Weighted Minimum Backward Fréchet Distance

Amin Gheibi[*†1], Anil Maheshwari[†1], and Jörg-Rüdiger Sack[*†1]

[1]School of Computer Science, Carleton University, Ottawa, ON, Canada
[agheibi,anil,sack]@scs.carleton.ca

Abstract

The minimum backward Fréchet distance (MBFD) problem is a natural optimization problem for the weak Fréchet distance, a variant of the well-known Fréchet distance. In this problem, a threshold ε and two polygonal curves, T_1 and T_2, are given. The objective is to find a pair of walks on T_1 and T_2, which minimizes the union of the portions of backward movements while the distance between the moving entities, at any time, is at most ε. In this paper, we generalize this model to capture scenarios when the cost of backtracking on the input polygonal curves is not homogeneous. More specifically, each edge of T_1 and T_2 has an associated non-negative weight. The cost of backtracking on an edge is the Euclidean length of backward movement on that edge multiplied by the corresponding weight. The objective is to find a pair of walks that minimizes the sum of the costs on the edges of the curves, while guaranteeing that the curves remain at weak Fréchet distance ε. We propose an exact algorithm whose run time and space complexity is $\mathcal{O}(n^3)$, where n is the maximum number of the edges of T_1 and T_2.

1 Introduction

Measures for similarity between two polygonal curves have been studied in areas such as computational geometry, Geographical Information Systems (GIS), pattern recognition, shape matching, and robotics. Finding measures that capture the requirements of a particular domain remains challenging, both in practice and theory. One of the widely used measures for similarity between curves is the Fréchet distance which takes into account global features of the curves [1]. In some applications (such as map matching) a global approach as taken e.g., by the Fréchet distance, achieves a better result than a local approach [2].

The Fréchet distance is typically illustrated via the person-dog metaphor. Assume that a person wants to

walk along one curve and his/her dog on another. Each curve has a starting and an ending point. The person and the dog walk, from the starting point to the ending point, along their respective curves. The standard Fréchet distance is the minimum leash length required for the person to walk the dog without backtracking. A variant of the standard Fréchet distance is the weak Fréchet distance, also known as non-monotone Fréchet distance [1]. In this variant, backtracking is allowed during the walks. In [1], Alt and Godau proposed algorithms to compute the weak Fréchet distance in $\mathcal{O}(n^2 \log n)$ time, where n is the maximum number of segments in the input polygonal curves. The time complexity is improved by Har-Peled and Raichel [6]. They proposed an algorithm with quadratic time complexity for computing a generalization of the weak Fréchet distance. In some applications, the weak Fréchet distance is preferable to the standard Fréchet distance (see [2]).

In [4], Gheibi et al. introduced and solved an optimization problem on the weak Fréchet distance, called the *minimum backward Fréchet distance (MBFD)* problem. Their problem is to determine the minimum total length of backward movements on both input polygonal curves, required for the walks to achieve the given leash length. In that paper, it is assumed that the cost (i.e., weight) of backward movement is uniform and depends only on the Euclidean distance traveled on each of the input polygonal curves. They proposed an algorithm with time complexity $\mathcal{O}(n^2 \log n)$ and space complexity $\mathcal{O}(n^2)$, to solve MBFD exactly. Here, in this paper, we generalize this model to capture scenarios when the cost of backtracking on the input polygonal curves is not homogeneous. These weights could represent, for example, the cost of moving against a flow, or the cost for a moving entity (e.g., a human) to move backwards because of the entity's physiology [3]. Thus, in the new model, each edge of the input polygonal curves has an associated non-negative weight for backward movement. Then, the cost of backtracking on an edge is the Euclidean length of backward movement on that edge multiplied by the corresponding weight. The objective is to design an algorithm that a) finds a pair of walks that minimizes the sum of the costs on the edges of the curves, while guaranteeing that the leash length is at most ε, b) halts
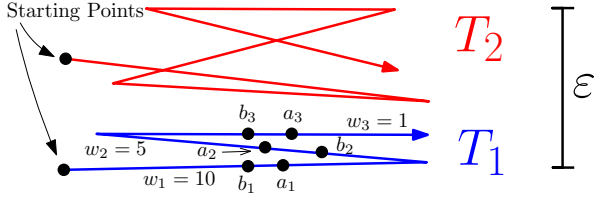
Figure 1: Moving backwards from a_3 to b_3 allows to walk on T_1 and T_2 and keeping the distance between moving entities at most ε during the walks while the cost is minimized.

with the answer of no feasible solution if such a pair of walks does not exist for the given leash length. We call this problem, the *weighted minimum backward Fréchet distance (WMBFD) problem*. Note that if the standard Fréchet distance between the input curves is already at most ε, then no backtracking is necessary and the optimal solution is identical to a pair of walks that realizes the Fréchet distance.

Figure 1 shows an example. In this figure, two polygonal curves, T_1 and T_2, and a length ε are drawn. The person walks on T_1 and the dog walks on T_2. The weights, w_i, $i = 1, 2, 3$, for segments of T_1 are given. For this illustration, we let the cost of backtracking on all segments of T_2 be 1. In this example, it is impossible to walk from the starting point to the end and maintain the leash length at most ε, without moving backwards. Six points, a_1, b_1, a_2, b_2, a_3, and b_3 are specified on T_1. If the person moves backwards, either from a_1 to b_1, or from a_2 to b_2, or from a_3 to b_3, then the curves are at weak Fréchet distance ε. In this example, the Euclidean length of $\overline{a_1 b_1}$ is less than the Euclidean length of $\overline{a_3 b_3}$. However, the weight of moving backwards on the first segment is 10, while that on the third one is 1. Therefore, the pair of walks that minimizes the cost is as follows: the dog and the person move forwards together from the starting point, until the dog reaches the end of the third segment of T_2 and the person reaches the point a_3 on T_1. Then, the dog keeps moving forwards until the end of the fourth segment of T_2, while the person moves backwards from a_3 to b_3. Finally, they move forwards again together until the end of the respective curves. The cost of this pair of walks is the Euclidean length of $\overline{a_3 b_3}$ multiplied by 1.

This paper is organized as follows. In Section 2, we discuss preliminaries and define the problem formally. In Section 3, we propose a polynomial time algorithm to solve the problem exactly. Then, in Section 4, we design an algorithm with improved time and space complexity. At the end, we conclude the paper.

2 Preliminaries and Problem Definition

In this section, first, preliminary concepts are discussed. Then, the WMBFD problem is defined formally. A geometric path in $\mathbb{R}^2$ is a sequence of points in the Eu-

clidean space, $\mathbb{R}^2$. A discrete geometric path, or a polygonal curve, is a geometric path, sampled by a finite sequence of points (i.e., vertices), which are connected by line segments (i.e., edges) in order. Let $T_1 : [0, n] \to \mathbb{R}^2$ and $T_2 : [0, m] \to \mathbb{R}^2$ be two polygonal curves of complexity (number of segments) n and m, respectively. W.l.o.g., assume that $m \leq n$. A vertex of T_1 (resp. T_2) is denoted by $T_1(i)$ (resp. $T_2(j)$), $i = 0, \ldots, n$ (resp. $j = 0, \ldots, m$). An edge of T_1 (resp. T_2) between two vertices $T_1(i - 1)$ and $T_1(i)$ (resp. $T_2(j - 1)$ and $T_2(j)$) is denoted by e_i (resp. e_j), $i = 1, \ldots, n$ (resp. $j = 1, \ldots, m$). Furthermore, each edge, e_i (resp. e_j), $i = 1, \ldots, n$ (resp. $j = 1, \ldots, m$), of T_1 (resp. T_2) has an associated non-negative weight (or cost) $w_i \in \mathbb{R}$ (resp. $w_j \in \mathbb{R}$). A parameterization of a polygonal curve, $T_1 : [0, n] \to \mathbb{R}^2$, is a continuous function $f : [0, 1] \to [0, n]$, where $f(0) = 0$ and $f(1) = n$ ($[0, 1]$ is a time interval). If f is non-decreasing, then the parameterization is monotone. The weak Fréchet distance, $\delta_w(T_1, T_2)$, is defined as Formula 1, where $d(.,.)$ is the Euclidean distance and f and g are two parameterization of $[0, n]$ and $[0, m]$, respectively. Note that f and g are not necessarily monotone. However, for the standard Fréchet distance, they must be monotone.

$$\delta_w(T_1, T_2) = \inf_{f,g} \max_{t \in [0,1]} d(T_1(f(t)), T_2(g(t))) \quad (1)$$

Weighted Quality. For a parameterization, f, of a polygonal curve, T_1, let $\mathcal{B}_{f,i} \subseteq [0, 1]$ be the closure of the set of times in which $f(t)$ is decreasing (i.e., the movement is backward), and $f(t) \in [i - 1\ i)$ (it is on edge e_i of T_1). $\mathcal{B}_{g,j} \subseteq [0, 1]$ is defined analogously for a parameterization, g, of T_2. For a pair of parameterizations, f and g, of two polygonal curves, T_1 and T_2, we define the weighted quality by Formula 2, where $\|.\|$ is the Euclidean length.

$$
\begin{aligned}
\mathcal{WQ}_{f,g}(T_1, T_2) := & \sum_{i=1}^{n} \|T_1(f(t))\|_{t \in \mathcal{B}_{f,i}} \cdot w_i \\
& + \sum_{j=1}^{m} \|T_2(g(t))\|_{t \in \mathcal{B}_{g,j}} \cdot w_j
\end{aligned}
\quad (2)
$$

Problem Definition. We formally define the WMBFD problem as follows. For a pair of weighted polygonal curves, T_1 and T_2, and a given leash length, ε, we are looking for a pair of optimal parameterizations, (f, g), as defined in Formula 3. We consider only pairs of parameterizations that guarantee to maintain the leash length at most ε, during the walks.

$$\mathcal{WQ}^{\varepsilon}(T_1, T_2) = \inf_{f,g} \mathcal{WQ}_{f,g}(T_1, T_2) \quad (3)$$

Weighted Deformed free-space diagram. A useful structure to decide whether the Fréchet distance between two polygonal curves is upper bounded by a given ε, is the free-space diagram [1]. For two polygonal curves, T_1 with n vertices and T_2 with m vertices,

and two corresponding parameterizations, f and g, the *free-space* is defined formally by Formula 4.

$$W = \{(t_1, t_2) \in [0,1]^2 \mid d(T_1(f(t_1)), T_2(g(t_2))) \le \varepsilon\} \tag{4}$$

The *free-space diagram* is the rectangle $[0,1] \times [0,1]$, partitioned into n columns and m rows. It consists of nm parameter cells $C^{i,j}$, for $i = 1, ..., n$ and $j = 1, ..., m$, whose interiors do not intersect with each other. The cell $C^{i,j}$ represents the multiplication of two subranges of $[0,1]$ that are mapped to the edge between vertices $T_1(i-1)$ and $T_1(i)$ and the edge between vertices $T_2(j-1)$ and $T_2(j)$. For each parameter cell $C^{i,j}$, there exists an ellipse such that the intersection of the area bounded by this ellipse with $C^{i,j}$ is equal to the free-space region of that cell. The boundary of this ellipse and the boundary of the cell, $C^{i,j}$, intersect at most eight times (i.e., at most two intersections per side of $C^{i,j}$). These intersection points form at most four intervals on the boundary of $C^{i,j}$. These intervals could be empty or contains only one point. In addition, two adjacent cells have the same interval on the shared side between the cells. The union of all cells' free-space builds the free-space (or white-space) of the diagram and is denoted by W. The complement of W is the forbidden-space (or black-space) of the diagram and is denoted by B. In this paper, we stretch and compress the columns and rows of the free-space diagram, such that their widths and heights are equal to the lengths of the corresponding segments of T_1 and T_2, respectively. Also, each cell, $C^{i,j}$, has two associated weights, w_x^i and w_y^j. The weight w_x^i is the weight of the edge between vertices $T_1(i-1)$ and $T_1(i)$ and the weight w_y^j is the weight of the edge between vertices $T_2(j-1)$ and $T_2(j)$. The resulting diagram is called the *weighted deformed free-space diagram* and is denoted by $\mathcal{F}$. The bottom left corner of $\mathcal{F}$ represents the starting points of T_1 and T_2 and is denoted by s. The top right corner of $\mathcal{F}$ represents the ending points of T_1 and T_2 and is denoted by t. For the given polygonal curves and ε in Figure 1, the corresponding weighted deformed free-space diagram is shown in Figure 2. As the diagram illustrates, to be able to walk on T_1 and T_2 with a leash length at most ε, there must be a backward movement on the polygonal curves (since there is no xy-monotone path from s to t in W). However, the possible walks are not unique. We are looking for a pair of walks that has the minimum backward movement cost, as we discussed in Section 1. In Figure 2, the red solid polygonal chain, called Π', is a path in W that realizes an optimal pair of walks on T_1 and T_2.

3 Algorithm

In this section, we propose a polynomial time algorithm, by transforming the WMBFD problem to a
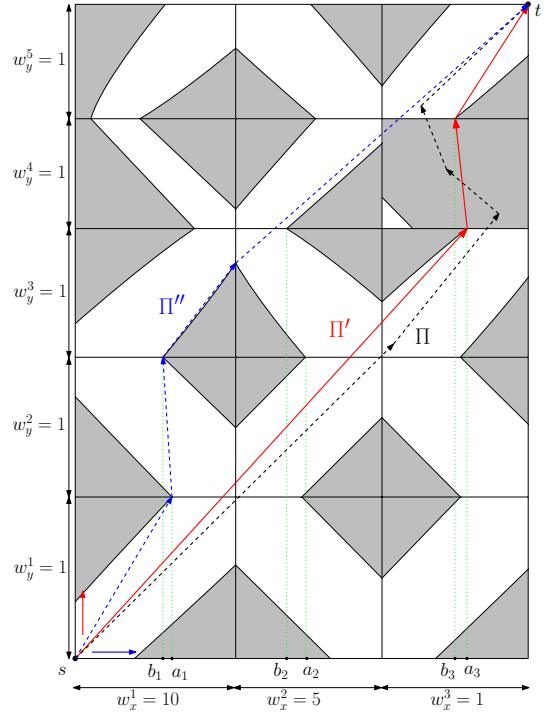

Figure 2: The corresponding weighted deformed free-space diagram of the given polygonal curves in Figure 1 is drawn. Π (the black dashed path) is an arbitrary path in W. $\Pi' \subset \mathcal{G}_w$ (the red solid path) is a path in W that realizes a pair of parameterizations which gives an optimal solution for WMBFD. Π'' (the blue dashed path) is a path in W that realizes the optimal solution for MBFD.

shortest path problem on a weighted directed graph, $\mathcal{G}_w = \langle V, E \rangle$, defined as follows.

Let $\mathcal{F}$ be the weighted deformed free-space diagram and W (resp. B) be the corresponding free-space (resp. forbidden-space) of $\mathcal{F}$. The vertices of W are the end points of the intervals on the boundary of the cells in $\mathcal{F}$ (i.e., at most 4 intervals per cell). The set of vertices, V, of $\mathcal{G}_w$, is the set of all vertices of W. Each vertex, v, has a x-coordinate (resp. y-coordinate), denoted by v_x (resp. v_y). Also, V contains s and t. We say two points are visible if it is possible to link them by a line segment in W. Every two visible vertices, v_1 and v_2, are linked by two directed edges in E, from v_1 to v_2, $\langle v_1, v_2 \rangle$, and vice versa, $\langle v_2, v_1 \rangle$. The weight of a directed edge $e = \langle v_1, v_2 \rangle \in E$ is a function of its direction, the x- and y-coordinates of v_1 and v_2, and the associated weights of the cells that e intersects. The weight function is defined as follows: suppose e intersects a sequence of k cells, $\langle C^{\sigma(1),\sigma'(1)}, C^{\sigma(2),\sigma'(2)}, \ldots, C^{\sigma(k),\sigma'(k)} \rangle$, of $\mathcal{F}$, where σ (resp. σ') is a function that maps the set $\{1, 2, \ldots, k\}$ to a sub-sequence (or a reversed sub-sequence) of the index sequence $\langle 1, 2, \ldots, n \rangle$ (resp. $\langle 1, 2, \ldots, m \rangle$). The line segment e enters a cell, $C^{\sigma(i),\sigma'(i)}$, $i = 1, \ldots, k$, at point $a^{\sigma(i),\sigma'(i)}$ and exits that cell at point $b^{\sigma(i),\sigma'(i)}$.

Note that $a^{\sigma(1),\sigma'(1)}$ (resp. $b^{\sigma(k),\sigma'(k)}$) is identical to v_1 (resp. v_2). The x-coordinate (resp. y-coordinate) of a point, a, is denoted by a_x (resp. a_y). Note that each cell, $C^{\sigma(i),\sigma'(i)}$, has two associated weights, $w_x^{\sigma(i)}$ and $w_y^{\sigma'(i)}$.

Let $|e|_{w_x} = \sum_{i=1}^{k} |a_x^{\sigma(i),\sigma'(i)} - b_x^{\sigma(i),\sigma'(i)}| \cdot w_x^{\sigma(i)}$ and $|e|_{w_y} = \sum_{i=1}^{k} |a_y^{\sigma(i),\sigma'(i)} - b_y^{\sigma(i),\sigma'(i)}| \cdot w_y^{\sigma'(i)}$. The weight of e, $|e|_w$, is calculated by the following function.

- If e is xy-increasing (i.e., it is non-decreasing from v_1 to v_2 in both x and y axes), then $|e|_w = 0$.

- If e is only x-increasing (resp. y-increasing), then $|e|_w = |e|_{w_y}$ (resp. $|e|_w = |e|_{w_x}$).

- Otherwise, $|e|_w = |e|_{w_x} + |e|_{w_y}$.

Finding an Optimal Solution. By construction of $\mathcal{G}_w$, both s and t are vertices in V. If either s or t is not in W, or there is no path from s to t in $\mathcal{G}_w$, then there is no solution for the given leash length. Otherwise, we prove that a shortest path from s to t, in $\mathcal{G}_w$, gives an optimal solution. Note that a vertex of the graph also corresponds to a point in $\mathcal{F}$. Therefore, the geometric embedding of a path in $\mathcal{G}_w$ is constructed by connecting the consecutive vertices of the path by line segments.

Observation 1 *Let $\Pi : [0,1] \rightarrow [0,n] \times [0,m]$ be a path in the free-space W, from s to t. Π is equivalent to a pair of parameterizations, $f : [0,1] \rightarrow [0 : n]$ and $g : [0,1] \rightarrow [0 : m]$, of the two polygonal curves, that maintains the leash length at most ε, for all $t \in [0,1]$.*

In this paper, we use norms in two spaces: (1) the Euclidean space of the input polygonal curves, called the input space, (2) the weighted deformed free-space diagram, called the configuration space. In the input space, we use the Euclidean length of a polygonal curve T and denote it by $||T||$. In the configuration space, a path from s to t in W, is denoted by its vertices, $\Pi : \langle s = p_1, p_2, \ldots, p_k = t \rangle$. All segments in Π are directed, $\overrightarrow{p_i p_{i+1}}$, $i = 1, \ldots, k-1$. The weighted length (or simply length), $|.|_w$, of each segment of Π is calculated by the weight function that we explained earlier in this section, for computing the weight of a directed edge in the graph. The weighted length (or simply length) of a path, $|\Pi|_w$, is the sum of the length of its segments. In addition, the notation Π_i is used to denote the sub-path of Π from p_1 to p_i.

Correctness. Lemma 3 is at the heart of the correctness proof. In order to prove that lemma, we need Lemmas 1 and 2. Their proofs are provided in the Appendix. This section is concluded by a corollary to Lemma 3 and Observation 1, that is, in order to find an optimal pair of parameterizations in our problem setting, it suffices to find a shortest path from s to t in $\mathcal{G}_w$.
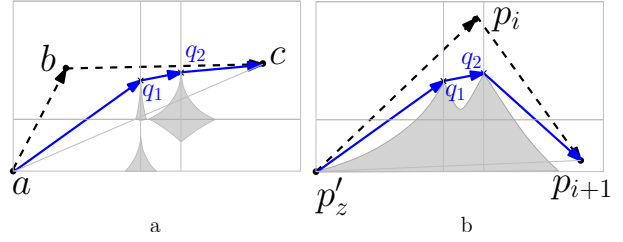

Figure 3: a) The visibility chain from a to c (the blue solid polygonal chain), $CC_a^c = \langle a, q_1, q_2, c \rangle$. b) The visibility chain from p_z' to p_{i+1}, $CC_{p_z'}^{p_{i+1}}$ (see Algorithm 1).

Definition 1 *A path $\Pi \in W$ is x-monotone (resp. y-monotone), if and only if, any vertical (resp. horizontal) line intersects it at most once. Π is xy-monotone, if and only if, it is both x- and y-monotone.*

Observation 2 *Let a and b be two points in W such that $a_x \neq b_x$ and $a_y \neq b_y$. Suppose Π is a xy-monotone path from a to b. In addition, let $R(a,b)$ be the axes-aligned rectangle uniquely determined by a and b as corners. Π lies inside $R(a,b)$.*

Lemma 1 *Let Π_1 and Π_2 be two xy-monotone paths in W, from a to b, where $a,b \in W$. Then, $|\Pi_1|_w = |\Pi_2|_w$. Furthermore, if Π_3 is an arbitrary path in W from a to b, then $|\Pi_1|_w \leq |\Pi_3|_w$.*

Definition 2 *Let a, b, and c be three distinct non-collinear points in W such that $\overrightarrow{ab} \in W$, $\overrightarrow{bc} \in W$ and $\overrightarrow{ac} \notin W$. We define the visibility chain from a to c, denoted by CC_a^c, as follows (see Figure 3a). Let B_{abc} denote the portion of B (the black-space) inside the triangle $\triangle abc$. Let CH be the convex hull of B_{abc} and the points a and c. Then, CC_a^c is defined to be the chain comprising the boundary of CH from a to c that lies inside $\triangle abc$. The visibility chain is directed from a to c, $CC_a^c = \langle a, q_1, \ldots, q_{last}, c \rangle$.*

Lemma 2 *Let $a, b, c \in W$ be three distinct non-collinear points that $\overrightarrow{ab}, \overrightarrow{bc} \in W$ and $\overrightarrow{ac} \notin W$. If $\triangle abc$ lies in $R(a,c)$, then CC_a^c is xy-monotone and $|CC_a^c|_w = |\overrightarrow{ab}|_w + |\overrightarrow{bc}|_w$ (Figure 3a).*

Lemma 3 *For any path $\Pi : \langle s = p_1, p_2, \ldots, p_{k_1} = t \rangle$ in W, there is a path $\Pi' : \langle s = p_1', p_2', \ldots, p_{k_2}' = t \rangle$ in W such that $\Pi' \subset \mathcal{G}_w$ and $|\Pi'|_w \leq |\Pi|_w$.*

Proof. We prove this lemma by designing an algorithm that constructs the path $\Pi' \subset \mathcal{G}_w$, through a transformation of path Π. Initially, Π' contains only $s = p_1' = p_1$ and $p_z' = s$. In this algorithm, p_z' is the latest vertex appended to the tail of Π'. Π' is constructed as follows. When considering the i-th vertex of Π, p_i, the algorithm tests if $\overrightarrow{p_z' p_{i+1}} \in W$. If so, p_i is skipped and Π' remains unchanged. Otherwise, the visibility chain from p_z' to

p_{i+1} is constructed (Figure 3b) and its vertices from q_1 to q_c are appended to the tail of Π'. The algorithm for constructing Π' is stated in Algorithm 1. The correctness of this algorithm is given in the Appendix, Lemma 5. The output, Π', of this algorithm is a path from s to t, such that $\Pi' \subset \mathcal{G}_w$ and $|\Pi'|_w \le |\Pi|_w$. $\qquad \square$

Algorithm 1 Constructing $\Pi' \in \mathcal{G}_w$

Input: The free-space W, A path $\Pi = \langle p_1, p_2, \ldots, p_{k_1} \rangle$, where $s = p_1$ and $p_{k_1} = t$.
Output: A path $\Pi' = \langle p'_1, p'_2, \ldots, p'_{k_2} \rangle$, where $s = p'_1$ and $p'_{k_2} = t$, such that $\Pi' \subset \mathcal{G}_w$ and $|\Pi'|_w \le |\Pi|_w$.

1: $\Pi' := \langle s \rangle$;
2: $p'_z = s$;
3: **for** i=2 to $k_1 - 1$ **do**
4:     **if** $\overrightarrow{p'_z p_{i+1}} \notin W$ **then**
5:        Compute the visibility chain from p'_z to p_{i+1}, $CC^{p_{i+1}}_{p'_z} = \langle p'_z, q_1, \ldots, q_c, p_{i+1} \rangle$, in $\triangle p'_z p_i p_{i+1}$;
6:        Append q_j, $j = 1, \ldots, c$, to Π';
7:        $p'_z = q_c$;
8: Append t to Π';
9: **return** Π';

Corollary 1 *There is a path from s to t in $\mathcal{G}_w$ with minimum weighted length in the free-space W.*

Proof. Assume Π is a path from s to t with minimum weighted length in the free-space W. If Π is not a subset of $\mathcal{G}_w$, then, by Lemma 3 there is a path from s to t, Π' in W such that $\Pi' \subset \mathcal{G}_w$ and $|\Pi'|_w \le |\Pi|_w$. Since Π has minimum weighted length, $|\Pi'|_w = |\Pi|_w$. $\qquad \square$

Theorem 1 *Let T_1 and T_2 be two polygonal curves and ε be a given leash length. Each segment of T_1 and T_2 has an associated weight, corresponding to the backward movement on that segment. A pair of parameterizations of T_1 and T_2 that minimizes the weighted sum of the backward movements during the walks can be found in polynomial time and space.*

Proof. It follows from Observation 1 and Corollary 1 that a shortest path in $\mathcal{G}_w$ yields an optimal pair of parameterizations for the WMBFD problem. Since $\mathcal{F}$ has a complexity of $\mathcal{O}(n^2)$, the number of edges of $\mathcal{G}_w$ is $\mathcal{O}(n^4)$. We construct the topology of the graph in $\mathcal{O}(n^4)$ time by the method in [5]. Since the weight of each edge of $\mathcal{G}_w$ is computed based on the projections onto x and y axes, it is possible to compute it in constant time using prefix sums [7]. We find a shortest path in the graph in $\mathcal{O}(n^4)$ time by Dijkstra's algorithm. Therefore, the total time complexity is $\mathcal{O}(n^4)$. $\qquad \square$

4 Improved Algorithm

In Section 3, we showed that the weighted graph $\mathcal{G}_w = \langle V, E \rangle$ contains a path that yields an optimal pair of parameterizations for the WMBFD problem. In this section, we will discuss that it is sufficient to compute only a subgraph of $\mathcal{G}_w$ to obtain an optimal solution. Let $\mathcal{G}_w' = \langle V, E' \rangle$ be a sub-graph of $\mathcal{G}_w$ such that $E' = \{e' \in E \mid e'$ lies completely within a row or within a column of $\mathcal{F}\}$, where $\mathcal{F}$ is the weighted free-space diagram.

Lemma 4 *There is a path in $\mathcal{G}_w'$ that realizes an optimal pair of parameterizations for our problem setting.*

Proof. We will show that, for any directed edge $e = \langle u_1, u_2 \rangle \in E$ that is not in E', we can construct a xy-monotone path from u_1 to u_2, π_{u_1, u_2}, in $\mathcal{G}_w'$ (see Figure 4). Then, by Lemma 1, $|\pi_{u_1, u_2}|_w = |e|_w$. By Theorem 1, a shortest path, Π' in $\mathcal{G}_w$ yields an optimal solution. Therefore, if for any directed edge, $e = \langle u_1, u_2 \rangle$, of Π', π_{u_1, u_2} exists in $\mathcal{G}_w'$, then there is a path in $\mathcal{G}_w'$ that realizes an optimal pair of parameterizations.

Now, we prove that π_{u_1, u_2} exists in $\mathcal{G}_w'$, for any directed edge $e = \langle u_1, u_2 \rangle \in E$. If e stays completely within a row or a column of $\mathcal{F}$, then $\pi_{u_1, u_2} = e$. Otherwise, e crosses several rows and columns. There are four cases, depending on the orientation of e: a) xy-increasing b) x-increasing and y-decreasing c) y-increasing and x-decreasing d) xy-decreasing. We prove this lemma for the last case. The proofs for the other 3 cases are analogous. Assume that e is xy-decreasing. The edge e intersects a sequence of intervals on the boundary of the cells of $\mathcal{F}$. We partition e into sub-edges so that each sub-edge is contained within a row or within a column of $\mathcal{F}$, as follows (Figure 4).

We traverse e from u_1 to u_2. The point $p_1 \in e$ is the point where we exit the row and the column that contain u_1. Therefore, any point after p_1 on e during the traversal is not in the row or the column that contains u_1. We continue the traversal from p_1 to u_2. The point $p_2 \in e$ is defined analogously. It is the point where we exit the row and the column that contain p_1. We define p_i, $i = 3, \ldots, z$, analogously with respect to p_{i-1}. Then, the sequence of sub-edges of e is $\langle \overrightarrow{u_1 p_1}, \overrightarrow{p_1 p_2}, \ldots, \overrightarrow{p_z u_2} \rangle$.

Denote the interval that contains p_i, $i = 1, \ldots, z$, by I_i. Let $u_1 = p_0 \in I_0$ and $u_2 = p_{z+1} \in I_{z+1}$. Note that I_i and I_{i+1}, $i = 0, \ldots, z$, are on the boundary of a row or a column. We say a point $q = (q_x, q_y)$ dominates a point $p = (p_x, p_y)$, if $p_x \le q_x$ and $p_y \le q_y$. In this proof, the endpoint of I_i that dominates p_i is denoted by q_i. We construct π_{u_1, u_2}, in two phases. In the first phase, we construct a xy-monotone path, π'_{u_1, u_2}, from u_1 to u_2 (the green dashed polygonal chain in Figure 4). Then, in the second phase, we transform it to a path, π_{u_1, u_2}, in $\mathcal{G}_w'$ (the blue dotted polygonal chain in Figure 4).

In the first phase, we start from $p_0 = u_1$. For each sub-edge of e, $\overrightarrow{p_i p_{i+1}}$, $i = 0, \ldots, z$, if we construct a xy-monotone path, $\pi'_{p_i, q_{i+1}}$, from p_i to q_{i+1}, then the concatenation of $\pi'_{p_i, q_{i+1}}$ and $\overrightarrow{q_{i+1} p_{i+1}}$ is a xy-monotone path from p_i to p_{i+1}, $\pi'_{p_i, p_{i+1}}$, because q_{i+1} dominates
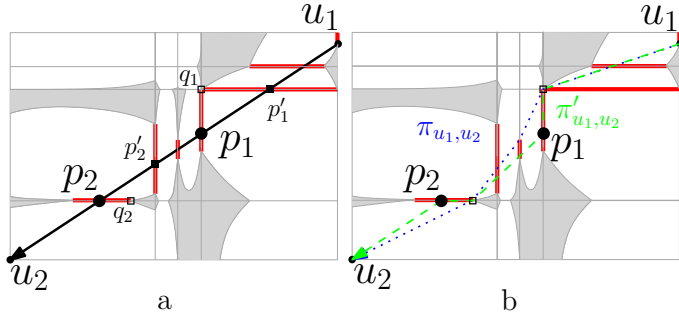
Figure 4: Illustration of proof of Lemma 4.

p_{i+1}. Then, the concatenation of $\pi'_{p_i,p_{i+1}}$, $i = 0, \ldots, z$ is a xy-monotone path, π'_{u_1,u_2}, from u_1 to u_2. Now, we explain how to construct $\pi'_{p_i,q_{i+1}}$. If $\overrightarrow{p_iq_{i+1}} \in W$, then $\pi'_{p_i,q_{i+1}} = \overrightarrow{p_iq_{i+1}}$. It is obviously xy-monotone. If $\overrightarrow{p_iq_{i+1}} \notin W$, then $\pi'_{p_i,q_{i+1}}$ is the visibility chain, $CC^{q_{i+1}}_{p_i}$, from p_i to q_{i+1}, in $\triangle p_i p'_{i+1} q_{i+1}$, where p'_{i+1} is a point, defined as follows. Let I'_{i+1} be the last interval that $\overrightarrow{p_ip_{i+1}}$ intersects before intersecting I_{i+1} and p'_{i+1} be the intersection point of I'_{i+1} and $\overrightarrow{p_ip_{i+1}}$. The point p'_{i+1} dominates q_{i+1}. Therefore, $\triangle p_i p'_{i+1} q_{i+1}$ lies in $R(p_i, q_{i+1})$, the axes-aligned rectangle that is determined by p_i and q_{i+1} as opposite corners. Thus, by Lemma 2, $CC^{q_{i+1}}_{p_i}$ is xy-monotone.

In the second phase, we transform π'_{u_1,u_2} to a xy-monotone path, π_{u_1,u_2}, in $\mathcal{G}_w'$. This transformation is done by replacing the edges in π'_{u_1,u_2} that are not in E'. These edges are $\overrightarrow{q_ip_i}$ and $\overrightarrow{p_i\mathcal{S}(p_i)}$, $i = 1, \ldots, z$, where $\mathcal{S}(.)$ is the successor operation and $\mathcal{S}(p_i)$ is the vertex after p_i in π'_{u_1,u_2}. These are the edges that connect p_i, $i = 1, \ldots, z$, to the previous and next vertex of p_i in π'_{u_1,u_2}. Note that $\mathcal{S}(p_i)$ is a vertex in V and could be identical to q_{i+1}. If $\overrightarrow{q_i\mathcal{S}(p_i)} \in W$, then the two edges, $\overrightarrow{q_ip_i}$ and $\overrightarrow{p_i\mathcal{S}(p_i)}$, are replaced by $\overrightarrow{q_i\mathcal{S}(p_i)} \in E'$. It is obviously xy-monotone. If $\overrightarrow{q_i\mathcal{S}(p_i)} \notin W$, then the two edges, $\overrightarrow{q_ip_i}$ and $\overrightarrow{p_i\mathcal{S}(p_i)}$, are replaced by the visibility chain, $CC^{\mathcal{S}(p_i)}_{q_i}$, from q_i to $\mathcal{S}(p_i)$, in $\triangle q_i p_i \mathcal{S}(p_i)$. Since π'_{u_1,u_2} is a xy-monotone path, the concatenation of $\overrightarrow{q_ip_i}$ and $\overrightarrow{p_i\mathcal{S}(p_i)}$ is also a xy-monotone path. Therefore, $\triangle q_i p_i \mathcal{S}(p_i)$ lies in $R(q_i, \mathcal{S}(p_i))$. Thus, by Lemma 2, $CC^{\mathcal{S}(p_i)}_{q_i}$ is xy-monotone. Also, $CC^{\mathcal{S}(p_i)}_{q_i} \subset \mathcal{G}_w'$ since the vertices of this visibility chain belong to one column or one row of $\mathcal{F}$. By repeating this process for every p_i, $i = 1, \ldots, z$, the resulting path, denoted by π_{u_1,u_2}, is in $\mathcal{G}_w'$. Since all sub-paths of π_{u_1,u_2} are xy-monotone, π_{u_1,u_2} is also xy-monotone. $\square$

Theorem 2 *Assume we are given two polygonal curves, T_1 and T_2, and a leash length, ε. Each segment of T_1 and T_2 has an associated weight, corresponding to the backward movement on that segment. A pair of parameterizations of T_1 and T_2 that minimizes the weighted sum of the backward movements during the*

walks can be found in $\mathcal{O}(n^3)$ time and space, where n is the number of segments in the input polygonal curves.

Proof. The correctness follows from Lemma 4. $\mathcal{F}$ has $\mathcal{O}(n^2)$ cells and each vertex of $\mathcal{G}_w'$ on the boundary of a cell is connected to at most $\mathcal{O}(n)$ vertices of $\mathcal{G}_w'$ that are in the same row or column. Therefore, the number of edges of $\mathcal{G}_w'$ is $\mathcal{O}(n^3)$. It is possible to find all the edges of $\mathcal{G}_w'$ that lie in a column or row of $\mathcal{F}$ in $\mathcal{O}(n^2)$ time by the method proposed in [5]. In addition, to compute the weight of the edges that are in one row or column, $\mathcal{O}(n^2)$ time and $\mathcal{O}(n)$ space suffice (by using prefix sums, see [7]). Using Dijkstra's algorithm, we find a shortest path in $\mathcal{G}_w'$ in $\mathcal{O}(n^3)$ time. Therefore, both time and space complexities of our algorithm are $\mathcal{O}(n^3)$. Note that if the representing nodes for s and t in $\mathcal{G}_w'$ are not in a connected component of $\mathcal{G}_w'$, then there is no feasible walk with the leash length of ε. $\square$

5 Conclusion

In this paper, we generalized the MBFD problem by capturing weighted scenarios. We established that this problem setting is dual to a weighted shortest path problem in a weighted deformed free-space diagram, $\mathcal{F}$. We proposed an exact algorithm to solve the problem in $\mathcal{O}(n^3)$ time and space. We are currently working on improving the time complexity.

References

[1] H. Alt and M. Godau. Computing the Fréchet distance between two polygonal curves. *Int. J. Comput. Geometry Appl.*, 5:75–91, 1995.

[2] S. Brakatsoulas, D. Pfoser, R. Salas, and C. Wenk. On map-matching vehicle tracking data. *31st VLDB*, pp. 853–864, 2005.

[3] T. Flynn, S. Connery, M. Smutok, R. Zeballos, and I. Weisman. Comparison of cardiopulmonary responses to forward and backward walking and running. *Med. Sci. Sports Exerc.*, 26(1):89–94, 1994.

[4] A. Gheibi, A. Maheshwari, J.-R. Sack, and C. Scheffer. Minimum Backward Fréchet Distance. *22nd ACM SIGSPATIAL*, pp. 381–388, 2014.

[5] S. K. Ghosh and D. M. Mount. An output-sensitive algorithm for computing visibility graphs. *SIAM J. Comput.*, 20(5):888–910, 1991.

[6] S. Har-Peled and B. Raichel. The Fréchet Distance Revisited and Extended. *27th ACM SoCG*, pp. 448–457, 2011.

[7] G. E. Blelloch. Prefix Sums and Their Applications. *Synthesis of Parallel Algorithms*, Morgan Kaufmann, 1990.

6   Appendix

Lemma 1 *Let Π_1 and Π_2 be two xy-monotone paths in W, from a to b, where $a, b \in W$. Then, $|\Pi_1|_w = |\Pi_2|_w$. Furthermore, if Π_3 is an arbitrary path in W from a to b, then $|\Pi_1|_w \leq |\Pi_3|_w$.*

Proof. If $a_x = b_x$ or $a_y = b_y$, then Π_1 and Π_2 are identical and the proof is trivial. Otherwise, by Observation 2, Π_1 and Π_2 lie in $R(a, b)$. Since Π_1 (also Π_2) is xy-monotone, its orthogonal projections onto x and y axes are not overlapping and equal to the width and height of $R(a, b)$, respectively. Because Π_1 and Π_2 have identical projections onto x and y axes and the weighted length of a path is defined based on its projection, then $|\Pi_1|_w = |\Pi_2|_w$. Also, any xy-monotone path from a to b has minimum weighted length among all paths from a to b, because its orthogonal projections onto x- and y-axis are non-overlapping. □

Lemma 2 *Let $a, b, c \in W$ be three distinct non-collinear points that $\overrightarrow{ab}, \overrightarrow{bc} \in W$ and $\overrightarrow{ac} \notin W$. If $\triangle abc$ lies in $R(a, c)$, then CC_a^c is xy-monotone and $|CC_a^c|_w = |\overrightarrow{ab}|_w + |\overrightarrow{bc}|_w$ (Figure 3a).*

Proof. Since $\triangle abc$ lies in $R(a, c)$, the path that consists of $\overrightarrow{ab}$ and $\overrightarrow{bc}$ is a xy-monotone path from a to c. If we show that CC_a^c is also a xy-monotone path from a to c, then by Lemma 1, $|CC_a^c|_w = |\overrightarrow{ab}|_w + |\overrightarrow{bc}|_w$.

To prove this, we need to define the angle of a vector. Suppose a directed segment in the free-space is a vector from the origin of the Cartesian coordinate system. The angle of a vector is defined as the angle between that vector and the positive direction of x-axis. Let α (resp. β) be the angle of $\overrightarrow{ab}$ (resp. $\overrightarrow{bc}$). Since $\triangle abc$ lies in $R(a, c)$, $|\alpha - \beta| = \pi/2$. In addition, since $R(a, c)$ is axes-aligned, precisely one of the four following cases is true: $0 \leq \alpha, \beta \leq \pi/2$, $\pi/2 \leq \alpha, \beta \leq \pi$, $\pi \leq \alpha, \beta \leq 3\pi/2$, $3\pi/2 \leq \alpha, \beta \leq 2\pi$. We denote the angles of segments, $\overrightarrow{aq_1}, \overrightarrow{q_1 q_2}, \ldots, \overrightarrow{q_{last}c}$, of $CC_a^c = \langle a, q_1, \ldots, q_{last}, c \rangle$ by θ_μ, $\mu = 1, \ldots, last + 1$. Since CC_a^c is a convex chain, the sequence of θ_μ, $\mu = 1, \ldots, last + 1$, is in a sorted order (either increasing or decreasing), between α and β. Therefore, all θ_μ, $\mu = 1, \ldots, last + 1$, are in one of the four mentioned quadrants. Thus CC_a^c is xy-monotone. This proves the lemma. □

Lemma 5 *Algorithm 1 is correct.*

Proof. We show that prior to the execution of the i-th iteration, $i = 2, \ldots, k_1 - 1$, of the **for**-loop in Algorithm 1 the following invariant holds:

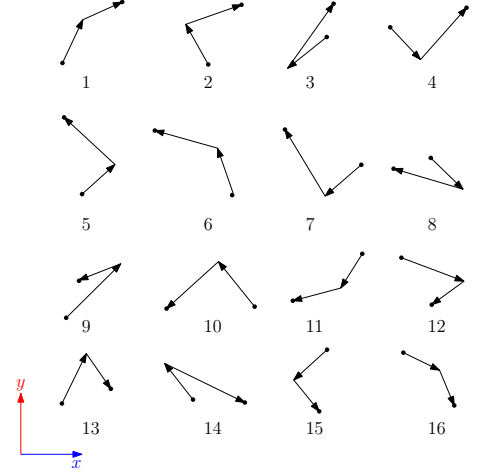I1. $\overrightarrow{p'_z p_i} \in W$

I2. $\Pi'_z \subset \mathcal{G}_w$ and

Figure 5: There are 16 cases for the combination of two directed segments.

I3. $|\Pi'_z|_w + |\overrightarrow{p'_z p_i}|_w \leq |\Pi_i|_w$.

We prove this by induction on i, the index of the vertices of Π (and index of the **for**-loop in Algorithm 1). The base case is $i = 2$. In this case, p'_z is equal to $p'_1 = s$. Clearly, $\overrightarrow{p'_1 p_2} \in W$, $\Pi'_1 \subset \mathcal{G}_w$ and $|\Pi'_1|_w + |\overrightarrow{p'_1 p_2}|_w = |\Pi_2|_w$, because $\overrightarrow{p'_1 p_2} = \overrightarrow{p_1 p_2}$. The induction hypothesis is that the invariant holds for all loop iterations before the i-th iteration of the **for**-loop. In the following, it is proved that it also holds before $(i + 1)$-th iteration of the **for**-loop.

In each iteration of the **for**-loop in Algorithm 1, we distinguish between the two cases: a) $\overrightarrow{p'_z p_{i+1}} \in W$, b) $\overrightarrow{p'_z p_{i+1}} \notin W$.

Case a) In Case a, p_i is skipped and Π' thus remains unchanged. Therefore, I1 and I2 hold, due to the induction hypothesis. In addition, since $\overrightarrow{p'_z p_{i+1}}$ is a segment in W and thus trivially xy-monotone, by Lemma 1, $|\overrightarrow{p'_z p_{i+1}}|_w \leq |\overrightarrow{p'_z p_i}|_w + |\overrightarrow{p_i p_{i+1}}|_w$. By induction hypothesis, we have $|\Pi'_z|_w + |\overrightarrow{p'_z p_i}|_w \leq |\Pi_i|_w$. By adding $|\overrightarrow{p_i p_{i+1}}|_w$ to the both sides of the inequality, we obtain $|\Pi'_z|_w + |\overrightarrow{p'_z p_{i+1}}|_w \leq |\Pi_i|_w + |\overrightarrow{p_i p_{i+1}}|_w = |\Pi_{i+1}|_w$. Therefore, I3 remains true after i-th iteration (i.e., before $(i + 1)$-th iteration).

Case b) In Case b, the **then** part of the **if** statement of the algorithm is entered and the visibility chain from p'_z to p_{i+1} is constructed. It is denoted by $CC_{p'_z}^{p_{i+1}}$: $\langle p'_z, q_1, \ldots, q_c, p_{i+1} \rangle \in W$, where $q_j \in V$, $j = 1, \ldots, c$. Then, the q_j, from $j = 1$ to $j = c$, is appended to the tail of Π'. Finally, p'_z is updated to q_c. In the remaining, it is proved that the invariant holds.

Since $CC_{p'_z}^{p_{i+1}}$ is the visibility chain, it is easy to see that all q_i, $i = 1, \ldots, last$, are represented by a node in the graph, $\mathcal{G}_w$, because they are vertices of W. Therefore, I1 and I2 hold. In order to check if I3 holds, we

need to analyze different cases. Each directed segment in W is of one of the following types: 1. xy-increasing 2. x-increasing and y-decreasing 3. y-increasing and x-decreasing 4. xy-decreasing. Therefore, there are 16 cases for the combination of two segments, $\overrightarrow{p'_z p_i}$ and $\overrightarrow{p_i p_{i+1}}$ (Figure 5). In all 16 cases, the orthogonal projection of $CC^{p_{i+1}}_{p'_z}$ onto the x-axis (resp. y-axis) is not longer than the sum of the orthogonal projections of $\overrightarrow{p'_z p_i}$ and $\overrightarrow{p_i p_{i+1}}$ onto the x-axis (resp. y-axis). Therefore, $|CC^{p_{i+1}}_{p'_z}|_w \leq |\overrightarrow{p'_z p_i}|_w + |\overrightarrow{p_i p_{i+1}}|_w$. Here we only show the proofs for two cases of Figure 5 as the proofs for the other cases are analogous.

Consider the case when both $\overrightarrow{p'_z p_i}$ and $\overrightarrow{p_i p_{i+1}}$ are y-increasing and x-decreasing (see Case 6 in Figure 5). In this case, since $\triangle p'_z p_i p_{i+1}$ lies in $R(p'_z, p_{i+1})$, by Lemma 2, $|CC^{p_{i+1}}_{p'_z}|_w = |\overrightarrow{p'_z p_i}|_w + |\overrightarrow{p_i p_{i+1}}|_w$. By inductive hypothesis we have $|\Pi'_z|_w + |\overrightarrow{p'_z p_i}|_w \leq |\Pi_i|_w$. Add now $|\overrightarrow{p_i p_{i+1}}|_w$ to both sides of the inequality. We obtain $|\Pi'_z|_w + |CC^{p_{i+1}}_{p'_z}|_w \leq |\Pi_i|_w + |\overrightarrow{p_i p_{i+1}}|_w$. It follows that $|\Pi'_{z+c}|_w + |\overrightarrow{q_c p_{i+1}}|_w \leq |\Pi_{i+1}|_w$, where q_c is the latest inserted vertex to the tail of Π' and Π'_{z+c} is the sub-path of Π' from index 1 to index $z + c$. Therefore, I3 holds. The proofs for cases 1,11 and 16 are similar.

Now consider Case 9, when $\overrightarrow{p'_z p_i}$ is xy-increasing and $\overrightarrow{p_i p_{i+1}}$ decreases in both x and y axes (illustrated in Figure 6). In this case, $|\overrightarrow{p'_z p_i}|_w + |\overrightarrow{p_i p_{i+1}}|_w = 0 + |\overrightarrow{p_i p_{i+1}}|_w$.

The vertical line that passes through p_{i+1} is denoted by $L^{\perp}_x$. The horizontal line that passes through p_{i+1} is denoted by $L^{\perp}_y$. Suppose these lines are directed toward $+\infty$. The following two properties hold. First, any directed segment of $CC^{p_{i+1}}_{p'_z}$ that lies on the left of $L^{\perp}_x$ is increasing in x. Therefore, they have a weighted length zero in the x-dimension. Second, any directed segment of $CC^{p_{i+1}}_{p'_z}$ that lies below $L^{\perp}_y$ is increasing in y. Therefore, they have a weighted length of zero in the y-dimension. By these two properties, any directed segment of $CC^{p_{i+1}}_{p'_z} : \langle p'_z, q_1, \ldots, q_c, p_{i+1} \rangle$ that lies on the left of $L^{\perp}_x$ and below $L^{\perp}_y$ is xy-increasing and has the weighted length zero.

Suppose $\overrightarrow{q_r q_{r+1}}$ is the first line segment in $CC^{p_{i+1}}_{p'_z}$ on the right side of $L^{\perp}_x$ that is x-decreasing. Since $CC^{p_{i+1}}_{p'_z}$ is a convex chain and is inside the triangle $\triangle p'_z p_i p_{i+1}$, the sub-chain $\langle q_r, \ldots, q_c, p_{i+1} \rangle$ is x-monotone and its weighted length in x-dimension is less than or equal to the weighted length of $\overrightarrow{p_i p_{i+1}}$ in x-dimension. Therefore, the weighted length of $CC^{p_{i+1}}_{p'_z}$ in x-dimension is less than or equal to the weighted length of $\overrightarrow{p_i p_{i+1}}$ in x-dimension.

It is analogous for the y-dimension. Suppose $\overrightarrow{q_u q_{u+1}}$ is the first line segment in $CC^{p_{i+1}}_{p'_z}$ above $L^{\perp}_y$ that is y-decreasing. Since $CC^{p_{i+1}}_{p'_z}$ is a convex chain and is inside the triangle $\triangle p'_z p_i p_{i+1}$, the sub-chain $\langle q_u, \ldots, q_c, p_{i+1} \rangle$

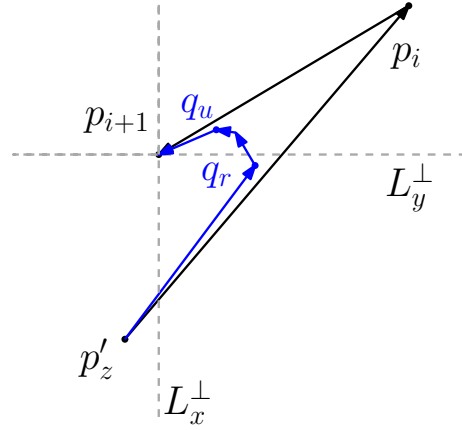Figure 6: *The segment from q_r to q_{r+1} is the first line segment in $CC^{p_{i+1}}_{p'_z}$ on the right side of $L^{\perp}_x$ that is x-decreasing. The segment from q_u to q_{u+1} is the first line segment in $CC^{p_{i+1}}_{p'_z}$ above of $L^{\perp}_y$ that is y-decreasing.*

is y-monotone and its weighted length in y-dimension is less than or equal to the weighted length of $\overrightarrow{p_i p_{i+1}}$ in y-dimension. Therefore, the weighted length of $CC^{p_{i+1}}_{p'_z}$ in y-dimension is less than or equal to the weighted length of $\overrightarrow{p_i p_{i+1}}$ in y-dimension.

To conclude, the weighted length of $CC^{p_{i+1}}_{p'_z}$, which is the sum of the weighted length of $CC^{p_{i+1}}_{p'_z}$ in x- and y-dimensions, is less than or equal to the weighted length of $\overrightarrow{p_i p_{i+1}}$, which is the sum of the weighted length of $\overrightarrow{p_i p_{i+1}}$ in x- and y-dimensions. Thus, $|CC^{p_{i+1}}_{p'_z}|_w \leq |\overrightarrow{p_i p_{i+1}}|_w = |\overrightarrow{p'_z p_i}|_w + |\overrightarrow{p_i p_{i+1}}|_w$. By inductive hypothesis, $|\Pi'_z|_w + |\overrightarrow{p'_z p_i}|_w \leq |\Pi_i|_w$. By adding $|\overrightarrow{p_i p_{i+1}}|_w$ to the both sides of the inequality, we conclude $|\Pi'_z|_w + |CC^{p_{i+1}}_{p'_z}|_w \leq |\Pi'_z|_w + |\overrightarrow{p'_z p_i}|_w + |\overrightarrow{p_i p_{i+1}}|_w \leq |\Pi_i|_w + |\overrightarrow{p_i p_{i+1}}|_w$. It follows that $|\Pi'_{z+c}|_w + |\overrightarrow{q_c p_{i+1}}|_w \leq |\Pi_{i+1}|_w$. Therefore, I3 also holds for this case. The proofs for the other remaining cases are similar. $\square$