Reality Distortion: Exact and Approximate Algorithms for Embedding into the Line*

Amir Nayyeri[†] Benjamin Raichel[‡]

Abstract

We describe algorithms for the problem of minimum distortion embeddings of finite metric spaces into the the real line (or a finite subset of the line). The time complexities of our algorithms are parametrized by the values of the minimum distortion, δ, and the spread, Δ, of the point set we are embedding.

We consider the problem of finding the minimum distortion bijection between two finite subsets of $\mathbb{R}$. This problem was known to have an exact polynomial time solution when δ is below a specific small constant, and hard to approximate within a factor of $\delta^{1-\epsilon}$, when δ is polynomially large. Let D be the largest adjacent pair distance, a value potentially much smaller than Δ. Then we provide a $\delta^{O(\delta^2 \log^2 D)} n^{O(1)}$ time exact algorithm for this problem, which in particular yields a quasipolynomial running time for constant δ, and polynomial D.

For the more general problem of embedding any finite metric space (X, d_X) into a finite subset of the line, Y, we provide a $\Delta^{O(\delta^2)}(mn)^{O(1)}$ time $O(1)$-approximation algorithm (where $|X| = n$ and $|Y| = m$), which runs in polynomial time provided δ is a constant and Δ is polynomial. This in turn allows us to get a $\Delta^{O(\delta^2)}(n)^{O(1)}$ time $O(1)$-approximation algorithm for embedding (X, d_X) into the continuous real line.

1 Introduction

Given two metric spaces (X, d_X) and (Y, d_Y), an ***embedding*** of X into Y is an injective map $f : X \to Y$. The ***expansion*** e_f and the ***contraction*** c_f of f are defined as follows.

$$e_f = \max_{\substack{x,x' \in X \\ x \neq x'}} \frac{d_Y(f(x), f(x'))}{d_X(x, x')}, \quad c_f = \max_{\substack{x,x' \in X \\ x \neq x'}} \frac{d_X(x, x')}{d_Y(f(x), f(x'))}.$$

The ***distortion*** of f is defined as $\delta_f = e_f \cdot c_f$.

Finding an embedding of minimum (or small) distortion is of interest due to its varied applications. Often the goal is to map a given metric space (X, d_X) into a "simpler" metric space (Y, d_Y). For example, by embedding (high dimensional) metrics into $\mathbb{R}^2$ or $\mathbb{R}^3$, one can visualize a data set, facilitating observations of patterns in the data (e.g., clusters). Moreover, if one can embed into a low dimensional space, then one gains access to a number of algorithmic tools which are prohibitively expensive in higher dimensions due to exponential dependence on the dimension.

*The full updated version of this paper is available online http://web.engr.illinois.edu/~raichel2/distortion.pdf

†School of Electrical Engineering and Computer Science; Oregon State University; nayyeria@eecs.oregonstate.edu; http://web.engr.oregonstate.edu/~nayyeria/.

‡Department of Computer Science; University of Illinois; 201 N. Goodwin Avenue; Urbana, IL, 61801, USA; raichel2@illinois.edu; http://illinois.edu/~raichel2/. Work on this paper was partially supported by the University of Illinois Graduate College Dissertation Completion Fellowship.

There are also natural applications when the two metric spaces are similar. For example, consider the problem of finding a minimum distortion bijection between two finite subsets of $\mathbb{R}^d$. In this case there are natural applications to problems such as shape matching and object recognition.

Due to its varied applications, the problem of computing small distortion embeddings of finite metric spaces has been studied extensively. Consequently, it is an area rich with results, perhaps most notably Bourgain's theorem stating that any finite metric space is embeddable into $\mathbb{R}^{O(\log^2 d)}$ (with the standard ℓ_2 norm) with $O(\log n)$ distortion [Bou85, LLR94]. Rather than giving an exhaustive list, we refer the reader to [Ind01] and [Mat13], for a list of applications, results, and a general introduction to the area of metric embeddings.

Finite Metrics into $\mathbb{R}$.

Here we focus on the specific case of embeddings of finite metric spaces into a subset of the Euclidean line. Results in this area fall into two main categories, namely, embeddings into the continuous Euclidean line or bijections into a finite subset of the line.

Bijections. Consider the case of finding the minimum distortion bijection between two finite subsets of $\mathbb{R}$. Kenyon et al. [KRS04] was the first to study this problem, providing a polynomial time exact algorithm when the minimum distortion, δ, is $< 3 + 2\sqrt{2}$. Furthering the dynamic programming approach of Kenyon et al., Chandran et al. [CMO$^+$08] provided a polynomial time exact algorithm when $\delta < 13.602$, and showed this approach breaks for $\delta > 13.928$. Surprisingly, these exact algorithms for the case when the distortion is smaller than a specific constant, are the only positive results that the authors are aware of for finding the minimum distortion bijection between two finite subsets of $\mathbb{R}$. Moreover, they also appear to be the only positive results for embedding more general metric spaces into a finite subset of $\mathbb{R}$.

On the other hand, a number of hardness results are known for bijections between point sets in $\mathbb{R}^d$, for $d = 1$ and higher. Hall and Papadimitriou [HP05] showed (among other things) that for any fixed dimension $d \geq 1$, if the distortion is polynomially large, i.e. $\delta = n^\epsilon$, then the optimal bijection is hard to approximate within a factor of $\delta^{1-\epsilon}$. Additionally, for $d = 2$ Edmonds et al. [ESZ10] showed that it is hard to distinguish between the cases $\delta < 3.61 + \epsilon$ or $\delta > 4 - \epsilon$, and for $d = 3$ Papadimitriou and Safra [PS05] showed that δ is hard to approximate within a factor of 3.

These results left open the question of whether efficient (exact or approximate) algorithms for intermediate values of the distortion are possible. In this paper, we take our first step toward answering this question by providing the following exact algorithm.

Theorem 1.1. *Let $X, Y \subset \mathbb{R}$ be point sets of cardinality n, such that the closest pair of points in X are distance 1 apart and the furthest adjacent pair are distance D apart. Then there is a $\delta^{O(\delta^2 \log^2 D)} n^{O(1)}$ time exact algorithm to compute an (X, Y)-bijection of minimum distortion, δ.*

Interpreting the above result, for any constant distortion δ, if D is polylogarithmic in size we get a polynomial running time and if D polynomial in size then we get a quasi-polynomial running time. Also note that D is the furthest distance between an *adjacent* pair, a value which is potentially a linear factor smaller than the ***spread***, Δ, defined as the ratio of the largest and smallest inter-point distances.

On the approximation side, our technique also yields a $O(1)$-approximation algorithm for the minimum distortion bijection when (X, d_X) is allowed to be any finite metric space. This result, however, more naturally fits into the setting of embeddings into $\mathbb{R}$.

Embedding into $\mathbb{R}$. Another category of results consider minimum distortion embeddings from more general metric spaces into the Euclidean line. Here the results can be further broken down into two cases, based on whether (X, d_X) is allowed to be any general finite metric space, or some restricted subclass. Specifically, for the case when (X, d_X) is restricted to be (the shortest path metric induced by) an unweighted graph, Badiou et al. [BDG⁺05] provided an $O(\sqrt{n})$-approximation algorithm and an $n^{O(\delta)}$ exact algorithm, which later Fellows et al. [FFL⁺13] improved to an $O(n\delta^4(2\delta + 1)^{2\delta})$ time exact algorithm. For the case of unweighted trees, Badiou et al. [BDG⁺05] provided an $\tilde{O}(n^{1/3})$-approximation. For weighted trees, Badiou et al. [BCIS05] provided a $\delta^{O(1)}$-approximation and showed that the minimum distortion embedding is hard to approximate within a polynomial factor in n for weighted tree metrics with $\Delta = n^{O(1)}$, where Δ denotes the spread of (X, d_X).

For general metric spaces, Badiou et al. [BCIS05] provided an $O(\Delta^{3/4}\delta^{O(1)})$-approximation algorithm (and their hardness result for weighted trees applies here as well). Perhaps the most relevant result for the current work, is the result of Fellows et al. [FFL⁺13], who provided a fixed parameter tractable algorithm with running time $n(\delta\Delta)^4(2\delta+1)^{2\delta\Delta}$. Moreover, they show that the exponential dependence on Δ is unavoidable for any exact algorithm. However, this did not rule out the possibility of an approximation algorithm with polynomial dependence on Δ. In particular, we show the following more general result.

Theorem 1.2. *There is a $\Delta^{O(\delta^2)}(mn)^{O(1)}$ time $O(1)$-approximation algorithm to compute a minimum distortion embedding of a metric space of cardinality n into a point set on the Euclidean line of cardinality m, where δ is the minimum distortion, Δ is the spread, and $m \geq n$.*

In particular, our results imply a polynomial time constant factor approximation algorithm, provided the distortion is constant and the spread is polynomial. Note that this result is strictly more general than the case when Y is the entire real line. Specifically, if one can approximately embed (X, d_X) into the integer line, then at the cost of an additional small constant factor, one can approximately embed (X, d_X) into the real line.

Theorem 1.3. *There is a $\Delta^{O(\delta^2)}(n)^{O(1)}$ time $O(1)$-approximation algorithm to compute a minimum distortion embedding of a metric space of cardinality n into the Euclidean line, where δ is the minimum distortion and Δ is the spread.*

Significance of our work. For the problem of finding the minimum distortion bijection between point sets on the line, previous results were on one of two extremes, either providing exact algorithms for $\delta < 14$, or strong inapproximability results for polynomially large δ. Thus our work bridges this knowledge gap by showing that for larger constant values of distortion, the problem is still polynomial time solvable, provided a polynomial spread. Our results are in part facilitated by parameterizing on the value D (the largest adjacent pair distance) for the $\mathbb{R}$ bijection problem, and the spread, Δ, for the problem of embedding into the line. Previous work on the embedding problem, such as the work of Fellow et al. [FFL⁺13], also achieved positive results by parameterizing on Δ. However, the key distinction is that in our work rather than Δ appearing in the exponent, only $\log^2 D$ or $\log \Delta$ appears in the exponent. Note that in the later case $\log \Delta$ is achieved at the cost of a constant factor approximation, however, the work of Fellow et al. also implies that such a running time requires that the algorithm be approximate.

For the embedding problem, our work also differs in another significant way. Previous work considered embedding either into the continuous real line or the integer line (which are related problems). Our work on the other hand more generally allows the subset of points on the line to have arbitrary spacing. In particular the scales of adjacent point distances can differ dramatically.

Note on the running time and approximation quality. In this paper we focus on improving the allowable ranges of the values δ, D, and Δ, which lead to polynomial running times, rather than optimizing the precise constant of the polynomial. In particular, when stating running times exponents are often written using $O(\cdot)$ notation. Additionally, in our $O(1)$-approximation section, rather than optimizing the precise value of the constant in the approximation, we make some assumptions which degrade the constant, but significantly simplify the presentation. As such, both the constants in the running time as well as in the approximation, can be improved.

2 Overview

We now provide a high level overview of our approach. The main idea is to use a sliding window, as was done by Badiou et al. [BCIS05] and Fellows et al. [FFL+13].

Bijections in $\mathbb{R}$. Suppose there was a non-contracting bijection f between $X, Y \subsetneq \mathbb{R}$ of expansion δ.[1] Consider the function f restricted to some interval on the Y side. Here we call this interval a window as it provides a partial view of the function f. Now we don't actually know the function f (or even if it exists), but if this window is small enough then one can simply guess all possible functions into the window. The goal then is to determine the entire function f by stitching together views as we slide our window from the leftmost point in Y all the way to the rightmost point. In order for this stitching process to work, we need our windows to be stateless (otherwise we end up guessing all possible bijections). In other words, guessing how f maps into a given window should allow us to partition X into three sets: (1) points mapped into the window by f, (2) points that must map to the left of the window by f, and (3) points that must map to the right of the window by f. To achieve this tri-partition property, the size of the sliding window should be sufficiently large. Specifically, if the window is at least δD wide, then no adjacent points in X can map to different sides of the window (as D is the distance between the furthest adjacent pair in X, and no pair can be expanded by more than δ). However, now we have a tug of war on the size of our window. We need the window to be large to break up the problem, and we need it to be small to guess all partial maps into the window. Specifically, (as we show later) one can assume the smallest inter-point distance is 1, and so the total number of maps into a δD length interval is $(\delta D)^{O(\delta D)} n^{O(1)}$, which is far too expensive.

The real difficulty is the various competing scales of adjacent pair distances. Specifically, if $D = O(1)$ then this approach yields an efficient algorithm (for constant δ), since windows can then be very small. On the other hand, if we have just one really large adjacent pair distance, our window must be large. To remedy the situation, we use the scales of adjacent pairs in X to determine the size of the windows they map into. Specifically, we view X as a graph (a path), and assign scales to edges based on their length. Namely, an edge of scale s has its length in $(\delta^{s-1}, \delta^s]$. The vertices (i.e. points in X) inherit the scales of their incident edges. Next, observe that only $O(\delta^2)$ points of scale s can map into an interval of length $2\delta^{s+1}$, which we call a window of scale s. Also observe that such a window is wide enough such that two points in X at distance $\leq \delta^s$ apart, cannot map to the left and right of such an interval (otherwise, the expansion is too large). So windows of length $2\delta^{s+1}$ give us a tri-partition for points at scale s, but what about points at other scales? To that end, we consider a collection of maps into a tower of concentric windows at all scales, which we call a multi-scale window. Our algorithm then slides this multi-scale window from left to right to view the pre-image of Y. Using this multi-scale window allows us to save

[1] Later on it is proven why one can assume the bijection is non-contracting.

a $D/\log^2 D$ factor in the exponent when compared to the single large window approach, overall yielding a $\delta^{O(\delta^2 \log^2 D)} n^{O(1)}$ running time.

General Metric Spaces. Now consider the problem of embedding any finite metric space (X, d_X) into a (potentially larger) finite subset $Y \subsetneq \mathbb{R}$. We want to adapt the algorithm above to this more general setting. However, now there is an issue. Previously, we used adjacency relations between points on the line to define scales. In general metric spaces, however, we lose the notion of points in X being adjacent along the line, and so we need a new way to define scales. Here we solve this problem by using the permutation of X defined by the standard Gonzalez algorithm for k-center clustering. Specifically, this permutation is defined recursively by setting x_1 to be an arbitrary starting point, and for all $1 < i \leq n$, defining x_i as the point furthest from the set $S_{i-1} = \{x_1, \ldots, x_{i-1}\}$. Clearly $d(x_i, S_{i-1}) > d(x_j, S_{j-1})$ for $i < j$, and so the intuition is that points earlier on in this permutation have larger scale. So the natural question then is if the Gonzalez permutation gives a way to define scales, why was it not also used for the case when $X \subsetneq \mathbb{R}$? To answer this, let x, x' be the pair in X realizing the expansion. The issue is that as we slide our multi-scale window, the Gonzalez permutation does not guarantee there will be a single moment in time in which both x and x' are seen in our window.[2] However, for the Gonzalez permutation one can at least argue that there is some moment in time where our window sees a pair $z, z' \in X$ such that x is close the z and x' is close to z'. (Specifically, we argue the pairs we see in X form an $O(1)$-spanner.)

By the above, we will have to settle for our algorithm being an $O(1)$-approximation to the optimal embedding of (X, d_X) into $Y \subsetneq \mathbb{R}$. However, now that we have moved to approximation, one can speed up our algorithm as follows. Previously, we used the fact that the mapping we are looking for is non-contracting, so points at scale s must be at least δ^s apart from one another in the image of f, producing a bound on the number of scale s points which can map into a window of length $2\delta^{s+1}$. However, a priori we do no which points in this window get mapped onto by scale s points, and such a window could potentially contain many points of Y. Previously we had to try all possibilities as the goal was an exact algorithm. However, in the approximate setting one can instead breakup this interval into δ^s length bins (each of which can contain the image of at most one scale s point), and map to these approximate bins rather than directly to points.

In order to make this idea go through, a lot machinery must be set up, which previously was not needed. However, there are two key advantages to all this extra work. First, our algorithm will now run in polynomial time (as opposed to quasipolynomial) when the spread is polynomial. The second, more surprising advantage, is that this additional machinery makes for a near seamless transition from the case of bijections to case of embeddings, either into larger finite point sets or the entire real line.

3 Preliminaries

Maps. Let A and B be two sets. A partial map f from A to B is denoted by $f : A \nrightarrow B$. The domain of f, denoted by $\mathrm{Dom}(f)$, is the set of all $a \in A$ for which $f(a)$ is defined. So, $\mathrm{Dom}(f) \subseteq A$. The Image of f, denoted by $\mathrm{Im}(f)$, is the set of all $b \in B$ such that $b = f(a)$ for some $a \in A$. So, $\mathrm{Im}(f) \subseteq B$. In the special case that $A = \mathrm{Dom}(f)$, we call f a total map, or simply a map, and we denote it by $f : A \to B$.

[2]Intuitively, this was inevitable, as X no longer lies on a line, and so one should not be able to define scales in a way that imposes a linear adjacency relationship on X.

Distortion. Let (X, d_X) and (Y, d_Y) be two metric spaces. An ***embedding*** of X into Y is an injective map $f : X \to Y$. The ***expansion*** e_f and the ***contraction*** c_f of f are defined as follows.

$$e_f = \max_{\substack{x,x' \in X \\ x \neq x'}} \frac{d_Y(f(x), f(x'))}{d_X(x, x')}, \quad c_f = \max_{\substack{x,x' \in X \\ x \neq x'}} \frac{d_X(x, x')}{d_Y(f(x), f(x'))}.$$

The ***distortion*** of f is defined as $\delta_f = e_f \cdot c_f$. It follows by definition that distortion is invariant under scaling of either of the sets. We say that the pair (x, x') ***realizes*** the expansion of the map f, if $d(f(x), f(x'))/d(x, x') = e_f$. Similarly, we say that the pair (y, y') realizes the contraction of the map f, if $d(f^{-1}(y), f^{-1}(y'))/d(y, y') = c_f$.

Basic distortion facts. Variants of the following two lemmas can be found in previous papers (e.g., [KRS04]), and we include proofs for the sake of completeness. For a finite subset $A \subsetneq \mathbb{R}$, $a, a' \in A$ are said to be adjacent in A if the interval $(a, a') \cap A = \emptyset$. The first lemma tells us that for embedding finite point sets into the line (or a subset of the line), one only needs to look at adjacent pairs in the image to know the contraction, and moreover if embedding from a finite subset of the line then one only needs to look at adjacent pairs in the domain to know the expansion.

Lemma 3.1. *Let f be an embedding of X into Y with expansion e_f and contraction c_f. The following properties hold.*

1. *If X is a finite subset of $\mathbb{R}$, then there are adjacent points $x, x' \in X$ that realize the expansion.*
2. *If $\mathrm{Im}(f) \subseteq Y$ is a finite subset of $\mathbb{R}$, then there are points $y, y' \in \mathrm{Im}(f)$ which are adjacent in $\mathrm{Im}(f)$ and that realize the contraction.*

Proof: Let $X = \{x_1, \ldots, x_n\} \subsetneq \mathbb{R}$, and suppose, to derive a contradiction, that the expansions of all adjacent pairs are smaller than e_f.

$$\text{for all } \ 1 \leq i < n, \qquad d(f(x_i), f(x_{i+1})) < e_f d(x_i, x_{i+1}) \tag{1}$$

Since the expansion of the map is e_f there are x_p and x_q such that $d(f(x_p), f(x_q))/d(x_p, x_q) = e_f$. Thus, we have:

$$e_f = \frac{d(f(x_p), f(x_q))}{d(x_p, x_q)} = \frac{d(f(x_p), f(x_q))}{\sum_{i=p}^{q-1} d(x_i, x_{i+1})} \leq \frac{\sum_{i=p}^{q-1} d(f(x_i), f(x_{i+1}))}{\sum_{i=p}^{q-1} d(x_i, x_{i+1})} < \frac{\sum_{i=p}^{q-1} e_f d(x_i, x_{i+1})}{\sum_{i=p}^{q-1} d(x_i, x_{i+1})} = e_f$$

The first inequality is implied by triangle inequality, and the second one is implied by (1). The derived contradiction implies the first part of the lemma statement. The proof for the second part is similar. Essentially, change x's to y's and f to f^{-1} in the previous proof. $\square$

An embedding is ***non-contracting*** if the contraction is ≤ 1. The following lemma allows us to restrict our attention to non-contracting embeddings of expansion $\leq \delta$, where δ is a known value.

Lemma 3.2. *Let (X, d_X) and (Y, d_Y) be finite metric spaces of sizes n and m, respectively. Then the problem of finding an embedding of X into Y with minimum distortion reduces to solving $(mn)^{O(1)}$ instances of the following problem: given a real value $\delta \geq 1$, compute a non-contracting embedding of X into Y with expansion at most δ, or correctly report that no such embedding exists.*

Proof: Let f be any optimal map with distortion $\delta = \delta_f$. Let $x, x' \in X$ and $y, y' \in Y$ realize the expansion e_f and contraction c_f, respectively. Let $Y = \{y_1, \ldots, y_m\}$, and let $Y' = \{c_f y_1, \ldots, c_f y_n\}$ be this set after scaling by c_f. Consider the function $f' : X \to Y'$, in which $f'(x) = c_f f(x)$, for any $x \in X$. Clear the contraction, $c_{f'}$, of this function is 1. As the distortion is invariant under scaling, it then must be that $e_{f'} = \delta$. Therefore, one can look for f' instead, after scaling Y.

As we do not know the function f, we must guess x, x', $f(x)$, $f(x')$, y, y', $f^{-1}(y)$ and $f^{-1}(y')$. However, there are $(mn)^{O(1)}$ choices for these values, which proves the lemma. □

4 An Exact Fixed-parameter Tractable Algorithm

Let $X = \{x_1, x_2, \ldots, x_n\}$ and $Y = \{y_1, y_2, \ldots, y_n\}$ be two point sets on the Euclidean line. To simplify explanations, we assume all members of X and Y are positive. The pairs (X, d) and (Y, d) are metric spaces, where d is the Euclidean distance on the line, that is for any $p, q \in \mathbb{R}^+$, $d(p, q) = |p - q|$. As scaling either X or Y does not affect the distortion, we assume for all $i \in \{1, 2, \ldots, n-1\}$, $1 \le d(x_i, x_{i+1}) \le D$, and $\min_i d(x_i, x_{i+1}) = 1$.

In this section we describe a fixed-parameter tractable algorithm for computing the minimum distortion bijection between X and Y. By Lemma 3.2, it suffices to instead solve the problem of finding a non-contracting bijection of expansion at most δ, where $\delta \ge 1$ is a known value (we will also detect when such a bijection does not exist).

4.1 Windows, Scales, and Partial Maps

Scales. The ***scale*** of an ***edge*** (x_i, x_{i+1}) $(1 \le i < n)$ is s if and only of $d(x_i, x_{i+1}) \in (\delta^{s-1}, \delta^s]$. Similarly, the ***scale*** of an ***edge*** (y_i, y_{i+1}) $(1 \le i < n)$ is s if and only of $d(y_i, y_{i+1}) \in (\delta^{s-1}, \delta^s]$. In turn, a point has scale s if it is adjacent to an edge of scale s. So, a point has one or two scales. For each $s \in \{0, 1, \ldots, S = \log_\delta(D)\}$, sets X_s and Y_s are composed of the scale s vertices of X and Y, respectively. Note that for $s \ne t$, X_s and X_t are not necessarily disjoint.

Windows. Let $B(c, r) = [c - r, c + r]$ be the ball of radius r centered at c, and let $Y[c, r] = B(c, r) \cap Y$. We refer to the set $Y[c, \delta^{s+1}]$ as a ***window*** of scale s with center c. In turn, we refer to the collection of concentric windows of scales 0 to S, $W_c = \{Y[c, \delta], Y[c, \delta^2], \ldots, Y[c, \delta^{S+1}]\}$, as a ***multi scale window*** with center c.

The multi scale window $W_{c'}$ ***succeeds*** the multi scale window W_c if there exists an $0 \le s \le S$ with the following properties.

1. For all $0 \le i \le S$, $i \ne s$, we have $Y[c, \delta^{i+1}] = Y[c', \delta^{i+1}]$.

2. Either $Y[c, \delta^{s+1}] = Y[c', \delta^{s+1}] \cup \{c - \delta^{s+1}\}$ or $Y[c', \delta^{s+1}] = Y[c, \delta^{s+1}] \cup \{c' + \delta^{s+1}\}$.

Intuitively, the set of points at all scales except s are identical. At scale s, they differ in one point, which is either $c - \delta^{s+1}$ or $c' + \delta^{s+1}$. Sometimes, we refer to this point as the ***difference*** of W_c and $W_{c'}$.

It is possible to find a sequence of multi-scale windows $\mathcal{W} = (W_{c_1}, W_{c_2}, \ldots, W_{c_k})$ with the following properties.

1. For each $i \in \{2, 3, \ldots, k\}$, W_{c_i} succeeds $W_{c_{i-1}}$.

2. $c_1 = y_1$ and $c_k = y_n$.

Specifically, let $\mathcal{C}' = \{y_i - \delta^{s+1}$ and $y_i + \delta^{s+1} \mid i \in \{1, \ldots, n\}$ and $s \in \{0, \ldots, S\}\}$. Then the set of centers of the multi-scale windows is the set $\mathcal{C} = (\mathcal{C}' \cup \{y_1, y_n\}) \cap [y_1, y_n]$.[3] Since in each succession a point either enters or leaves the window of some scale, we have $k = |\mathcal{C}| = O(nS)$. We fix such a sequence $\mathcal{W}$ for the rest of this paper, and refer to it as the ***standard sequence of multi scale windows***. Figure 1 illustrates a prefix of this sequence for an example with two scales and five points.

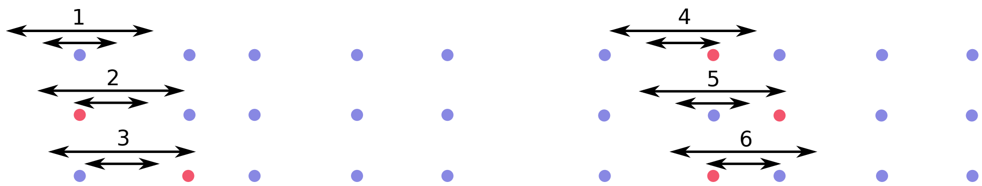

Figure 1. A prefix of the sequence of multi scale windows. The red point is entering or leaving a single scale window.

Partial maps. A ***scale s partial map*** into the neighborhood of c is a non-contracting partial map $f_s : X_s \twoheadrightarrow Y[c, \delta^{s+1}]$ with expansion at most δ.

Remark 4.1. *Note that f_s being non-contracting implies that f_s is one-to-one, since if two points were to map to the same $y \in Y$ then the contraction would be infinite. The same will be true in the definition of multi-scale partial maps below.*

Lemma 4.2. *Let f_s be a partial map at scale s and into the neighborhood of c, for $s \in \{0, 1, 2, \ldots, S\}$, and $c \in \mathbb{R}$. We have, $|\mathrm{Dom}(f_s)| = |\mathrm{Im}(f_s)| = O(\delta^2)$.*

Proof: Let $\mathrm{Dom}(f_s) = \{x_{\iota_1}, x_{\iota_2}, \ldots, x_{\iota_k}\}$, where $x_{\iota_1} < x_{\iota_2} < \ldots < x_{\iota_k}$. Since, each of these points is adjacent to at least one edge of scale s, for any $1 \leq i \leq k-2$, we have $d(x_{\iota_i}, x_{\iota_{i+2}}) \geq \delta^{s-1}$. Since f_s is non-contracting, we have $d(f_s(x_{\iota_i}), f_s(x_{\iota_{i+2}})) \geq \delta^{s-1}$. Consequently, $d(f_s(x_{\iota_1}), f_s(x_{\iota_k})) \geq \lfloor k/2 \rfloor \delta^{s-1}$. Since both x_{ι_1} and x_{ι_k} are mapped into $Y[c, \delta^{s+1}]$, a length $2\delta^{s+1}$ interval, $k = O(\delta^2)$. ☐

Multi-scale map. Let W_c be a multi-scale window in the standard sequence of multi-scale windows, $\mathcal{W}$. A function $F : X \twoheadrightarrow Y$ is a ***multi-scale partial map*** into W_c if:

1. $F = f_1 \cup f_2 \cup \ldots \cup f_S \cup g$, such that f_i is a scale i partial map into c.

2. If $z_r = \min(\{y \in Y \mid y > c + \delta\})$ is defined, then $\mathrm{Im}(g) = \{z_r\}$ and $|\mathrm{Dom}(g)| = 1$, otherwise g is empty. We refer to z_r as the ***center-right point*** of W_c (or of F).

3. $Y[c, \delta] \subseteq \mathrm{Im}(F)$.

4. F is non-contracting and has expansion at most δ.

Remark 4.3. *Let $h : A \to B$ be a non-contracting bijection of expansion $\leq \delta$. Then for any subset $C \subseteq A$, the restriction $f \restriction_C$ is a non-contracting bijection of expansion at most δ. In particular, the requirements that F is non-contracting and expansion at most δ implies these requirements on the partial maps. However, for clarity we include these redundant requirements for partial maps.*

[3]Technically, depending on the distances between adjacent points in Y, more than one window may change between W_{c_i} and $W_{c_{i+1}}$. This is easily resolved by slightly increasing the width of each window by some small ϵ.

Lemma 4.4. *Let W_c be any multi-scale window in $\mathcal{W}$. There are $\delta^{O(S^2\delta^2)} \cdot n$ multi-scale partial maps into W_c.*

Proof: Let y be the point in $\mathrm{Im}(F)$ which is closest c. We can assume $y \in [c, \delta^{S+1}]$, as otherwise $F = g$, and $\mathrm{Dom}(g)$ is either empty or one of $O(n)$ possibilities from X. As y is the closest point to c in $\mathrm{Im}(F)$, it must be contained in all windows in W_c with non-empty intersection with $\mathrm{Im}(F)$ (since the windows are concentric and all have center c). Let t be the index of the smallest window such that $y \in Y[c, \delta^{t+1}]$. For any $s \geq t$, we have $\mathrm{Im}(f_s) \subseteq Y[c, \delta^{s+1}] \subseteq Y[y, 2\delta^{s+1}]$.

Let $h : F^{-1}(y) \to \{y\}$, and for all $f_i \in F$, let $f'_i = f_i \cup h$. By Remark 4.3, since F is non-contracting and has expansion at most δ, the same holds for all f'_i. Therefore, $\mathrm{Dom}(f_s) \subseteq \mathrm{Dom}(f'_s) \subseteq X[F^{-1}(y), 2\delta^{s+1}]$. Since the minimum distance between any pair of points is at least one, we have $|Y[y, 2\delta^{s+1}]| = O(\delta^{s+1})$, and $|X[F^{-1}(y), 2\delta^{s+1}]| = O(\delta^{s+1})$. Lemma 5.3 implies that the size of the domain (and also the image) of f_s is $O(\delta^2)$. Therefore, the total number of possibilities for f_s, once y is fixed, is as follows.

$$\binom{O(\delta^{s+1})}{O(\delta^2)}\binom{O(\delta^{s+1})}{O(\delta^2)}O(\delta^2)^{O(\delta^2)} = \delta^{O(S\delta^2)}$$

As $y \in [c, \delta^{S+1}]$, there are at most δ^{S+1} possibilities for y, and at most n possibilities for the pre-image of y. Once y and its pre-image are fixed, there are $\delta^{O(S\delta^2)}$ number of possibilities for each f_s. Therefore, overall, the number of multi-scale partial maps is bounded by:

$$\left(\delta^{O(S\delta^2)}\right)^S \cdot \delta^{S+1} \cdot n = \delta^{O(S^2\delta^2)}n$$

□

4.2 Left, Right, and Center

Feasibility. Call a bijection $f : X \to Y$ ***feasible*** if it is non-contracting and has expansion at most δ. A multi-scale partial map $F = f_1 \cup \cdots \cup f_S \cup g$ into W_c, is called a ***view*** of f in W_c if for all s, f_s is the restriction $f \restriction_{X_s}$ whose range is restricted to the interval $Y[c, \delta^{s+1}]$ (formally $f_s = f \restriction_{f^{-1}(Y[c,\delta^{s+1}]) \cap X_s}$), and $g : \{f^{-1}(z_r)\} \to \{z_r\}$. (Note that by Remark 4.3, the non-contracting and expansion $\leq \delta$ properties of multi-scale partial maps are satisfied.) In this case we call f a ***feasible extension*** of F. Similarly, we call F ***feasible*** if there exists a feasible f, such that F is a view of f.

For a point $x \in X$, let $\mathsf{ms}(x)$ be the larger of the (at most) two scales that x participates in. For any multi-scale partial map, $F = f_1, \ldots, f_S \cup g$, with a feasible extension f, let $L_f = L_f(F) = \{x \in X \mid f(x) < c - \delta^{\mathsf{ms}(x)+1}\}$, $C_f = C_f(F) = \mathrm{Dom}(F)$, and $R_f = R_f(F) = \{x \in X \mid f(x) > c + \delta^{\mathsf{ms}(x)+1}\} \setminus \mathrm{Dom}(g)$. (Note that for R we must remove $\mathrm{Dom}(g)$, as $\mathrm{Dom}(g) \subseteq \mathrm{Dom}(F)$ and it may be that $z_r > c + \delta^{\mathsf{ms}(f^{-1}(z_r))+1}$.) We have the following.

Lemma 4.5. *Let F be a feasible multi-scale partial map into $W_c \in \mathcal{W}$. Let f and f' be two distinct feasible extensions of F, then:*
1) $C_f \neq \emptyset$.
2) The sets L_f, C_f, R_f, form a tri-partition of X.
3) $L_f = L_{f'}$ and $R_f = R_{f'}$.

Proof: We first prove $C_f \neq \emptyset$. For any $W_c \in \mathcal{W}$, it holds that $Y \cap [c, y_n] \neq \emptyset$. As such, the definition of multi-scale partial maps requires that $C_f \neq \emptyset$. Specifically, either $y_n > c + \delta$ in which case z_r and hence g is defined, or $y_n \in Y[c, \delta]$ and so $y_n \in \mathrm{Im}(F)$.

Now we prove that L_f, C_f, R_f is a partition of X. Consider a point $x \in C_f = \text{Dom}(F)$. If $x = f^{-1}(z_r)$ then by definition $x \notin L_f, R_f$. For any other point $x \in C_f = \text{Dom}(F)$, since F is a multi-scale partial map, $F(x) \in Y[c, \delta^{\mathsf{ms}(x)+1}]$, and therefore $C_f \cap L_f = C_f \cap R_f = \emptyset$. Additionally, trivially $L_f \cap R_f = \emptyset$. On the other hand since f is a feasible extension of F, if $f(x) \in Y[c, \delta^{\mathsf{ms}(x)+1}]$, then $x \in C_f = \text{Dom}(F)$, and hence for any $x \in X$, $x \in L_f \cup C_f \cup R_f$.

We now prove that $L_f = L_{f'}$, the case for R will then follow by a similar argument. Suppose otherwise that $L_f \neq L_{f'}$, and without loss of generality assume that there exists some $x \in X$ such that $x \in L_f$ and $x \notin L_{f'}$. Since f and f' are both feasible extensions of F, $C_f = C_{f'}$. Therefore, if $x \in L_f$, since L_f, C_f, R_f, is a tri-partition, then $x \notin C_{f'}$. Therefore, $x \in L_f$ and $x \in R_{f'}$.

Let a be a neighbor of x. If $f(a) \in Y[c, \delta^{\mathsf{ms}(a)+1}]$, then since F is a view of f, it holds that $a \in C_f = C_{f'}$ and so $f(a) = f'(a)$. Similarly if $f'(a) \in [c, \delta^{\mathsf{ms}(a)+1}]$, then $f(a) = f(a')$. However, if $f(a)$ lies to the right of c then $d(f(a), f(x)) > \delta d(x, a)$ and if $f'(a)$ lies to the left of c then $d(f'(a), f'(x)) > \delta d(x, a)$. Therefore $a \notin C_f = C_{f'}$. An even simpler argument also implies that $a \notin R_f$ and $a \notin L_{f'}$. So since L, C, R is a tri-partition, it has been argued that, for any $x \in L_f$ and $x \in R_{f'}$, the immediate neighbors of x in either direction must also be in L_f and $R_{f'}$. Hence if this argument is recursively applied to x's neighbors, then one can conclude $L_f = X$ and $R_{f'} = X$. However, $C_f = C_{f'} \neq \emptyset$, and so we get a contradiction. □

The above lemma implies that all feasible extensions of F induce the same tri-partition of X. Specifically, $X = L \cup C \cup R$, which we call the ***left***, ***center***, and ***right*** of F, respectively. The center is the domain of F, i.e. $C = \text{Dom}(F)$, and the left set, L, and right set, R, are as defined in the above lemma.

Remark 4.6. *Let $\mathcal{W} = \{W_{c_1}, \ldots, W_{c_k}\}$ be the standard sequence of multi-scale windows. Let F_1 be a feasible multi-scale partial map into W_{c_1}, then $L(F_1) = \emptyset$. Specifically, $c_1 = y_1$, hence $Y \cap (-\infty, c_1) = \emptyset$, and so $L(F_1) = \{x \in X \mid f(x) < c_1 - \delta^{\mathsf{ms}(x)+1}\} = \emptyset$ (where f is any feasible extension). A similar logic implies that for a feasible F_k into $W_{c_k} \in \mathcal{W}$, $R(F_k) = \emptyset$.*

Below we describe an algorithm which given a feasible F, outputs the tri-partition $L \cup C \cup R$. If F is not feasible, then ideally this could be detected, however being able to do so without knowing the extension f seems unlikely. Therefore, if F is not feasible the algorithm either returns that F is infeasible or outputs some bogus tri-partition.

Lemma 4.7. **LCR(F)**: *Given a multi-scale partial map F into $W_c \in \mathcal{W}$, there is a polynomial time algorithm such that if F is feasible it outputs the corresponding partition of X into sets $L \cup C \cup R$, as described above. If F is not feasible it either outputs a tri-partition $L \cup C \cup R$ or returns that F is infeasible.*

Proof: We say the status of a point in X is known if we know which set in $L \cup C \cup R$ it belongs to. We now prove that we can determine the status of any $x \in X$ if it is adjacent to some x' whose status is known. By induction this will imply we can determine that status of all points in X. Note that we must know the status of at least one point in X since $\text{Dom}(F) = C \neq \emptyset$, by Lemma 4.5.

Let f be a feasible extension of F. The key observation is that for any $x \in X$, if $x \notin \text{Dom}(F)$ then $f(x) \notin Y[c, \delta^{\mathsf{ms}(x)+1}]$. So let x' be a point whose status is known, and let x be a point adjacent to x' whose status is unknown. Let s be the scale of the edge xx'. Without loss of generality, suppose $f(x')$ lies to the right of c. In this case $f(x) \geq c - \delta^{s+1}$, as otherwise the edge xx' was stretched more than δ. Since $x \notin \text{Dom}(F)$, we know $f(x) \notin Y[c, \delta^{\mathsf{ms}(x)+1}]$, and therefore can conclude $f(x) > \delta^{\mathsf{ms}(x)+1}$, i.e. $x \in R$. For the case when x' lies to the left of c, one can similarly conclude that $x \in L$.

After defining the sets $L \cup C \cup R$, the algorithm performs a couple of simple consistency checks. First, the algorithm verifies $L \cup C \cup R$ is a tri-partition, and otherwise returns infeasible. Note the check will not return infeasible for any feasible F, since by Lemma 4.5 any feasible F induces such a tri-partition. Next, for all adjacent $x_i, x_{i+1} \in X$, the algorithm checks and returns infeasible if either $x_i \in L, x_{i+1} \in R$ or $x_{i+1} \in L, x_i \in R$. This check is valid since any feasible F cannot map adjacent x_i, x_{i+1} to L and R, as this would violate that the expansion is at most δ. Specifically, let s be the scale of the edge x_i, x_{i+1}, then $\mathsf{ms}(x_i), \mathsf{ms}(x_{i+1}) \geq s$, however, if they were mapped to the left and right then $d(f(x), f(x)) > 2\delta^{s+1}$ (for any feasible extension f). □

We use $\mathsf{LCR}(F)$ to denote the algorithm of the above lemma which for any multi-scale partial map, F, either returns a tri-partition $L \cup C \cup R$ or that F is infeasible.

Succession of maps. Let $W_{c_{i+1}}$ succeed W_{c_i} in the standard sequence $\mathcal{W}$. By definition, there is exactly one scale s such that $Y[c_i, \delta^{s+1}]$ and $Y[c_{i+1}, \delta^{s+1}]$ differ. Specifically, either $Y[c_i, \delta^{s+1}] = Y[c_{i+1}, \delta^{s+1}] \cup \{c_i - \delta^{s+1}\}$ or $Y[c_{i+1}, \delta^{s+1}] = Y[c_i, \delta^{s+1}] \cup \{c_{i+1} + \delta^{s+1}\}$. In the former case we say $W_{c_{i+1}}$ ***drops*** the point $y = \{c_i - \delta^{s+1}\}$, and in the latter case we say $W_{c_{i+1}}$ ***adds*** the point $y' = \{c_{i+1} + \delta^{s+1}\}$.

Let $z_{r_{i+1}}$ and z_{r_i} be the center-right points of $W_{c_{i+1}}$ and W_{c_i}, respectively. We abuse notation slightly and say $z_{r_i} = z_{r_{i+1}}$ if neither are defined, and $z_{r_i} \neq z_{r_{i+1}}$ if z_{r_i} is defined but not $z_{r_{i+1}}$. Note that if $z_{r_i} \neq z_{r_{i+1}}$ then it must be that $W_{c_{i+1}}$ adds $z_{r_i} = y' = \{c + \delta\}$ (and so $F_i^{-1}(z_{r_i}) \in C_i$ and $F_{i+1}^{-1}(z_{r_i}) \in C_{i+1}$).

Let F_i and F_{i+1} be multi-scale partial maps into W_{c_i} and $W_{c_{i+1}}$, respectively. If neither $\mathsf{LCR}(F_i)$ nor $\mathsf{LCR}(F_{i+1})$ returns infeasible, then we say F_{i+1} ***succeeds*** F_i if the following conditions hold:

1. For any $x \in X$ we have the following properties.
   (a) If $x \in L_i$ then $x \in L_{i+1}$.
   (b) If $x \in C_i$ then $x \in C_{i+1}$ or $x \in L_{i+1}$.
   (c) If $x \in R_i$ then $x \in R_{i+1}$ or $x \in C_{i+1}$.
   (d) If $x \in C_i$ and $F_i(x) \cap B(c_{i+1}, \delta^{\mathsf{ms}(x)+1}) \neq \emptyset$ then $x \in C_{i+1}$ and $F_i(x) = F_{i+1}(x)$.
   (e) If $x \in C_{i+1}$ and $F_{i+1}(x) \cap B(c_i, \delta^{\mathsf{ms}(x)+1}) \neq \emptyset$ then $x \in C_i$ and $F_{i+1}(x) = F_i(x)$.
2. For g_i and g_{i+1} we have the following properties:
   (a) If $z_{r_i} = z_{r_{i+1}}$ then $g_i = g_{i+1}$.
   (b) If $z_{r_i} \neq z_{r_{i+1}}$ then $z_{r_i} \in B(c_{i+1}, \delta)$ and $F_{i+1}^{-1}(z_{r_i}) = F_i^{-1}(z_{r_i}) = g_i^{-1}(z_{r_i})$

We say that the ***status*** of a point $x \in X$ changes if (1) $x \in L_i$ and $x \notin L_{i+1}$, (2) $x \in C_i$ and $x \notin C_{i+1}$, or (3) $x \in R_i$ and $x \notin R_{i+1}$. The above conditions for succession then imply at most one $x \in X$ changes its status between F_i and F_{i+1}. Specifically, as $W_{c_i}, W_{c_{i+1}} \in \mathcal{W}$, there is at most one scale, s, where the set of points in a window can change and at scale s either $W_{c_{i+1}}$ drops $y = \{c_i - \delta^{s+1}\}$ or $W_{c_{i+1}}$ adds $y' = \{c_{i+1} + \delta^{s+1}\}$. Therefore, if $g_i = g_{i+1}$, then only the status of the point which maps onto y or y' can change. If $g_i \neq g_{i+1}$, then $W_{c_{i+1}}$ adds $\{c_{i+1} + \delta\}$, and so only the status of a point in $\mathrm{Dom}(g_{i+1})$ can change.

Remark 4.8. *Let $F_1, \ldots, F_k$ be multi-scale partial maps into $\mathcal{W} = (W_{c_1}, W_{c_2}, \ldots, W_{c_k})$, such that F_{i+1} succeeds F_i. We have the following consequences of the definition of succession,*

1. $\forall i \in \{1, \ldots k-1\}, L_i \subseteq L_{i+1}$ and $R_i \supseteq R_{i+1}$

2. If $x \in L_{i+1}$ and $x \notin L_i$ then $x \in C_i$.

3. If $x \in R_i$ and $x \notin R_{i+1}$ then $x \in C_{i+1}$.

The above conditions imply that if $L_1 = R_k = \emptyset$, then $\forall x \in X, \exists i \in \{1, \dots k\}$ such that $x \in C_i$, and moreover the set $T = \{i \in \{1, \dots, k\} \mid x \in C_i\}$ forms a consecutive subsequence of $(1, 2, \dots, k)$, such that $F_i(x) = F_j(x)$, for all $i, j \in T$.

Lemma 4.9. *Let F_i and F_{i+1} be multi-scale partial maps into $W_{c_i}, W_{c_{i+1}} \in \mathcal{W}$, respectively. If F_i and F_{i+1} are views of some feasible bijection f, then F_{i+1} succeeds F_i.*

Proof: First observe that since F_i and F_{i+1} are views of a single well defined bijection f, that if $x \in C_i, C_{i+1}$ then $F_i(x) = F_{i+1}(x) = f(x)$. Moreover, by the definition of feasible view, if $f(x) \in B(c_i, \delta^{\mathsf{ms}(x)+1})$ then $x \in C_i$ and if $f(x) \in B(c_{i+1}, \delta^{\mathsf{ms}(x)+1})$ then $x \in C_{i+1}$. These facts combined immediately imply that properties 1d and 1e of the definition of succession hold. Similarly, property 2 holds since (a) if $z_{r_i} = z_{r_{i+1}}$ then either z_{r_i} is not defined (in which case g_i and g_{i+1} are empty) or $g_i^{-1}(z_{r_i}) = f^{-1}(z_{r_i}) = f^{-1}(z_{r_{i+1}}) = g_{i+1}^{-1}(z_{r_{i+1}})$, and (b) if $z_{r_i} \neq z_{r_{i+1}}$ then by the definition of center-right points it holds that z_{r_i} is defined and $z_{r_i} \in B(c_{i+1}, \delta)$, and therefore $f^{-1}(z_{r_i}) \in C_i, C_{i+1}$ (by the definition of multi-scale partial maps).

What remains is to prove properties 1a, 1b, and 1c hold. First observe that the functions g_i, g_{i+1} cannot affect the validity of these properties. Specifically, it is easy to argue that if $x \in \mathrm{Dom}(g_i)$ then $x \in C_i, C_{i+1}$, and if $x \in \mathrm{Dom}(g_{i+1})$ then $x \in C_{i+1}$ and either $x \in C_i$ (if $g_i = g_{i+1}$) or $x \in R_i$. Therefore, we only need to consider the set $X' = X \setminus \{\mathrm{Dom}(g_i) \cup \mathrm{Dom}(g_{i+1})\}$. By definition, for $x \in X'$, (i) $x \in L_i$ if $f(x) < c_i - \delta^{\mathsf{ms}(x)+1}$, (ii) $x \in R_i$ if $f(x) > c_i + \delta^{\mathsf{ms}(x)+1}$, and (iii) $x \in C_i$ otherwise. Similarly, (i) $x \in L_{i+1}$ if $f(x) < c_{i+1} - \delta^{\mathsf{ms}(x)+1}$, (ii) $x \in R_{i+1}$ if $f(x) > c_{i+1} + \delta^{\mathsf{ms}(x)+1}$, and (iii) $x \in C_{i+1}$ otherwise. Properties 1a, 1b, and 1c now hold as $c_{i+1} > c_i$. □

4.3 The Algorithm and Analysis

Our algorithm builds a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of all non-empty multi-scale partial maps into the set of windows $\mathcal{W} = \{W_{c_1}, \dots, W_{c_k}\}$, and $(F_i \to F_j) \in \mathcal{E}$ if and only if F_j succeeds F_i. A vertex $F_i \in \mathcal{V}$ is called ***starting*** if it maps into W_{c_1} and $L(F_i) = \emptyset$. Similarly, $F_i \in \mathcal{V}$ is called ***ending*** if it maps into W_{c_k} and $R(F_i) = \emptyset$. The following lemma ensures that a path from a starting vertex to an ending vertex translates to a feasible bijection and vice versa.

Lemma 4.10. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be as described above. There is a non-contracting map of distortion at most δ if and only if there is a directed path from a starting vertex to an ending vertex in $\mathcal{G}$.*

Proof: ($\Rightarrow$) Let f be an map of distortion δ, and let $\mathcal{W} = \{W_{c_1}, \dots, W_{c_k}\}$ be the sequence of multi-scale windows. For each $i \in \{1, 2, \dots, k\}$, let F_i be the view of f in W_{c_i}. Clearly, for all i, F_i is a multi-scale partial map (it is in fact feasible) and hence $F_i \in \mathcal{V}$. Moreover, F_{i+1} succeeds F_i by Lemma 4.9. Finally, observe that by Remark 4.6, since F_1 and F_k are feasible views of f, $L(F_1) = \emptyset$ and $R(F_k) = \emptyset$, and so they are starting and ending vertices, respectively. Therefore, there exists a directed path from a starting vertex to an ending vertex in $\mathcal{G}$.

($\Leftarrow$) Now, suppose there is a directed path $F_1 \to F_2 \to \dots \to F_k$ from a starting vertex F_1 to an ending vertex F_k in $\mathcal{G}$ (note that the definition of succession enforces that F_i is a multi-scale partial map into $W_{c_i} \in \mathcal{W}$). Let $L_i = L(F_i)$, $R_i = R(F_i)$, and $C_i = \mathrm{Dom}(F_i)$. By Remark 4.8 we know (1) for any $x \in X$ there is an $i \in \{1, \dots, k\}$ such that $x \in C_i = \mathrm{Dom}(F_i)$, and (2) the set $\{i \in \{1, \dots, k\} \mid x \in C_i\}$ forms a consecutive subsequence $(i, i+1, \dots, j)$ of $(1, 2, \dots, k)$, such that $F_i(x) = F_{i+1}(x) = \dots = F_j(x)$.

We define the function $f : X \to Y$ as follows: $f(x) = y$ if and only if there is an i such that $F_i(x) = y$. Properties (1) and (2) ensure that f is a well-defined bijection. It remains to prove that f is non-contracting, and has expansion at most δ. To do so we now argue that for each edge $(x_\ell, x_{\ell+1}) \in X$ (resp. $(y_{\ell'}, y_{\ell'+1}) \in Y$), that there exists some F_i such that $x_\ell, x_{\ell+1} \in \mathrm{Dom}(F_i)$ (resp. $y_{\ell'}, y_{\ell'+1} \in \mathrm{Im}(F_i)$). Then since the F_i are non-contracting and have expansion $\leq \delta$, and since f is a bijection consistent with the F_i, Lemma 3.1 implies f is non-contracting and has expansion $\leq \delta$.

So let $(x_\ell, x_{\ell+1})$ be an edge of X. We want to argue that for some F_i, $x_\ell, x_{\ell+1} \in C_i$. As already argued above we know that for x_ℓ (resp. $x_{\ell+1}$) the set of C_i in which x_ℓ (resp. $x_{\ell+1}$) appears, is a non-empty consecutive subsequence of the C_i. For the sake of contradiction assume these two subsequences for x_ℓ and $x_{\ell+1}$ are disjoint. Without loss of generality assume that out of x_ℓ and $x_{\ell+1}$, that x_ℓ is the first to appear in the center. Let F_t be the last of the F_i such that x_ℓ is in the center. It must be that $x_{\ell+1} \in R_t$, since we assumed that x_ℓ and $x_{\ell+1}$ do not appear in the center at the same time, and in order for $x_{\ell+1}$ to be in L_r it must pass through the center (but x_ℓ entered the center first). Therefore $x_\ell \in L_{t+1}$ and $x_{\ell+1} \in R_{t+1}$, as F_t was the last time x_ℓ was in the center and since by the definition of succession only one element in X can change its status in the sets L, C, R. However, in this case the algorithm $\mathsf{LCR}(F_{t+1})$ should have returned invalid (since it specifically checks for this contradiction).

Now let $(y_{\ell'}, y_{\ell'+1})$ be an edge of Y. Consider the multi-scale window $W_{c_i} \in \mathcal{W}$ such that $y_{\ell'}$ is the rightmost point in the window $Y[c_i, \delta]$, in which case $y_{\ell'+1}$ is the point realizing $z_r = \min(\{y \in Y \mid y > c + \delta\})$. Then by the definition of multi-scale partial map $(y_{\ell'}, y_{\ell'+1}) \in \mathrm{Im}(F_i)$. □

Now, we are ready to present the proof for Theorem 1.1, which summarizes our algorithm for finding the minimum distortion bijection between two finite point sets on the line.

Proof (of Theorem 1.1): Lemma 3.2 reduces the problem to $n^{O(1)}$ instances of the following problem: given a real value $\delta \geq 1$, compute a non-contracting (X, Y) bijection with expansion at most δ, or correctly report that no such embedding exists.

To find a non-contracting bijection of distortion δ, we build $\mathcal{G}$ and look for a directed path from a starting vertex to an ending vertex, based on Lemma 4.10. The vertices of $\mathcal{G}$ are multi-scale partial maps on the elements of the standard sequence of multi-scale windows of $\mathcal{W}$, which has length $O(Sn)$. Lemma 4.4 implies there are at most $\delta^{O(\delta^2 S^2)} n$ multi-scale partial maps into a single multi-scale window, and so the size of the graph $\mathcal{G}$ is $\delta^{O(\delta^2 S^2)} n^{O(1)}$. Thus, in the same asymptotic running time, we can check whether there is a path from a starting vertex to an ending vertex.

Since $S = O(\log D)$, the running time of the algorithm is $\delta^{O(\delta^2 \log^2 D)} n^{O(1)}$. □

5 An Approximate Algorithm for Embedding into the Line

Let (X, d_X) be a discrete metric space over $X = \{x_1, x_2, \ldots, x_n\}$. Suppose that for any pair $x, x' \in X$, $1 \leq d_X(x, x') \leq \Delta$, and $\min_{x,x' \in X} d_X(x, x') = 1$. Δ will be called the ***spread*** of X. Also, let $Y = \{y_1, y_2, \ldots, y_m\} \subset \mathbb{R}^+$ be a set of points, where $m \geq n$, and let (Y, d_Y) be the induced metric space.

In this section, we describe an algorithm which computes an embedding of X into Y, whose distortion is an $O(1)$-approximation to the minimum distortion embedding. This algorithm has polynomial running time for polynomial Δ and constant δ. At the end of the section we show this result can be extended to the case when Y is the entire real line.

By Lemma 3.2, it suffices to instead solve the problem of finding a non-contracting embedding of expansion at most δ, where $\delta \geq 1$ is a known value (we will also detect when such an embedding does not exist). Furthermore, we assume that such an embedding is onto y_1 and y_m. This is fine because the first and the last point in the image of the desired embedding can be guessed from $O(m^2)$ choices.

Remark 5.1. *The algorithm in this section can be seen as an extension of the algorithm from the previous section. In particular, much of the terminology is the same (e.g. scales, windows, succession, etc.). However, all the definitions here differ, and are often more complicated. Specifically, rather than defining the terms once in a manner consistent with both problems, here we choose to redefine the terms so that the previous section was as simple as possible.*

5.1 Scales, Windows, and Bins

Scales on X. Let $(x_1, x_2, \ldots, x_n)$ be the Gonzales permutation of X computed as follows. The point x_1 is an arbitrary point in X. For every $2 \leq i \leq n$, the point $x_i \in X \backslash \{x_1, \ldots, x_{i-1}\}$ is the farthest point from the set $\{x_1, \ldots, x_{i-1}\}$. For each $s \in \{0, 1, \ldots, S = \lfloor \log_\delta \Delta \rfloor + 1\}$, the set $\boldsymbol{X_{\geq s}}$ is composed of the points in the maximal subsequence of $(x_1, x_2, \ldots, x_n)$, in which the mutual distances of the points are at least δ^s. Note that $X_{\geq S} = \emptyset$, as $\Delta < \delta^S$. The scale of a point $x \in X$, is the largest s such that $x \in X_{\geq s}$.

For the remainder of the paper we assume that we have precomputed the sets $X_s = X_{\geq s} \backslash X_{\geq s+1}$, for all $s \in \{0, 1, \ldots, S-1\}$. Additionally, we also precompute for each $x \in X$ and for each $0 \leq s < S$, the x's nearest neighbor in $X_{\geq s}$. This can all be done in $(Sn)^{O(1)}$ time.

Windows. Let $B(c, r) = [c - r, c + r]$ be the ball of radius r centered at c, and let $Y[c, r] = B(c, r) \cap Y$. We refer to the set $Y[c, 3\delta^{s+2}]$ as a ***window*** of scale s with center c. In turn, we refer to the collection of concentric windows of scales 0 to S, $W_c = \{Y[c, 3\delta^2], Y[c, 3\delta^3], \ldots, Y[c, 3\delta^{S+2}]\}$, as a ***multi-scale window*** with center c.

The multi-scale window $W_{c'}$ ***succeeds*** the multi-scale window W_c if there exists an $0 \leq s \leq S$ with the following properties.

1. For all $0 \leq i \leq S$, $i \neq s$, we have $Y[c, 3\delta^{i+2}] = Y[c', 3\delta^{i+2}]$.

2. Either $Y[c, 3\delta^{s+2}] = Y[c', 3\delta^{s+2}] \cup \{c - 3\delta^{s+2}\}$ or $Y[c', 3\delta^{s+2}] = Y[c, 3\delta^{s+2}] \cup \{c' + 3\delta^{s+2}\}$.

Intuitively, the set of points at all scales except s are identical. At scale s, they differ in one point, which is either $c - 3\delta^{s+2}$ or $c' + 3\delta^{s+2}$. Sometimes we refer to this point as the ***difference*** of W_c and $W_{c'}$.

It is possible to find a sequence of multi-scale windows $\mathcal{W} = (W_{c_1}, W_{c_2}, \ldots, W_{c_k})$ with the following properties.

1. For each $i \in \{2, 3, \ldots, k\}$, W_{c_i} succeeds $W_{c_{i-1}}$.

2. $c_1 = y_1$ and $c_k = y_m$.

Specifically, let $\mathcal{C}' = \{y_i - 3\delta^{s+2}$ and $y_i + 3\delta^{s+2} \mid i \in \{1, \ldots, n\}$ and $s \in \{0, \ldots, S\}\}$. Then the set of centers of the multi-scale windows is the set $\mathcal{C} = (\mathcal{C}' \cup \{y_1, y_n\}) \cap [y_1, y_n]$.[4] Since in each succession a point either enters or leaves the window of some scale, we have $k = |\mathcal{C}| = O(nS)$. We fix such a sequence $\mathcal{W}$ for the rest of this paper, and refer to it as the ***standard sequence of multi-scale windows***.

[4]Technically, depending on the distances between adjacent points in Y, more than one window may change between W_{c_i} and $W_{c_{i+1}}$. This is easily resolved by slightly increasing the width of each window by some small ϵ.

Bins. For each $i \in \{0, 1, \ldots, S\}$, we partition $\mathbb{R}^+$ into intervals of length δ^s: $\sqcup_s = \{[0, \delta^s), [\delta^s, 2\delta^s), \ldots\}$. In other words, $\sqcup_s$ is the one dimensional grid on $\mathbb{R}^+$, of cell width δ^s. We refer to these intervals as ***bins*** of scale s. In turn, we define $\sqcup_s[c, r]$ as the subset of intervals in $\sqcup_s$ that intersect $B(c, r)$. Some of these bins are contained in $B(c, r)$, but some of them may only partially intersect $B(c, r)$.

Remark 5.2. *In the remainder of this section we will assume that δ is a power of 2. Note that this can be done at the cost of a factor of 2 in the final approximation since we know $\delta \geq 1$. This assumption is not necessary, but it will make the presentation simpler as bins from different scales will now be concentric.*

In order to obtain a faster algorithm, we consider approximate maps into bins rather than into the actual points in Y. The following lemma, is immediate from the definitions above.

Lemma 5.3. *Let $f_s : X_{\geq s} \to \sqcup_s[c, 3\delta^{s+2}]$ be a one-to-one partial map, for $s \in \{0, 1, \ldots, S\}$, and $c \in \mathbb{R}$. We have, $|\text{Dom}(f_s)| = |\text{Im}(f_s)| = O(\delta^2)$.*

Proof: Each bin in $\sqcup_s$ has length δ^s, so there are $O(\delta^2)$ bins intersecting $B[c, 3\delta^{s+2}]$. Thus, $|\text{Im}(f_s)| = O(\delta^2)$. In addition, $|\text{Dom}(f_s)| = |\text{Im}(f_s)|$ since f_s is one-to-one. ☐

5.2 Approximate Maps, Restrictions, and Extension

Approximate maps. Consider a collection of maps $Z = \{f_0, f_1, \ldots, f_S, g\}$ with the following properties:

1. For each $1 \leq s \leq S$, the map $f_s : X_{\geq s} \to \sqcup_s[c, 3\delta^{s+2}]$ is one-to-one.
2. The map $f_0 : X \to Y[c, 3\delta^2]$ is one-to-one[5].
3. For any $x \in X$ and $0 \leq s_1 < s_2 \leq S$, if $x \in \text{Dom}(f_{s_1})$ then $x \in \text{Dom}(f_{s_2})$ and $f_{s_1}(x) \subseteq f_{s_2}(x)$.
4. If $[i\delta^s, (i+1)\delta^s) \in \text{Im}(f_s)$ then $[i\delta^s, (i+1)\delta^s) \cap Y \neq \emptyset$.
5. The map $g : X \to Y \cap (c + 3\delta^2, \infty)$ acts on exactly one point in X if $y_m \in (c + 3\delta^2, \infty)$, and it is empty otherwise.
6. If $y_1 \in [c, 3\delta^2]$ then f_0 maps onto y_1, and if $y_m \in [c, 3\delta^2]$ then f_0 maps onto y_m.

We define the ***approximate map***, F, induced by Z to be the function such that:

1. $\text{Dom}(F) = \cup_{h \in Z} \text{Dom}(h)$
2. $F(x) = \cap_{h \in Z,\ x \in \text{Dom}(h)} h(x)$

In particular, for any $x \notin \text{Dom}(g)$, if $F(x)$ is defined then it is just the smallest bin containing x (which may be a single point). Also note that if $x \in \text{Dom}(g)$, then it may be that $F(x) = \emptyset$.

We will use the notation $F = \langle f_0, f_1, \ldots, f_S, g \rangle$ to refer to the approximate map induced by a set $\{f_0, f_1, \ldots, f_S, g\}$ (satisfying the above properties). We say that such an F as defined above is an approximate map into c (or rather into the multi-scale window with center c).

Remark 5.4. *Note that the last two conditions on g and f_0, respectively, imply that for any $W_c \in \mathcal{W}$, if F is an approximate map into W_c, then $\text{Dom}(F) \neq \emptyset$.*

[5] Note that f_0 can be defined into $\sqcup_s[c, 3\delta^2]$ as well, however, explicitly defining it as a map into the points makes the exposition of the ideas simpler.

Restriction and Extension. Let f be an embedding of X into Y, and let $f_s : X_{\geq s} \to \sqcup_s[c, 3\delta^{s+2}]$ be a scale s map. Also, let $\amalg_s[c, 3\delta^{s+2}]$ be the union of all bins in $\sqcup_s[c, 3\delta^{s+2}]$ (i.e. $\amalg_s[c, 3\delta^{s+2}]$ is an interval on the real line). Note that this might properly contain $B(c, 3\delta^{s+2})$ because the extreme bins might partially intersect $B(c, 3\delta^{s+2})$. The partial function f_s is an ***approximate restriction*** of f into $X_{\geq s} \times \sqcup_s[c, 3\delta^{s+2}]$ if the following properties hold:

1. For any $x \in X_{\geq s}$ and $f(x) \in \amalg_s[c, 3\delta^{s+2}]$, we have $f(x) \in f_s(x)$.

2. For all other $x \in X$, $x \notin \mathrm{Dom}(f_s)$.

Under these conditions, we also say that f is an extension of f_s.

Let $F = \langle f_0, f_1, \ldots, f_S, g\rangle$. The approximate map F is the ***approximate restriction*** of the embedding f into W_c if the following properties hold.

1. For each $s \in \{0, \ldots, S\}$, f_s is the approximate restriction of f into $X_{\geq s} \times \sqcup_s[c, 3\delta^{s+2}]$.

2. For $x \in \mathrm{Dom}(g)$, we have $g(x) = f(x)$.

Under these conditions, we also say that f is an ***extension*** of F.

An approximate map $F = \langle f_0, f_1, \ldots, f_S, g\rangle$ is ***feasible*** if it has a non-contracting extension, f, of expansion at most δ.[6] In this case f is a feasible extension of F.

5.3 Left, Right, and Center

Left, right. For any approximate map, $F = \langle f_0, \ldots, f_S, g\rangle$, with a feasible extension f, let $L_{s,f} = L_{s,f}(F) = \{x \in X_{\geq s} | x \notin \mathrm{Dom}(f_s) \text{ and } f(x) < c - 3\delta^{s+2}\}$, $C_{s,f} = C_{s,f}(F) = \mathrm{Dom}(f_s)$, and $R_{s,f} = R_{s,f}(F) = \{x \in X_{\geq s} | x \notin \mathrm{Dom}(f_s) \text{ and } f(x) > c + 3\delta^{s+2}\}$. We have the following.

Lemma 5.5. *Let F be a feasible approximate map into $W_c \in \mathcal{W}$. Let f and f' be two distinct feasible extensions of F, then:*

1. *The sets $L_{s,f}, C_{s,f}, R_{s,f}$, form a tri-partition of $X_{\geq s}$.*

2. *$L_{s,f} = L_{s,f'}$ and $R_{s,f} = R_{s,f'}$.*

Proof: Since F is a restriction of f, if $f(x) \in B(c, 3\delta^{s+2})$ then $x \in \mathrm{Dom}(f_s) = C_{f,s}$. Otherwise, if $x \notin \mathrm{Dom}(f_s)$, then if $x < c - 3\delta^{s+2}$ then $x \in L_{f,s}$ and if $x > c + 3\delta^{s+2}$ then $x \in R_{f,s}$.

To prove the second part, we show that given F the condition of any point in any feasible extension is uniquely determined. We present the following algorithmic proof for this statement.

We use induction on s, the scale of x, starting from $s = S - 1$ and going to $s = 0$ (recall that $X_{\geq S} = \emptyset$).

Since there is an embedding from X into Y of expansion at most δ, it follows that $d_Y(y_1, y_m) \leq \delta^{S+1}$ (recall that we only consider embeddings which are required to map onto both y_1 and y_m). Hence, for any $W_c \in \mathcal{W}$, since $c \in [y_1, y_m]$, it follows that the scale $S - 1$ window of any standard multi-scale window contains all points. Consequently, this window intersects all non-empty bins of scale $S - 1$. Therefore, for all $x \in X_{\geq S-1}$, we have $x \in \mathrm{Dom}(f_{S-1})$ (as F is feasible), and so the base case for induction holds.

Now, assume that the condition of any point in $X_{\geq s+1}$ is uniquely determined. Let $x \in X_{\geq s} \backslash X_{\geq s+1}$, and suppose $x \notin \mathrm{Dom}(f_s)$, then we show that simultaneously for all feasible extensions of F, either the image of x lies strictly to the left of c or strictly to the right of c. Since

[6] Conversely, it is easy to verify that the restriction of a non-contracting and expansion at most δ embedding of X into Y satisfies all the requirements defining an approximate map.

$x \notin \mathrm{Dom}(f_s)$, this will imply that simultaneously for all feasible extensions of F, either x is on the left or right side of f_s.

Let x' be x's nearest neighbor in $X_{\geq s+1}$. By the definition of $X_{\geq s+1}$, we have $d_X(x, x') \leq \delta^{s+1}$ (as otherwise $x \in X_{\geq s+1}$), and so $d_Y(f(x), f(x')) \leq \delta^{s+2}$, for any feasible extension f. By the induction hypothesis, exactly one of the following three conditions hold for x'.

1. x' is on the left side of f_{s+1}: Therefore, for any feasible extension, $f(x') < c - 3\delta^{s+3}$.

$$\Rightarrow f(x) < c - 3\delta^{s+3} + d_Y(f(x), f(x')) \leq c - 3\delta^{s+3} + \delta^{s+2} \leq c.$$

2. x' is on the right side of f_{s+1}: Therefore, for any feasible extension, $f(x') > c + 3\delta^{s+3}$.

$$\Rightarrow f(x) > c + 3\delta^{s+3} - d_Y(f(x), f(x')) \geq c + 3\delta^{s+3} - \delta^{s+2} \geq c.$$

3. x' is in $\mathrm{Dom}(f_{s+1})$. We consider three cases:

   (a) The interval $f_{s+1}(x')$ is completely on the left side of c: Therefore, for any feasible extension, $f(x') < c$.

   $$\Rightarrow f(x) < f(x') + d_Y(f(x), f(x')) \leq c + \delta^{s+2}$$

   Since $x \notin \mathrm{Dom}(f_s)$, we have $f(x) \notin B(c, 3\delta^{s+2})$, so $f(x) < c$.

   (b) The interval $f_{s+1}(x')$ is completely on the right side of c: Therefore, for any feasible extension, $f(x') > c$.

   $$\Rightarrow f(x) > f(x') - d_Y(f(x), f(x')) \geq c - \delta^{s+2}$$

   Since $x \notin \mathrm{Dom}(f_s)$, we have $f(x) \notin B(c, 3\delta^{s+2})$, so $f(x) > c$.

   (c) The interval $f_{s+1}(x')$ contains the point c: Therefore, for any feasible extension, $f(x') \in [c - \delta^{s+1}, c + \delta^{s+1}]$, as $f_{s+1}(x')$ is an interval of length δ^{s+1}. Therefore,

   $$f(x) \in [c - \delta^{s+1} - \delta^{s+2}, c + \delta^{s+1} + \delta^{s+2}] \subseteq [c - 2\delta^{s+2}, c + 2\delta^{s+2}].$$

   So, x must be in the domain of f_s, as f_s is a restriction of f into $X_{\geq s} \times \sqcup_s [c, 3\delta^{s+2}]$. Hence, this case cannot happen.

□

The above lemma implies that all feasible extensions of F induce the same tri-partition of each $X_{\geq s}$. Specifically, $X_{\geq s} = L_s \cup C_s \cup R_s$, which we call the ***left***, ***center***, and ***right*** of F at scale s, respectively. The center is the domain of f_s, i.e. $C_s = \mathrm{Dom}(f_s)$, and the left set, L_s, and right set, R_s, are as defined in the above lemma.

This lemma also gives an algorithm which given a feasible F, outputs the tri-partitions $L_s \cup C_s \cup R_s$ at all scales $0 \leq s \leq S$. If F is not feasible, then ideally this could be detected, however being able to do so without knowing the extension f seems unlikely. Therefore, if F is not feasible the algorithm either returns that F is infeasible if a tri-partition is not produced or outputs some bogus tri-partition. The following corollary formalizes this statement.

Corollary 5.6. *Given an approximate map F into $W_c \in \mathcal{W}$, there is a polynomial time algorithm such that if F is feasible it outputs the corresponding partition of $X_{\geq s}$ into sets $L_s \cup C_s \cup R_s$ for all $0 \leq s \leq S$, as described above. If F is not feasible it either outputs a tri-partition $L_s \cup C_s \cup R_s$ or returns that F is infeasible.*

Remark 5.7. *Note that in the previous section our algorithm for determining left, right, and center (i.e. Lemma 4.9), also performed a consistency check to make sure adjacent points did not get assigned to the left and right. Here the check is more involved (as left and right are with respect to individual scales), and so instead we perform a second post-processing step. This is the "plausible" check as described below.*

Plausible approximate maps. An approximate map F assigns an interval to each point in its domain. These intervals, intuitively indicate estimations of the image of the points in an optimal solution. Lemma 5.5 extends these estimations to the points that are not in the domain of F. It assigns intervals (of infinite lengths) to the points not in the domain of F. Therefore, F together with Lemma 5.5 gives an estimation for the image of all points in X in an optimal solution. Corollary 5.6 ensures that these estimations are computable in polynomial time. The following definition, formalizes this idea by introducing a function $\widetilde{F} : X \to \mathcal{I}$, where $\mathcal{I}$ is the set of all intervals on the real line. (We emphasize that $\emptyset \in \mathcal{I}$.) For each $x \in X$:

1. If $x \in \mathrm{Dom}(F)$ then $\widetilde{F}(x) = F(x)$.
2. If $x \notin \mathrm{Dom}(F)$ then

$$\widetilde{F}(x) = \left(\bigcap_{x \in L_s} (-\infty, c - 3\delta^{s+2}) \right) \cap \left(\bigcap_{x \in R_s} (c + 3\delta^{s+2}, \infty) \right).$$

Intuitively, F is plausible if one cannot conclude that it is not feasible by examining $\widetilde{F}$. Formally, F is ***plausible*** if the following properties hold.

1. For all $x \in X$, $\widetilde{F}(x) \neq \emptyset$.
2. For all $x, x' \in X$, there are $y \in \widetilde{F}(x)$ and $y' \in \widetilde{F}(x')$ such that:

$$1 \leq d_Y(y, y')/d_X(x, x') \leq \delta.$$

In particular, if F is feasible then it is plausible, but not necessarily the other way around. The following lemma ensures that plausibility can be checked in polynomial time.

Lemma 5.8. *Let $F = \langle f_0, f_1, \ldots, f_S, g \rangle$ be an approximate map into W_c. It can be decided in polynomial time whether F is plausible or not.*

Proof: It takes $O(Sn)$ time to compute $F(x)$ for all $x \in X$, given the functions $f_0, f_1, \ldots, f_S, g$. By Corollary 5.6, $\widetilde{F}$ can be computed in polynomial time. Finally, given $\widetilde{F}$, plausibility conditions can be checked in $O(n^2)$ time. □

Next, we bound the number of plausible approximate maps into a fixed multi-scale window. To that end, we need the following lemma that bounds the density of points in a metric space, provided it admits an embedding of small distortion into the Euclidean line.

Lemma 5.9. *Let (X, d_X) be any metric space that can be embedded into the line with distortion δ, and let $R, r \in \mathbb{R}^+$. Let $K \subseteq X$ be a subset of the points with the property that for any $x, x' \in K$, we have $d_X(x, x') > r$. Finally, let $B = B(c, R)$ be any ball with center $c \in X$ and radius R. B contains at most $2\delta R/r$ points of K.*

Proof: Let f be any embedding of contraction 1 and expansion δ. Such an embedding exists because distortion is invariant under scaling of X. Let I be the shortest interval on the line that contains $\mathrm{Im}(f)$. The length of I is at most $2\delta R$, because the expansion of f is δ.

In addition, because f is non-contracting, for any $x, x' \in K$, we have $|f(x) - f(x')| > r$. Furthermore, f maps all points of K into I. Thus, by a packing argument, $|K| \leq 2\delta R/r$. □

Lemma 5.10. *For any $c \in \mathbb{R}^+$, there are are $\delta^{O(S\delta^2)} \cdot (mn)^{O(1)}$ plausible approximate maps into W_c. Moreover, there is an algorithm to enumerate these approximate maps in the same asymptotic running time.*

Proof: Let $F = \langle f_0, f_1, \ldots, f_S, g\rangle$ be a plausible approximate map into W_c. Let a be the smallest integer in $\{0, 1, \ldots, S\}$, for which f_a is not empty, and let $x \in \mathrm{Dom}(f_a)$. (Note there are at most $O(mn)$ ways to define g, and hence this bound the number of possible F where a is not defined.)

First, for each $s \in \{a, \ldots, S\}$, we show that given x and $f_a(x)$ are fixed, then there is a small set, T_s, which must contain all points in $\mathrm{Dom}(f_s)$. Specifically, T_s will be all points from $X_{\geq s}$ which are in the ball of radius $8\delta^{s+2}$ centered at x. By Lemma 5.9, since the distance between any pair of points in $X_{\geq s}$ is $\Omega(\delta^s)$, we know $|T_s| = O(\delta^3)$. Therefore, if we can show $\mathrm{Dom}(f_s) \subseteq T_s$ (given that x is already fixed), then we can limit the possible choices for the points in $\mathrm{Dom}(f_s)$.

Specifically, we show $d_X(x, x') \leq 8\delta^{s+2}$, for any $x' \in \mathrm{Dom}(f_s)$. By definition, $f_a(x)$ and $f_s(x')$ are intervals of length δ^a and δ^s, respectively, and both $f_a(x)$ and $f_s(x')$ intersect $B(c, 3\delta^{s+2})$. In addition, $F(x) \subseteq f_a(x)$ and $F(x') \subseteq f_s(x')$. Thus, for any $y \in F(x)$ and $y' \in F(x')$ we have:

$$d_Y(y, y') \leq 6\delta^{s+2} + \delta^a + \delta^s \leq 8\delta^{s+2}. \tag{2}$$

Because F is plausible, there are $z \in F(x)$ and $z' \in F(x')$ such that $d_X(x, x') \leq d_Y(z, z')$. Also, as $z \in F(x)$ and $z' \in F(x')$, we have $d_Y(z, z') \leq 8\delta^{s+2}$. Therefore, we have the desired bound:

$$d_X(x, x') \leq d_Y(z, z') \leq 8\delta^{s+2}.$$

By Lemma 5.3, $|\mathrm{Dom}(f_s)| = |\mathrm{Im}(f_s)| = O(\delta^2)$. Therefore, the total number of candidates for f_s can be bounded as follows, given that x and $f_a(x)$ are fixed.

$$\binom{|T_s|}{|\mathrm{Dom}(f_s)|} \cdot |\mathrm{Dom}(f_s)|^{|\mathrm{Im}(f_s)|} = \binom{O(\delta^3)}{O(\delta^2)} \cdot O(\delta^2)^{O(\delta^2)} = \delta^{O(\delta^2)}.$$

Our algorithm guesses a, $f_a(x)$, and x, in this order. There are, S options for a, $O(\delta^2)$ options for $f_a(x)$, and $O(n)$ options for x. Once x, a, and $f_a(x)$ are fixed, there are $\delta^{O(\delta^2)}$ possibilities for each f_s. So, the total number of candidates for $\{f_0, \ldots, f_S\}$ considered by our algorithm is

$$S \cdot O(\delta^2) \cdot n \cdot \left(\delta^{O(\delta^2)}\right)^S = \delta^{O(S\delta^2)} n.$$

Finally, our algorithm guesses g from $mn + 1$ options. It either maps exactly one point in X to exactly one point in Y, or it is empty. Therefore, overall the total number of candidates for plausible approximate maps into W_c, enumerated by our algorithm, is $\delta^{O(S\delta^2)} mn^2$. □

Succession of maps. Let $W_{c_{i+1}}$ succeed W_{c_i} in the standard sequence $\mathcal{W}$. By definition, there is exactly one scale s such that $Y[c_i, 3\delta^{s+2}]$ and $Y[c_{i+1}, 3\delta^{s+2}]$ differ. Specifically, either $Y[c_i, \delta^{s+1}] = Y[c_{i+1}, 3\delta^{s+2}] \cup \{c_i - 3\delta^{s+2}\}$ or $Y[c_{i+1}, 3\delta^{s+2}] = Y[c_i, 3\delta^{s+2}] \cup \{c_{i+1} + 3\delta^{s+2}\}$. In the former case we say $W_{c_{i+1}}$ ***drops*** the point $y = \{c_i - 3\delta^{s+2}\}$ at scale s, and in the latter case we say $W_{c_{i+1}}$ ***adds*** the point $y' = \{c_{i+1} + 3\delta^{s+2}\}$ at scale s.

Let $F_i = \langle f_0^i, f_1^i, \ldots, f_S^i, g^i\rangle$ and $F_{i+1} = \langle f_0^{i+1}, f_1^{i+1}, \ldots, f_S^{i+1}, g^{i+1}\rangle$ be plausible approximate maps into W_{c_i} and $W_{c_{i+1}}$, respectively. We say F_{i+1} ***succeeds*** F_i if the following conditions hold.

1. For any $s \in \{0, \dots, S\}$, and for any $x \in X_{\geq s}$ we have the following properties.
   (a) If $x \in L_s^i$ then $x \in L_s^{i+1}$.
   (b) If $x \in C_s^i$ then $x \in C_s^{i+1}$ or $x \in L_s^{i+1}$.
   (c) If $x \in R_s^i$ then $x \in R_s^{i+1}$ or $x \in C_s^{i+1}$.
   (d) If $x \in C_s^i$ and $f_s^i(x) \cap B(c_{i+1}, 3\delta^{s+2}) \neq \emptyset$ then $x \in C_s^{i+1}$ and $f_s^i(x) = f_s^{i+1}(x)$.
   (e) If $x \in C_s^{i+1}$ and $f_s^{i+1}(x) \cap B(c_i, 3\delta^{s+2}) \neq \emptyset$ then $x \in C_s^i$ and $f_s^i(x) = f_s^{i+1}(x)$.
2. For g^i and g^{i+1} we have the following properties:
   (a) If $W_{c_{i+1}}$ does *not* add a point at scale zero then $g^i = g^{i+1}$.
   (b) If $W_{c_{i+1}}$ adds y at scale zero then
      i. If $y \notin \mathrm{Im}(g^i)$ then $y \notin \mathrm{Im}(f_0^{i+1})$ and $g^i = g^{i+1}$.
      ii. If $y \in \mathrm{Im}(g^i)$ and $g^i(x) = y$ then $f_0^{i+1}(x) = y$.

Remark 5.11. *Let F_i, F_{i+1} be as defined above. Then one can show that if F_i and F_{i+1} are restrictions of some feasible f, then F_{i+1} succeeds F_i. Indeed, the definition of succession was chosen to have the strictest rules possible which are consistent with feasible maps (i.e. we want to weed out non-feasible maps). This can be formally proven, as was done in Lemma 4.9, but we spare the reader (and the writer) the tedious details.*

5.4 The Algorithm and Analysis

Our algorithm builds a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of all plausible approximate maps into the set of windows $\mathcal{W} = \{W_{c_1}, \dots, W_{c_k}\}$, and $(F_i \to F_j) \in \mathcal{E}$ if and only if F_j succeeds F_i. A vertex $F_i \in \mathcal{V}$ is called ***starting*** if it maps into W_{c_1} and $L(F_i) = \emptyset$. Similarly, $F_i \in \mathcal{V}$ is called ***ending*** if it maps into W_{c_k} and $R(F_i) = \emptyset$.

Recall that $c_1 = y_1$ and $c_k = y_m$, and also that the last property of approximate maps requires that an embedding hits y_1 and y_m (i.e. this holds for any starting and ending vertex, respectively). This can be ensured by guessing the first and the last point of the image of an optimal embedding among a polynomial number of choices.

Our algorithm computes a directed path in $\mathcal{G}$ from any starting vertex to any ending vertex. The following lemma ensures that such a path exists if X admits a feasible embedding into Y.

Lemma 5.12. *If there is a non-contracting map of distortion at most δ then there is a directed path from a starting vertex to an ending vertex in $\mathcal{G}$.*

Proof: Let f be a feasible map, and let $\mathcal{W} = \{W_{c_1}, \dots, W_{c_k}\}$ be the standard sequence of multi-scale windows. For each $i \in \{1, 2, \dots, k\}$, let F_i be the restriction of f into W_{c_i}. For each $1 \leq i \leq k$, F_i is plausible (in fact, feasible) since it is a restriction of the feasible map f, and hence $F_i \in \mathcal{V}$. Moreover, F_{i+1} succeeds F_i by Remark 5.11. Finally, observe that since F_1 and F_k are restrictions of a feasible f, $L(F_1) = \emptyset$ and $R(F_k) = \emptyset$, and so they are starting and ending vertices, respectively. Therefore, there exists a directed path from a starting vertex to an ending vertex in $\mathcal{G}$. ☐

On the other hand, we now want to show that given a directed path $\mathcal{P} = F_1 \to F_2 \to \dots \to F_k$ from a starting vertex F_1 to an ending vertex F_k, that an embedding of X into Y with distortion $O(\delta)$ can be computed. The following lemma describes some properties of the approximate maps in $\mathcal{P}$ that will be helpful towards this end.

Lemma 5.13. *Let $\mathcal{P} = F_1 \to F_2 \to \ldots \to F_k$ be a directed path from a starting vertex F_1 to an ending vertex F_k in $\mathcal{G}$, where $F_i = \langle f_0^i, \ldots, f_S^i, g^i \rangle$, for all $1 \le i \le k$. For each $1 \le i \le k$ and for each $0 \le s \le S$, let $L_s^i = L_s(F_i)$, $R_s = R_s^i(F_i)$, and $C_s^i = \mathrm{Dom}(f_s^i)$. The following properties hold.*

1. *$\emptyset = L_s^1 \subseteq L_s^2 \subseteq \ldots \subseteq L_s^k$.*
2. *$R_s^1 \supseteq R_s^2 \supseteq \ldots \supseteq R_s^k = \emptyset$.*
3. *For any $x \in X_{\ge s}$ there is a non-empty contiguous subsequence $(\ell, \ell+1, \ldots, r)$ of $(1, 2, \ldots, k)$ such that for all $i \in \{\ell, \ell+1, \ldots, r\}$ we have $x \in C_s^i = \mathrm{Dom}(f_s^i)$. Moreover, for all $i, j \in \{\ell, \ell+1, \ldots, r\}$, $f_s^i(x) = f_s^j(x)$.*
4. *For any $x \in X_{\ge s}$, either $\ell = 1$ or $f_s^\ell(x)$ is the rightmost bin at scale s intersection $B(c_\ell, 3\delta^{s+2})$. Similarly, either $r = k$ or $f_s^r(x)$ is the leftmost bin at scale s intersecting $B(c_r, 3\delta^{s+2})$.*

Proof: The fact that $L_s^1 = \emptyset$ and $R_s^k = \emptyset$ is immediate from the definition of starting and ending vertices. The remaining statements of the lemma are then all direct consequences of the definition of succession. Specifically, the containment properties, i.e. $L_s^1 \subseteq L_s^2 \subseteq \ldots \subseteq L_s^k$ and $R_s^1 \supseteq R_s^2 \supseteq \ldots \supseteq R_s^k$ are implied by the first three properties of succession (i.e. 1a, 1b, and 1c). Statement 3 above then follows as 1c of the definition of succession does not allow any $x \in X$ to go directly from the right to the left (at any scale), and moreover, each $x \in X$ must be in the center at some point as $L_s^1 = \emptyset$ and $R_s^k = \emptyset$ for all s. The last statement then follows by properties 1d and 1e from the definition of succession. □

Given $\mathcal{P}$, we define the relation h from X to Y as follows: we say $h(x) = y$ if and only if there is an $1 \le i \le k$ such that $f_0^i(x) = y$.

Lemma 5.14. *The relation h defined above is an embedding of X into Y.*

Proof: First, we show that h is a function. Note that (x, y) pairs are added to h only by looking at the smallest scale functions f_0^i. The three properties of Lemma 5.13 guarantee that the set $\{F_i | 1 \le i \le k,\ x \in C_0^i = \mathrm{Dom}(f_0^i)\}$ is a non-empty contiguous subpath $F_i \to \ldots \to F_j$ of $\mathcal{P}$. So, by the definition of succession, $f_0^i(x) = \ldots = f_0^j(x)$, so x is mapped to at most one point in Y. Since, the subpath is not empty, h acts on all members of X. Therefore, h is a function.

To show that h is injective, we prove that the pre-image of each $y \in Y$, if it exists, is unique. To that end, let $x, x' \in X$, and suppose $h(x) = h(x') = y \in Y$. Therefore, there exist i and j such that $f_0^i(x) = y$ and $f_0^j(x') = y$, and without loss of generality assume $i < j$. It follows, from the definition of the standard sequence of multi-scale windows, that y is in the smallest scale window of $\{W_{c_i}, W_{c_{i+1}}, \ldots, W_{c_j}\}$. Therefore, because of the definition of succession $f_0^i(x) = f_0^{i+1}(x) = \ldots = f_0^j(x)$. Finally, $f_0^j(x) = y$ and $f_0^j(x') = y$ implies $x = x'$, as approximate maps are injective. □

Now that we have ensured that h is an embedding of X into Y, what remains is to bound the distortion of h. To this end, images of x (that are intervals on the real line) under different F_i's are considered as estimates for $h(x)$. The following lemma ensures that these estimates are consistent with $h(x)$.

Lemma 5.15. *Let $x \in X$, and let $F_i = \langle f_0^i, f_1^i, \ldots, f_S^i, g^i \rangle \in \mathcal{P}$. For any $0 \le s \le S$, if $f_s^i(x)$ is defined then $h(x) \in f_s^i(x)$. Moreover, if $x \in L_s^i$ then $h(x) < c_i - 3\delta^{s+2}$, and if $x \in R_s^i$ then $h(x) > c_i + 3\delta^{s+2}$.*

Proof: Let F_j be any approximate map in $\mathcal{P}$ such that $x \in \mathrm{Dom}(f_0^j)$. Property 3 of approximate maps implies that for all $0 \le s \le S$, $x \in \mathrm{Dom}(f_s^j)$, and $h(x) = f_0^j(x) \in f_s^j(x)$. Now, let F_i be any approximate map of $\mathcal{P}$ and suppose that $x \in \mathrm{Dom}(f_s^i)$ for an $0 \le s \le S$. Lemma 5.13(3) implies that $f_s^i(x) = f_s^j(x)$, and so it contains $h(x)$.

Next, we show that $x \in L_s^i$ implies $h(x) < c_i - 3\delta^s$. Choose j so that $x \in C_s^j$ and $x \in L_s^{j+1}$. Such a j exists by Lemma 5.13(3). By Lemma 5.13(4) $f_s^j(x)$ is the leftmost bin that intersects $B(c_j, 3\delta^{s+2})$. Therefore, $f_s^{j+1}(x)$ is completely on the left side of $c_{j+1} - 3\delta^s$. Moreover, for any $i > j$, it is completely on the left side of $c_i - 3\delta^s$. By the first part of this lemma, $h(x) \in f_s^j(x)$, therefore, $h(x)$ is on the left side of $c_i - 3\delta^s$, for any $i > j$. By Lemma 5.13(1), $x \in L_s^i$ implies $i > j$. Thus, $x \in L_s^i$ implies $h(x) < c_i - 3\delta^s$. The proof for $x \in R_s^i$ is symmetric. □

In the following couple of lemmas we bound the expansion of h. We start by bounding the expansion of pairs of points in X that are close to each other relative to their scales.

Lemma 5.16. *Let $x, x' \in X$ be of scales s and s' respectively. If $d_X(x, x') \le 3\delta^{\max(s,s')+1}$ then*

$$d_Y(h(x), h(x')) = O(\delta) \cdot d_X(x, x')$$

Proof: Let $h(x) < h(x')$. First consider the case that $s' = \max(s, s')$. Let F_j (into W_{c_j}) be the last approximate map of $\mathcal{P}$ such that $h(x) \in B(c_j, 3\delta^2)$. The properties of the standard sequence imply $F_j(x) = h(x) \le c_j$. Consequently, for any $y' > c_j + 3\delta^{s'+2}$, we have $d_Y(y', F_j(x)) > 3\delta^{s'+2}$. Thus, plausibility of F_j implies $x' \notin R_{s'}^j$ (as we assumed $d_X(x, x') \le 3\delta^{s'+1}$). Additionally, since $h(x') > h(x) \ge c_j - 3\delta^{s'+2}$, Lemma 5.15 implies $x' \notin L_{s'}^j$. Therefore, $x' \in C_{s'}^j = \mathrm{Dom}(f_{s'}^j)$.

Let a be smallest scale such that $x' \in C_a^j$. If $a = 0$ then $x' \in \mathrm{Dom}(f_0^j)$ and its image has to be $h(x')$ (again by Lemma 5.15), and the lemma statement immediately follows from the definition of plausibility. So we assume that $0 < a \le s'$, in the rest of the proof.

The plausibility of F_j implies the existence of a $z' \in f_a^j(x')$ such that

$$d_Y(z', h(x)) \le \delta d_X(x', x).$$

Since, $f_a^j(x')$ is an interval of length δ^a that contains $h(x')$, this in turn implies

$$d_Y(h(x'), h(x)) \le \delta d_X(x', x) + \delta^a. \tag{3}$$

Because $h(x') > h(x) \ge c_j - 3\delta^{a+1}$, Lemma 5.15 implies $x' \notin L_{a-1}^j$. Also, $x' \notin C_{a-1}^j$ because of the choice of a. Thus, $x' \in R_{a-1}^j$. It follows, by the plausibility of F_j, that a point $w' > c_j + 3\delta^{a+1}$ exists such that

$$d_Y(w', h(x)) \le \delta d_X(x', x).$$

In particular, we have,

$$3\delta^{a+1} \le d_Y(w', h(x)) \le \delta d_X(x', x),$$

which implies

$$\delta^a \le d_X(x', x).$$

Substituting in (3), we conclude:

$$d_Y(h(x'), h(x)) \le \delta d_X(x', x) + \delta^a \le 2\delta d_X(x', x).$$

The proof for the other case, in which $s = \max(s, s')$, follows by a symmetric argument, via considering the *first* approximate map F_j (into W_{c_j}) of $\mathcal{P}$ such that $h(x') \in B(c_j, 3\delta^2)$.

□

In order to bound the expansion for all pairs of points in X, we show that the set of relatively close pairs form a spanner of the metric space. The following lemma formalizes this idea. Note that this lemma is only about the metric space (X, d_X), and it has nothing to do with the embedding or Y.

Lemma 5.17. *Let $G_S = (X, E_S)$ be a weighted graph, with vertex set X. For $x, x' \in X$, of scales s and s' respectively, we have $(x, x') \in E_S$ (with weight $d_X(x, x')$) if and only if $d_X(x, x') \leq 3\delta^{\max(s,s')+1}$. Let (X, d_S) be the shortest path metric on G_S. We have*

$$d_S(x, x') \leq O(1) \cdot d_X(x, x').$$

Proof: Let the scale of x, x' be of scales s and s', respectively. Also, let $\delta^k \leq d_X(x, x') < \delta^{k+1}$. Finally, let z and z' be the closest points of $X_{\geq k}$ to x and x', respectively. By the definition of $X_{\geq k}$,

$$d_X(x, z) \leq \delta^k,$$

and,

$$d_X(x', z') \leq \delta^k.$$

On the other hand, since $z, z' \in X_{\geq k}$ their scale is at least k. So, $(x, z), (x', z') \in E_S$. Moreover, by the triangle inequality, we have:

$$d_X(z, z') \leq d_X(z, x) + d_X(x, x') + d_X(x', z') \leq 2\delta^k + d_X(x, x') \leq 3\delta^{k+1}. \tag{4}$$

Again, by the definition of E_S, $(z, z') \in E_S$ because the scale of z is at least k, and $d_X(z, z') \leq 3\delta^{k+1}$. Thus, (x, z, z', x') is an (x, x')-path in G_S. Therefore, we have

$$d_S(x, x') \leq d_X(x, z) + d_X(z, z') + d_X(z', x') \leq \delta^k + d_X(z, z') + \delta^k = 2\delta^k + d_X(z, z') \leq 2d_X(x, x') + d_X(z, z').$$

Also, from (4), we have:

$$d_X(z, z') \leq 2\delta^k + d_X(x, x') \Rightarrow d_X(z, z') \leq 3d_X(x, x').$$

Therefore, combining the above two inequalities,

$$d_S(x, x') \leq 2d_X(x, x') + d_X(z, z') \leq 2d_X(x, x') + 3d_X(x, x') = 5d_X(x, x')$$

□

Lemma 5.16 and Lemma 5.17 imply an $O(\delta)$ expansion for h, which is formalized in the following lemma. To bound the contraction of h, the following lemma considers the g part of the approximate maps.

Lemma 5.18. *Let X, Y, and $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be as described above. Given a directed path from a starting vertex to an ending vertex in $\mathcal{G}$, an embedding of X into Y of distortion $O(\delta)$ can be computed in polynomial time.*

Proof: First, we bound the expansion between an arbitrary pair of points $x, x' \in X$. Lemma 5.17 implies the existence of a sequence $(x = x_1, x_2, \ldots, x_t = x')$ with following properties.

1. $\sum_{i=1}^{t-1} d_X(x_i, x_{i+1}) = O(1) \cdot d_X(x, x')$.

2. For all $1 \leq i < t$, $d_X(x_i, x_{i+1}) \leq 3\delta^{\max(s_i, s_{i+1})+1}$, where s_i and s_{i+1} are the scales of x_i and x_{i+1}, respectively.

The second property and Lemma 5.16 imply for each $1 \leq i < t$, $d_Y(h(x_i), h(x_{i+1})) = O(\delta) \cdot d_X(x_i, x_{i+1})$. So, we have:

$$d_Y(h(x), h(x')) \leq \sum_{i=1}^{t-1} d_Y(h(x_i), h(x_{i+1})) = O(\delta) \cdot \sum_{i=1}^{t-1} d_X(x_i, x_{i+1}) = O(\delta) \cdot d_X(x, x').$$

Next, we bound the contraction of h. By Lemma 3.1, it suffices to bound the contraction for adjacent pairs of points in $\mathrm{Im}(h) \subseteq Y$. Let $(y_\ell, y_{\ell+1})$ be a consecutive pair in $\mathrm{Im}(h)$. Also, let $h(x) = y_\ell$, and $h(x') = y_{\ell+1}$. Consider any multi-scale window W_{c_i} such that $y_\ell \in B(c_i, \delta^2)$, and the approximate map F_i into W_{c_i} in $\mathcal{P}$. Either $y_{\ell+1} \in B(c_i, \delta^2)$ or $y_{\ell+1}$ is the leftmost point in $\mathrm{Im}(h) \cap (c_i + \delta^2, \infty)$. In the former case $F_i(x') = f_0^i(x') = y_{\ell+1}$, and in the latter case $F_i(x') = g^i(x') = y_{\ell+1}$. In both cases, $F_i(x) = h(x)$. Since F_i is plausible, we have

$$d_Y(y_\ell, y_{\ell+1}) = d_Y(h(x), h(x')) = d_Y(F_i(x), F_i(x')) \geq d_X(x, x').$$

Hence, h is non-contracting. □

Now, we are ready to present a proof for Theorem 1.2, which summarizes our algorithm for embedding a finite metric space into a finite subset of the Euclidean line.

Proof (of Theorem 1.2): Lemma 3.2 reduces the problem to $(nm)^{O(1)}$ instances of the following problem: given a real $\delta \geq 1$, compute a non-contracting embedding of X into Y with expansion at most δ, or correctly report that no such embedding exists

To find a non-contracting embedding of distortion δ, we build $\mathcal{G}$ and look for a directed path from an starting vertex to an ending vertex, based on Lemma 5.12 and Lemma 5.18. The vertices of $\mathcal{G}$ are approximate maps into standard multi-scale windows of the sequence $\mathcal{W}$, which has length $O(Sm)$. Lemma 5.10 implies there are at most $\delta^{O(\delta^2 S)}(mn)^{O(1)}$ approximate maps into a single multi-scale window. So, the the number of vertices of $\mathcal{G}$ is $\delta^{O(\delta^2 S)}(mn)^{O(1)}$. Thus, in $\delta^{O(\delta^2 S)}(mn)^{O(1)}$ time, we can check whether there is a path from a starting vertex to an ending vertex. Finally, since $S = O(\log_\delta \Delta)$, the running time of the algorithm is $\Delta^{O(\delta^2)}(mn)^{O(1)}$. □

Corollary 5.19. *There is a $\Delta^{O(\delta^2)} n^{O(1)}$ time $O(1)$-approximation algorithm to compute a minimum distortion bijection between a metric space to a point set on the real line, both of cardinality n, where δ is the minimum distortion and Δ is the spread.*

5.5 Embedding into the Euclidean line

We now extend Theorem 1.2 to obtain an algorithm for embedding a metric space into the Euclidean line. The following key lemma contains similar observations to those found in [FFL+13] and [BCIS05].

Lemma 5.20. *Let (X, d_X) be any metric space that can be embedded into the real line with distortion δ. There exists an embedding f of X into $\mathbb{Z}$ with distortion $O(\delta)$.*

Proof: Suppose, that after scaling d_X, that $d_X(x, x') \geq 1$ for all distinct $x, x' \in X$. Let f be an embedding of X into the line with distortion δ. By scaling the image of f we can assume that $c_f = 1$, and $e_f = \delta$. Therefore, for any $x, x' \in X$ we have $d_Y(f(x), f(x')) \geq 1$.

Now let $f' : X \to \mathbb{R}$ be defined as $f'(x) = \lceil f(x) \rceil$. We show that $e_{f'}$ is at most 2δ and that $c_{f'}$ is at most 2. For any distinct $x, x' \in X$, by the definition of f',

$$d_Y(f'(x), f'(x')) \leq d_Y(f(x), f(x')) + 1 \leq \delta d_X(x, x') + 1 \leq 2\delta d_X(x, x').$$

Also, by the same definition, we have

$$d_Y(f'(x), f'(x')) \geq 1.$$

Thus, if $d_X(x, x') \leq 2$ then its contraction is clearly bounded by 2. Otherwise, if $d_X(x, x') > 2$, then

$$d_Y(f'(x), f'(x')) \geq d_Y(f(x), f(x')) - 1 \geq d_X(x, x') - 1 \geq d_X(x, x')/2$$

Hence, the contraction is again bounded by 2. □

Now, we are ready to prove Theorem 1.3, that gives an algorithm for embedding a finite metric space into the real line.

Proof (of Theorem 1.3): Assume that d_X has been scaled such that $\min_{x \neq x' \in X} d_X(x, x') = 1$. Then by the definition of spread, $\Delta = \max_{x,x' \in X} d_X(x, x')$.

By Lemma 5.20, there is an embedding f into $\mathbb{Z}$, with distortion $O(\delta)$. Let $y_1, y_m \in \mathbb{Z}$ be the leftmost point and rightmost point, respectively, that f maps onto. By the definition of distortion, $|y_1 - y_m| = O(\delta\Delta)$. Therefore, if there is an embedding f into $\mathbb{Z}$, with distortion $O(\delta)$, then there is an embedding f' into $\{1, \ldots, c\delta\Delta\}$ of distortion $O(\delta)$, for some sufficiently large constant c.

Therefore, by Theorem 1.2, one can compute an $O(1)$-approximation to f' in time

$$\Delta^{O(\delta^2)}(c\delta\Delta n)^{O(1)} = \Delta^{O(\delta^2)} n^{O(1)}.$$

□

Acknowledgement

The authors would like to thank Anastasios Sidiropoulos for helpful discussion and comments.

References

[BCIS05] Mihai Badoiu, Julia Chuzhoy, Piotr Indyk, and Anastasios Sidiropoulos. Low-distortion embeddings of general metrics into the line. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, STOC '05, pages 225–233. ACM, 2005.

[BDG+05] Mihai Badoiu, Kedar Dhamdhere, Anupam Gupta, Yuri Rabinovich, Harald Räcke, R. Ravi, and Anastasios Sidiropoulos. Approximation algorithms for low-distortion embeddings into low-dimensional spaces. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '05, pages 119–128, Philadelphia, PA, USA, 2005. Society for Industrial and Applied Mathematics.

[Bou85] J. Bourgain. On lipschitz embedding of finite metric spaces in hilbert space. *Israel Journal of Mathematics*, 52(1-2):46–52, 1985.

[CMO+08] Nishanth Chandran, Ryan Moriarty, Rafail Ostrovsky, Omkant Pandey, Mohammad Ali Safari, and Amit Sahai. Improved algorithms for optimal embeddings. *ACM Trans. Algorithms*, 4(4):45:1–45:14, August 2008.

[ESZ10] Jeff Edmonds, Anastasios Sidiropoulos, and Anastasios Zouzias. Inapproximability for planar embedding problems. In *Proceedings of the Twenty-first Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '10, pages 222–235, Philadelphia, PA, USA, 2010. Society for Industrial and Applied Mathematics.

[FFL+13] Michael Fellows, Fedor V. Fomin, Daniel Lokshtanov, Elena Losievskaja, Frances Rosamond, and Saket Saurabh. Distortion is fixed parameter tractable. *ACM Trans. Comput. Theory*, 5(4):16:1–16:20, November 2013.

[HP05] Alexander Hall and Christos H. Papadimitriou. Approximating the distortion. In *APPROX-RANDOM*, pages 111–122, 2005.

[Ind01] P. Indyk. Algorithmic applications of low-distortion geometric embeddings. In *Proceedings of the 42Nd IEEE Symposium on Foundations of Computer Science*, FOCS '01, pages 10–, Washington, DC, USA, 2001. IEEE Computer Society.

[KRS04] Claire Kenyon, Yuval Rabani, and Alistair Sinclair. Low distortion maps between point sets. In *Proceedings of the Thirty-sixth Annual ACM Symposium on Theory of Computing*, STOC '04, pages 272–280, New York, NY, USA, 2004. ACM.

[LLR94] N. Linial, E. London, and Y. Rabinovich. The geometry of graphs and some of its algorithmic applications. In *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, SFCS '94, pages 577–591, Washington, DC, USA, 1994. IEEE Computer Society.

[Mat13] Jirí Matousek. *Lecture notes on metric embeddings*, 2013.
Available at: http://kam.mff.cuni.cz/~matousek/ba-a4.pdf.

[PS05] Christos Papadimitriou and Shmuel Safra. The complexity of low-distortion embeddings between point sets. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '05, pages 112–118, Philadelphia, PA, USA, 2005. Society for Industrial and Applied Mathematics.