely in $\sigma(G_r)$ or entirely in $G' \backslash \sigma(G_r)$.*
3. *For any $r \in G$, for any adjacent pair (u', v') in G' with $u' \in \sigma(G_r)$ and $v' \notin \sigma(G_r)$, both $\sigma^{-1}(u')$ and $\sigma^{-1}(v')$ are in $B(r, \alpha)$*

Proof. For the first statement, for any $v \in G'$, the expansion of σ^{-1} implies that $\sigma^{-1}(B'(v, 1)) \subseteq B(\sigma^{-1}(v), \alpha)$, and the cardinality of this ball is at most b^α by the degree bound on G.

For the second statement, let G_r be the subgraph graphrooted at $r \in G$. Let $v' = \sigma(v)$ be a vertex in $\Gamma'(\sigma(G_r))$. By the definition of Γ', v' is adjacent to some vertex $u' = \sigma(u)$ of $\sigma(G_r)$. From the inverse expansion bound, we have $d(u, v) \leq \alpha$. Now, assume that the shortest path from u to v goes through r. Then, clearly $d(r, v) \leq \alpha$. Thus, we have $d'(\sigma(r), v') \leq \alpha^2$. From this we get

$$\Gamma'(\sigma(G_r)) \subseteq B'(\sigma(r), \alpha^2)$$

from which we get the second statement.

For the third statement, note that by the expansion of σ^{-1}, we get that $d(\sigma^{-1}(u'), \sigma^{-1}(v')) \leq \alpha$. Again assuming that the shortest path from u to v goes through r, we get that $d(r, \sigma^{-1}(u')) \leq \alpha$ and $d(r, \sigma^{-1}(v')) \leq \alpha$.

Now, the proof of Lemma 3.4 completes this proof. □

Lemma 3.4. *Let $u \in G_r$ and $v \notin G_r$, then the shortest path from u to v goes through r.*

Proof. We shall prove this by contradiction. Suppose the shortest path from u to v does not go through r. In this case, this path has to go through a node (r') such that $l(r') \leq l(r)$ (otherwise, v is a vertex of G_r). Note that there is a path from r to r' such that any vertex w on this path ($w \neq r, r'$) has $l(w) < l(r)$. Hence, there is a path from r to r' of length at least 2 that does not overlap with the paths from u to r and u to r'. Now, consider the non-overlapping parts of the paths from u to r and u to r'. The lengths of these parts are at least 1 each and hence we get a cycle of length at least 4 (by joining the path from r to r' completely at lower levels and the path from r to r' completely at higher levels). This is a contradiction to the maximum cycle length restriction of 3 on G. Hence, the shortest path from u to v goes through r. □

We now present the algorithm, its analysis, and proof of theorem 3.2. This follows from the algorithm in the case of bounded degree trees present in [KRS04].

3.2 Algorithm and Proof of Theorem 3.2

3.2.1 Algorithm

The algorithm is a dynamic programming algorithm in the same way as given in [KRS04]. The graph G is graphrooted arbitrarily at a node r_0. The dynamic programming table T is indexed by the following parameters

1. $r \in \{u_1, ..., u_n\}$, the root of the subgraph G_r (with respect to the graphrooting G)
2. $r' \in \{v_1, ..., v_n\}$
3. An injection τ from $B(r, \alpha) \cap G_r$ into $B'(r', \alpha^2)$
4. A subset S of the vertices of G' with the property that each connected component of $G' \backslash B'(r', \alpha^2)$ lies entirely within S or entirely outside S.

An entry of the table is true if and only if there exists an injection $\sigma : G_r \rightarrow G'$ such that $\sigma(r) = r'$, σ coincides with r on $B(r, \alpha) \cap G_r$, $\sigma(G_r) = S$, and expansion of every edge of G_r and inverse expansion of every edge of $\sigma(G_r)$ are each at most α. To compute $T(r, r', \tau, S)$, we run through all combination of entries $T(r_i, r'_i, \tau_i, S_i)_i$ all of which have value true. r_i are the children of a given root r. We set the result to be true if at least one of these combinations satisfies the conditions below and false otherwise.

1. The map τ is consistent with all the maps τ_i, the τ_is are consistent among themselves, the S_i do not include r', and S is the union of the S_i plus the vertex r'.
2. For each r'_i, we have $d'(r', r'_i) \leq \alpha d(r, r_i)$.
3. For each adjacent pair v', w' in G', that belong to different sets S_i (or with $v' = r'$), both v' and w' are in the image of τ and satisfy $d(\tau^{-1}(v'), \tau^{-1}(v')) \leq \alpha$.

After all entries of the dynamic programming table are computed, the algorithm checks if some table entry $T(r_0, ., ., .)$ is true.

3.2.2 Running time and Correctness.

The degree bound on G implies that $B'(v, \alpha^2)$ has size at most b^{α^3} for any v. We claim that the size of the table T is at most

$$n \times n \times (b^{\alpha^3})^{b^\alpha} \times 2^{2b^{\alpha^3}} = O(n^2)$$

The two n terms come from the r and r' in the table. The third factor bounds the number of maps from $B(r, \alpha)$ to $B'(r', \alpha^2)$. From the second part of the lemma, we get the number of possibilities for the set S as the fourth factor. Filling the table entries takes constant time as given r and r', we only have to consider r'_i such that $r'_i \in B'(r', \alpha)$ (For further details see [KRS04]). Thus the overall running time is $O(n^2)$.

The correctness of the algorithm follows in the same way as in [KRS04] by an induction (bottom-up the levels in G). This also completes the proof of theorem 3.2.

Acknowledgments

We thank Claire Kenyon [Ken] for pointing us to the Albert-Atkinson lemma, which was also used by [KRS09] to obtain results similar to ours. This replaced our previous independent proof of a similar lemma that we needed to establish our result. We also thank Claire Kenyon, Yuval Rabani, and Alistair Sinclair for pointing to us inadvertent problems with attributions in the previous version [CMO$^+$08] of this paper which resulted in this revision. We thank William Evans for collaboration in the preliminary stages of this research. Finally, we thank Milan Stanojević for helping with the computation of the eigenvalues.

References

[AA05] M.H. Albert and M.D. Atkinson. Simple permutations and pattern restricted permutations. *Discrete Mathematics*, 300(1-3):1–15, 2005.

[AKOF03] T. Akutsu, K. Kanaya, A. Ohyama, and A. Fujiyama. Point matching under non-uniform distortions. *Discrete and Applied Mathematics*, 127:5–21, 2003.

[BBL98] P. Bose, J.F. Buss, and A. Lubiw. Pattern matching for permutations. *Inf. Process. Lett.*, 65(5):277–283, 1998.

[BCIS05] M. Badoiu, J. Chuzhoy, P. Indyk, and A. Sidiropoulos. Low-distortion embeddings of general metrics into the line. In *STOC'05*. ACM, 2005.

[BDG$^+$05] M. Badoiu, K. Dhamdhere, A. Gupta, Y. Rabinovich, H. Racke, R. Ravi, and A. Sidiropoulos. Approximation algorithms for low-distortion embeddings into low-dimensional spaces. In *SODA'05*. ACM, 2005.

[BMP02] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(4):509–522, 2002.

[CCDG82] P.Z. Chinn, J. Chvatalova, A.K. Dewdney, and N.E. Gibbs. The bandwidth problem for graphs and matrices – a survey. *Journal of Graph Theory*, 6:223–254, 1982.

[CLM03] B. Chazelle, D. Liu, and A. Magen. Sublinear geometric algorithms. In *STOC'03*. ACM, 2003.

[CMO$^+$] N. Chandran, R. Moriarty, R. Ostrovsky, O. Pandey, and A. Sahai. Improved algorithms for optimal embeddings. ECCC Technical Report number TR06-110, 2006.

[CMO$^+$08] N. Chandran, R. Moriarty, R. Ostrovsky, O. Pandey, M.A. Safari, and A. Sahai. Improved algorithms for optimal embeddings. *ACM Transactions on Algorithms*, 4(4), 2008.

[DPS02] J. Diaz, J. Petit, and M. Serna. A survey on graph layout problems. *ACM Computing Surveys*, 34(3):313–356, 2002.

[ES] W. Evans and M.A. Safari. k-separable permutations and their applications to metric embeddings, pattern matching, and stock sorting. Manuscript, September 2005.

[For96] S. Fortin. The graph isomorphism problem. Technical Report 96-20, University of Alberta, Edomonton, Alberta, Canada, 1996.

[HJ86] R.A. Horn and C.R. Johnson. *Matrix analysis*. Cambridge University Press, New York, NY, USA, 1986.

[HKW98] F. Hoffmann, K. Kriegel, and C. Wenk. Matching 2d patterns of protein spots. In *ACM Symposium on Computational Geometry*, pages 231–239. ACM, 1998.

[HP05] A. Hall and C. Papadimitriou. Approximating the distoriton. In *APPROX and RANDOM*, volume 3624 of *LNCS*, pages 111–122. Springer, 2005.

[JL03] W.B. Johnson and J. Lindenstrauss. *Handbook of Geometry of Banach Spaces*. North-Holland, 2003.

[Ken] C. Kenyon. Private communication with Safari, M.A., February 2006.

[KRS04] C. Kenyon, Y. Rabani, and A. Sinclair. Low Distortion Maps Between Point Sets. In *STOC'04*. ACM, 2004.

[KRS09] C. Kenyon, Y. Rabani, and A. Sinclair. Low distortion maps between point sets. *SIAM Journal on Computing (To Appear)*, 2009.

[Kru64a] J.B. Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Pyschometrika*, 29:1–27, 1964.

[Kru64b] J.B. Kruskal. Nonmetric multidimensional scaling: A numerical method. *Pyschometrika*, 29:115–129, 1964.

[Lin02] N. Linial. Finite metric spaces – combinatorics, geometry and algorithms. In *International Congress of Mathematicians III*, pages 573–586, 2002.

[PS05] C. Papadimitriou and S. Safra. The complexity of low-distortion embeddings between point sets. In *SODA'05*. ACM, 2005.

[Saf07] M.A. Safari. *D-Width Metric embeddings and their connections*. PhD in Computer Science, The University of British Columbia, 2007.

[She62a] R.N. Shepard. The analysis of proximities: Multidimensional scaling with an unknown distance function 1. *Pyschometrika*, 27:125–140, 1962.

[She62b] R.N. Shepard. The analysis of proximities: Multidimensional scaling with an unknown distance function 2. *Pyschometrika*, 27:216–246, 1962.

[ST93] J.H. Schmerl and W.T. Trotter. Critically indecomposable partially ordered sets, graphs, tournaments and other binary relational structures. *Discrete Math.*, 113(1-3):191–205, 1993.