Local Fréchet Permutation

Jonathan James Perry* Benjamin Raichel*

Abstract

In this paper we consider computing the Fréchet distance between two curves where we are allowed to locally permute the vertices. Specifically, we limit each vertex to move at most k positions from where it started, and give fixed parameter tractable algorithms in this parameter k, whose running times match the standard Fréchet distance computation running time when k is a constant. Furthermore we also show that computing such a local permutation Fréchet distance is NP-hard[1] when considering the weak Fréchet distance.

1 Introduction

A polygonal curve in $\mathbb{R}^d$ is defined by linearly interpolating an ordered sequence of points. In this paper we consider the well studied topic of polygonal curve similarity, as measured by the standard Fréchet distance, but where we are allowed to permute the ordered sequence of vertices defining each curve. Specifically, we seek to determine if there are permutations such that the Fréchet distance of the resulting curves is at most δ, where $\delta > 0$ is some given distance parameter. Here we will limit each permutation to only allow for local reordering of the points. Namely, we consider *k-permutations*, where a point at position i before permutation must end up at a position j after permutation such that $|i - j| \leq k$. Limiting to the case when k is small, intuitively means the orderings before and after the permutation are close, and thus in some sense the resulting curves as well. In particular, it can model scenarios where the input curve has some local corruptions or where local faults in the ordering occurred when collecting the data. Critically, limiting k allows us to achieve efficient algorithms for the (strong) Fréchet distance, whereas conversely it appears to aid our proof of hardness for weak Fréchet.

Prior Work.
Several previous papers ([1, 2, 6, 9, 10]) have considered a variant of the Fréchet distance sometimes referred to as the Curve/Point Set Matching problem (CPSM). Here one is given a curve π (i.e. an ordered sequence of points) of length n and an unordered point set P of size m, and asked whether a subset of P can be selected and ordered such that its Fréchet distance to π is at most some given threshold δ. The results in these papers vary based on whether (i) discrete or continuous Fréchet distance is used, (ii) a proper subset or all points of P is used (subset vs. all-points), and (iii) points can be repeated or not (non-unique vs. unique).

We briefly review prior results on CPSM. [9] gave an $O(nm^2)$ time algorithm for the continuous, subset, non-unique variant. Subsequently, [10] considered the discrete version, showing NP-hardness for both the subset and all-points unique cases, and giving polynomial time algorithms for both the subset and all-points non-unique cases. The papers [1, 2] then filled out the remaining continuous cases showing that both unique and non-unique all-points as well as the unique subset variant are NP-hard. Finally, [6] showed that if both curves are point sets, then the all-points unique discrete Fréchet distance problem can be solved in $O((m + n)\log(mn))$ time, but is NP-hard in the continuous case.

Related to our motivation of having errant data, other methods have been proposed for fixing curves so as to minimize Fréchet distance. There are many prior works in this direction, though they are perhaps further away from our permutation problem than the CPSM problem. Here we mention only [5] and [7], as we will utilize their techniques as discussed below (for other related works see references in [5, 7]). [5] considered the strong and weak Fréchet distance where points on the curves are uncertain, meaning there is some given set of potential locations where the point may be realized. [7] considered the strong and weak Fréchet distance when deletions or insertions are allowed on one or both curves.

Our Results.
We introduce and study the k, ℓ-permutation Fréchet distance, where one is allowed to k-permute the first curve and ℓ-permute the second curve, with the goal of making the Fréchet distance of the resulting curves below some threshold δ. For both the continuous and discrete variants, we provide fixed parameter tractable algorithms in terms of k and ℓ. Our running times match the corresponding quadratic running times for the standard Fréchet distance algorithms when k and ℓ are constants. Namely, for curves of lengths n and m, and for both the discrete and continuous cases, we

*Department of Computer Science, University of Texas at Dallas, USA {jperry,benjamin.raichel}@utdallas.edu.
Work on this paper was partially supported by NSF CAREER Award 1750780 and CCF Award 2311179.

[1]As it is trivially in NP, NP-hardness implies NP-completeness, though for consistency throughout we use NP-hardness.

give a running time of $O(nmk^2 4^k)$ when one curve is permuted and running time of $O(nmk^2 4^k \ell^2 4^\ell)$ if both curves are permuted.

Observe that the unrestricted permutation case, i.e. when $k = n$ and $\ell = m$, has already been studied as this is equivalent to the all-points unique CPSM problem discussed above. The prior results on unrestricted permutations can then be compared to our results on restricted permutations, as seen in Table 1. This connection to prior work implies our restriction to k-permutations is required in order to achieve polynomial time algorithms for many of the cases. Furthermore, for our restricted setting, if we set $k = \ell = 0$, then our problem is equivalent to the decision version of computing the Fréchet distance, for which [3] proved that for both the discrete or continuous version there is no strongly subquadratic algorithm unless the Strong Exponential Time Hypothesis (SETH) fails. (In fact it was shown that assuming SETH no constant factor approximation exists, and the constant was subsequently improved in [4].) Thus our $O(mn)$ running time (i.e. $O(n^2)$ when $n = m$) for constant k and ℓ is essentially tight up to lower order factors assuming SETH. In particular, while [6] achieve a near linear running time for the special case when unrestricted permutations are allowed on both curves for discrete Fréchet distance, such a result is not possible in our restricted permutation setting assuming SETH.

Unrestricted Prior Results

	One Curve	Both Curves
Discrete	NP-hard [10]	$O((m+n)\log(mn))$ [6]
Continuous	NP-hard [2]	NP-hard [6]

Our Restricted Results : Theorem 5

	One Curve	Both Curves
Discrete	$O(nmk^2 4^k)$	$O(nmk^2 4^k \ell^2 4^\ell)$
Continuous	$O(nmk^2 4^k)$	$O(nmk^2 4^k \ell^2 4^\ell)$

Table 1: Comparison of our restricted k, ℓ-permutation results with prior unrestricted permutation results. The one curve results are obtained by dropping ℓ terms.

Additionally, we provide algorithmic results for these problems when the goal is minimization rather than decision, either when minimizing k when δ is fixed, or minimizing δ when k is fixed. Specifically, to find the minimal k, which we call κ, for both continuous and discrete Fréchet distance, we give a running time of $O(nm\kappa^2 4^\kappa)$ when one curve is permuted and $O(nm\kappa^4 16^\kappa)$ when both curves are permuted (in which case we require both curves to be restricted to κ-permutations). When minimizing δ for discrete Fréchet distance we provide a running time of $O(nmk^2 4^k)$ for one curve and $O(nmk^2 4^k \ell^2 4^\ell)$ for two curves. When minimizing δ for continuous Fréchet distance, let $n = \max\{|\pi|, |\sigma|\}$ and let $\tau = \max\{k, \ell\}$. We then provide a running time of $O(n^2\tau^4 16^\tau(\log n + \tau))$ regardless of whether one or both curves are permuted, where the time bound holds with probability at least $1 - 1/n^c$, for any constant $c > 0$. These results are summarized in Table 2.

Min-K : Corollary 6

	One Curve	Both Curves
Disc./Cont.	$O(nm\kappa^2 4^\kappa)$	$O(nm\kappa^4 16^\kappa)$

Min-δ : Corollary 9 (Disc.) : Corollary 8 (Cont.)

	One Curve	Both Curves
Disc.	$O(nmk^2 4^k)$	$O(nmk^2 4^k \ell^2 4^\ell)$
Cont.	$O(n^2\tau^4 16^\tau(\log n + \tau))$	$O(n^2\tau^4 16^\tau(\log n + \tau))$

Table 2: Minimization results, where κ is the minimum k, $n = \max\{|\pi|, |\sigma|\}$, and $\tau = \max\{k, \ell\}$.

Finally, we consider the four decision problems but using weak Fréchet distance, which allows one to backtrack while traversing the curves, unlike the standard strong Fréchet considered in our algorithmic results. For the weak discrete Fréchet distance we reduce from 3SAT to show that permuting one curve or both curves is NP-hard even in $\mathbb{R}^1$. This reduction carries over to weak continuous Fréchet distance when raised to $\mathbb{R}^2$. In comparison, the prior NP-hardness results in Table 1 are all done in $\mathbb{R}^2$.

Both our algorithmic and hardness results follow a similar approach to that used in [7] for Fréchet edit distance. Specifically, our algorithmic results solve the problem by modeling potential solutions using DAG complexes, introduced in [8] and further utilized and expanded in [7]. Our hardness results for the weak case are inspired by the reduction used in [5], which was also used in [7] to prove weak variants of the Fréchet edit distance problem are NP-hard.

2 Preliminaries

Throughout, given points $p, q \in \mathbb{R}^d$, $||p - q||$ denotes their Euclidean distance. Moreover, given two (closed) sets $P, Q \subseteq \mathbb{R}^d$, $||P - Q|| = \min_{p\in P, q\in Q} ||p - q||$ denotes their distance, where for a single point $x \in \mathbb{R}^d$ we write $||x - P|| = ||\{x\} - P||$. We use angled brackets to denote an ordered list $\langle x_1, \ldots, x_n\rangle$, and use $L_1 \circ L_2$ to denote the concatenation of ordered lists L_1 and L_2. We use $[n]$ to denote the set $\{1, \ldots, n\}$.

Fréchet Distance.
The following definitions are standard, but in particular here we state the definitions directly as given in [7]. A *polygonal curve* is a sequence of n points $\pi = \langle \pi_1, \ldots, \pi_n\rangle$ where $\pi_i \in \mathbb{R}^d$ for all i. Such a sequence induces a continuous mapping from $[1, n]$ to $\mathbb{R}^d$, which we also denote by π, such that for any integer $1 \le i < n$, the restriction of π to the interval $[i, i+1]$ is defined by $\pi(i + \alpha) = (1 - \alpha)\pi_i + \alpha\pi_{i+1}$ for any $\alpha \in [0, 1]$, i.e. a straight line segment. We will view π as both a discrete

point sequence and a continuous function interchangeably, and when it is clear from the context, we also may use π to denote the image $\pi([1,n])$. We use $\pi[i,j]$, for $i \leq j$, to denote the restriction of π to the interval $[i,j]$. Given a curve $\pi = \langle \pi_1, \ldots, \pi_n \rangle$, we write $|\pi| = n$ to denote its size.

A reparameterization for a curve π of length n is a continuous non-decreasing bijection $f : [0,1] \to [1,n]$ such that $f(0) = 1, f(1) = n$. Given reparameterizations f, g of an n length curve π and an m length curve σ, respectively, the *width* between f and g is defined as

$$width_{f,g}(\pi,\sigma) = \max_{\alpha \in [0,1]} ||\pi(f(\alpha)) - \sigma(g(\alpha))||.$$

The (standard, i.e. continuous and strong) *Fréchet distance* between π and σ is then

$$\mathsf{d}_{\mathcal{F}}(\pi,\sigma) = \inf_{f,g} width_{f,g}(\pi,\sigma).$$

where f, g range over all possible reparameterizations of π and σ. Informally, the Fréchet distance is often described as the shortest leash length needed for a man on one curve and a dog on the other to walk from their respective starting to ending points of the curves.

The *discrete Fréchet distance* is similar to the above defined Fréchet distance, except that we do not traverse the edges but rather discontinuously jump to adjacent vertices. Specifically, define a monotone correspondence as a sequence of index pairs $\langle (i_1, j_1), \ldots, (i_k, j_k) \rangle$ such that $(i_1, j_1) = (1,1)$, $(i_k, j_k) = (n, m)$, for any $1 \leq z \leq k$ we have $1 \leq i_z \leq n$ and $1 \leq j_z \leq m$, and for any $1 \leq z < k$ we have $(i_{z+1}, j_{z+1}) \in \{(i_z+1, j_z), (i_z, j_z+1), (i_z+1, j_z+1)\}$. Let C denote the set of all monotone correspondences, then the discrete Fréchet distance is $\mathsf{d}_{\mathcal{DF}}(\pi,\sigma) = \inf_{c \in C} \max_{(i,j) \in c} ||\pi_i - \sigma_j||$.

Both the Fréchet distance and the discrete Fréchet distance have a corresponding *weak* variant, which is defined analogously except that one is allowed to backtrack on the curves. Specifically, the *weak Fréchet distance*, denoted $\mathsf{d}^{\mathrm{w}}_{\mathcal{F}}(\pi,\sigma)$, is defined similarly to the standard Fréchet distance above, except that when defining the width f and g are no longer required to be non-decreasing bijections, but are still required to be continuous and have $f(0) = 1, g(0) = 1$ and $f(1) = n, g(1) = m$. Similarly, the *weak discrete Fréchet distance*, denoted $\mathsf{d}^{\mathrm{w}}_{\mathcal{DF}}(\pi,\sigma)$, is defined similarly to the discrete Fréchet distance above, except that we no longer require the correspondence to be monotone. Specifically, a (non-monotone) correspondence is a sequence of index pairs $\langle (i_1, j_1), \ldots, (i_k, j_k) \rangle$ such that $(i_1, j_1) = (1,1)$, $(i_k, j_k) = (n,m)$, for any $1 \leq z \leq k$ we have $1 \leq i_z \leq n$ and $1 \leq j_z \leq m$, and for any $1 \leq z < k$ we have $(i_{z+1}, j_{z+1}) \in \{(i_z \pm 1, j_z), (i_z, j_z \pm 1), (i_z \pm 1, j_z \pm 1)\}$.

Permutation Fréchet Distance.
Viewing a polygonal curve $\pi = \langle \pi_1, \ldots, \pi_n \rangle$ as a sequence of points, a permutation of π is a bijection $f : [n] \to [n]$, which induces a new curve $f(\pi) = \langle \pi_{f(1)}, \ldots, \pi_{f(n)} \rangle$. Given a permutation f of π, we refer to f as a *k-permutation* if $|f(i) - i| \leq k$ for all $1 \leq i \leq n$, and we let $\mathcal{P}_k(\pi)$ denote the set of all k-permutations of π.

The *k,ℓ-permutation Fréchet distance* is then

$$\mathsf{d}^{k,\ell}_{\mathcal{PF}}(\pi,\sigma) = \min_{f \in \mathcal{P}_k(\pi), g \in \mathcal{P}_\ell(\sigma)} \mathsf{d}_{\mathcal{F}}(f(\pi), g(\sigma))$$

Observe that $\mathcal{P}_0(\pi)$ consists only of the identity function. Thus if we wish to consider the problem where permutations are only allowed on π then we write $\mathsf{d}^{k}_{\mathcal{PF}}(\pi,\sigma) = \mathsf{d}^{k,0}_{\mathcal{PF}}(\pi,\sigma) = \min_{f \in \mathcal{P}_k(\pi)} \mathsf{d}_{\mathcal{F}}(f(\pi),\sigma)$. Moreover, observe that $\mathsf{d}_{\mathcal{F}}(\pi,\sigma) = \mathsf{d}^{0,0}_{\mathcal{PF}}(\pi,\sigma)$.

We now define several problems based on the above:

- In the min-k permutation Fréchet distance problem, denoted MinK-PF, for a given $\delta > 0$ we seek the smallest value k such that $\mathsf{d}^{k,k}_{\mathcal{PF}}(\pi,\sigma) \leq \delta$.
- In the min-δ permutation Fréchet distance problem, denoted Minδ-PF, for given k and ℓ we seek the smallest value δ such that $\mathsf{d}^{k,\ell}_{\mathcal{PF}}(\pi,\sigma) \leq \delta$.
- For the MinK-PF and Minδ-PF problems if we instead ask if $\mathsf{d}^{k}_{\mathcal{PF}}(\pi,\sigma) = \mathsf{d}^{k,0}_{\mathcal{PF}}(\pi,\sigma) \leq \delta$, then we respectively refer to it as the *one-sided* MinK-PF or Minδ-PF problem.

All of the above definitions and problems immediately extended to the discrete, weak, or discrete weak Fréchet distance by replacing $\mathsf{d}_{\mathcal{F}}(f(\pi), g(\sigma))$ respectively with $\mathsf{d}_{\mathcal{DF}}(f(\pi), g(\sigma))$, $\mathsf{d}^{\mathrm{w}}_{\mathcal{F}}(f(\pi), g(\sigma))$, or $\mathsf{d}^{\mathrm{w}}_{\mathcal{DF}}(f(\pi), g(\sigma))$, in the definition of k,ℓ-permutation Fréchet distance.

DAG Complexes.
We will utilize the work of [8] and [7], the first of which defines the following generalization of a curve. Consider a directed acyclic graph (DAG) with vertices in $\mathbb{R}^d$, where a directed edge $\mathsf{p} \to \mathsf{q}$ is realized by the directed segment pq. We refer to such an embedded graph as being a *DAG complex*, denoted $\mathcal{C}$, with embedded vertices $V(\mathcal{C})$ (i.e. points) and embedded edges $E(\mathcal{C})$ (i.e. line segments). We denote the size of the complex as $|\mathcal{C}| = |E(\mathcal{C})| + |V(\mathcal{C})|$. Note that a DAG complex is allowed to have crossing edges and overlapping vertices. Call a polygonal curve $\pi = \langle \pi_1, \ldots, \pi_k \rangle$ *compliant* with $\mathcal{C}$ if $\pi_i \in V(\mathcal{C})$ for all i and $\pi_i \pi_{i+1} \in E(\mathcal{C})$ for all $1 \leq i < k$. (Note this implies π traverses each edge in the direction compliant with its orientation from the DAG.)

The following theorem was given in [7], who observed that the original theorem from [8] easily generalizes to the case where one allows sets of points for the start and end rather than individual points.

Theorem 1 *Given two* DAG *complexes $\mathcal{C}_1$ and $\mathcal{C}_2$, initial vertices $S_1 \subseteq V(\mathcal{C}_1)$ and $S_2 \subseteq V(\mathcal{C}_2)$, target vertices $T_1 \subseteq V(\mathcal{C}_1)$ and $T_2 \subseteq V(\mathcal{C}_2)$, and a value δ, then in $O(|\mathcal{C}_1||\mathcal{C}_2|)$ time one can determine the set of all pairs $t_1 \in T_1$ and $t_2 \in T_2$, such that there are curves π_1 and π_2 such that*

- *π_i is compliant with $\mathcal{C}_i$ for $i = 1, 2$.*
- *π_i starts at some $s_i \in S_i$ and ends at t_i, for $i = 1, 2$.*
- *$\mathsf{d}_{\mathcal{F}}(\pi_1, \pi_2) \leq \delta$.*

The standard Fréchet distance decision problem is typically computed by considering the product complex of two curves (i.e. a grid), and propagating the the reachable space according to a topological ordering of the cells in this product. The above theorem is thus obtained by observing that when the input consists of DAG complexes (which generalize curves), then the same approach works as the product complex still consists of cells with a topological order.

For the discrete Fréchet distance between two curves, one can again propagate reachability through the product, except the product is no longer a continuous space but rather simply a discrete grid graph. Thus again we can generalize to the case when the input is a pair of DAG's. Specifically, given graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, their product is $G_1 \times G_2 = (V_1 \times V_2, F)$, where $(u_1, u_2) \to (w_1, w_2) \in F$ if and only if (i) $u_1 = w_1$ and $u_2 \to w_2 \in E_2$, (ii) $u_2 = w_2$ and $u_1 \to w_1 \in E_1$, or (iii) $u_1 \to w_1 \in E_1$ and $u_2 \to w_2 \in E_2$. Observe, that since $G_1 \times G_2$ is also a DAG, there is thus again a topological ordering and so we immediately have the following corollary.

Corollary 2 *Given* DAG*'s $G_1 = (V_1, E_2)$ and $G_2 = (V_2, E_2)$, with vertices in $\mathbb{R}^d$, initial vertices $S_1 \subseteq V_1$ and $S_2 \subseteq V_2$, target vertices $T_1 \subseteq V_1$ and $T_2 \subseteq V_2$, and a value δ, then in $O(|G_1||G_2|)$ time one can determine the set of all pairs $t_1 \in T_1$ and $t_2 \in T_2$, such that there are paths π_1 in G_1 and π_2 in G_2 such that*

- *π_i starts at some $s_i \in S_i$ and ends at t_i, for $i = 1, 2$.*
- *$\mathsf{d}_{\mathcal{DF}}(\pi_1, \pi_2) \leq \delta$.*

3 Permutation Fréchet Distance

Given a polygonal curve $\pi = \langle \pi_1, \ldots, \pi_n \rangle$ and a parameter k, our goal is to construct a DAG complex where the set of compliant paths between specified start and end vertices is the same as the set of all k-permutations of π. To this end, consider constructing an arbitrary k-permutation of π, denoted $f(\pi)$, one vertex at a time. Observe that since f is a k-permutation, there are at most $2k + 1$ possible candidates for the vertex $\pi_{f(i)}$ for any i, namely the set $\{\pi_j \mid |j - i| \leq k\}$. Viewing these candidates as an ordered set $\langle \pi_{i-k}, \ldots, \pi_{i+k} \rangle$, any subset can be represented by a binary vector $v = \langle v_1, \ldots, v_{2k+1} \rangle \in \{0, 1\}^{2k+1}$, where $v_j = 1$ represents that $\pi_{i-k+(j-1)}$ is in the subset and $v_j = 0$ represent that it is not. In particular, when considering the possibilities for $\pi_{f(i)}$ we wish to restrict to the subset of $\langle \pi_{i-k}, \ldots, \pi_{i+k} \rangle$ which did not occur already in $\langle \pi_{f(1)}, \ldots, \pi_{f(i-1)} \rangle$. Thus rather than remembering this entire prior sequence, it suffices to pass a single $2k + 1$ length binary vector representing the subset of $\langle \pi_{i-k}, \ldots, \pi_{i+k} \rangle$ that has occurred already.

The construction of our DAG complex $\mathcal{C}$ is thus as follows. The vertices of $\mathcal{C}$ are copies of the vertices from π. Specifically, for each original vertex π_i we create a copy $\pi_i^{j,v}$ (i.e. π_i and $\pi_i^{j,v}$ have the same location) where j represents that π_i is the jth vertex in the permutation f, and $v \in \{0, 1\}^{2k+1}$ represents the subset of the $(2k + 1)$ possible vertices for $\pi_{f(j+1)}$ that have already occurred, as described above. Thus we have $V(\mathcal{C}) = \{\pi_i^{j,v} \mid 1 \leq i \leq n, |j - i| \leq k, v \in \{0, 1\}^{2k+1}\}$, and observe that $|V(\mathcal{C})| = O(nk2^{2k}) = O(nk4^k)$.

For the edge set $E(\mathcal{C})$, consider some vertex $\pi_i^{j,v}$. We add an edge from $\pi_i^{j,v}$ to $\pi_z^{j+1,w}$ if and only if $|j + 1 - z| \leq k$, π_z did not occur already (i.e. was not represented by a 1 in v), and w is consistent with v. In order for w to be consistent with v, w must represent the subset of the $(2k + 1)$ possible vertices for $\pi_{f(j+2)}$ that have already occurred, given that v represented the subset of the $(2k + 1)$ possible vertices for $\pi_{f(j+1)}$ that have already occurred. Thus $w = \langle w_1, \ldots, w_{2k+1} \rangle$ being consistent with $v = \langle v_1, \ldots, v_{2k+1} \rangle$ means that $w_{2k+1} = 0$ and $w_i = v_{i+1}$ for all $i < 2k + 1$, with the exception that the entry in w for π_z must be set to 1. Thus the degree of vertex $\pi_i^{j,v}$ is $O(k)$ and so $|\mathcal{C}| = O(|V(\mathcal{C})| + |E(\mathcal{C})|) = O(k \cdot nk4^k) = O(nk^2 4^k)$.

The last thing we must define is the allowed starting vertices S and ending vertices T in the DAG complex. Naturally, S consists of all vertices of the form $\pi_i^{1,v}$ where v is the all 0 vector except for the position representing π_i being 1, as such a vertex represents that π_i will be the first vertex in the permutation and it thus is the only vertex excluded from possibilities for $\pi_{f(2)}$. Similarly, T consists of all vertices of the form $\pi_i^{n,v}$, where here we can allow v to be any binary vector (since if the bit setting is invalid, it will not be reachable from a vertex in S).

As the above described complex $\mathcal{C}$ can be constructed in linear time in its size, we have the following.

Lemma 3 *Given a polygonal curve $\pi = \langle \pi_1, \ldots, \pi_n \rangle$ and a parameter k, in $O(nk^2 4^k)$ time one can construct a* DAG *complex $\mathcal{C}$ of size $O(nk^2 4^k)$ with vertex subsets S and T, such that the set of compliant paths in $\mathcal{C}$ that start at a vertex in S and end at a vertex in T is the same as the set of all k-permutations of π.*

It is easy to see that the above immediately applies to discrete Fréchet distance, by simply constructing the corresponding DAG's for the curves, rather than the DAG complexes. Thus we have the following.

Corollary 4 *Given a curve $\pi = \langle \pi_1, \ldots, \pi_n \rangle$ and a parameter k, in $O(nk^2 4^k)$ time one can construct a DAG G of size $O(nk^2 4^k)$ with vertex subsets S and T, such that the set of paths in G that start at a vertex in S and end at a vertex in T is the same as the set of all k-permutations of π.*

3.1 Algorithms and Results

Given curves π and σ of lengths n and m, respectively, along with integer parameters k and ℓ, and a value $\delta > 0$, Lemma 3 can be used to determine if $\mathsf{d}^{k,\ell}_{\mathcal{PF}}(\pi, \sigma) \leq \delta$. First, using Lemma 3, for the curve π and the parameter k we build a complex $\mathcal{C}_\pi$ along with sets S_π and T_π. Similarly for the curve σ and the parameter ℓ we build a complex $\mathcal{C}_\sigma$ with sets S_σ and T_σ. By definition, $\mathsf{d}^{k,\ell}_{\mathcal{PF}}(\pi, \sigma) \leq \delta$ if there exists $f \in \mathcal{P}_k(\pi)$ and $g \in \mathcal{P}_\ell(\sigma)$ such that $\mathsf{d}_{\mathcal{F}}(f(\pi), g(\sigma)) \leq \delta$, which by Lemma 3 is true if and only if there is a compliant path in $\mathcal{C}_\pi$ that starts at a vertex in S_π and ends at a vertex in T_π along with a compliant path in $\mathcal{C}_\sigma$ that starts at a vertex in S_σ and ends at a vertex in T_σ, such that their Fréchet distance is at most δ. Thus by applying Theorem 1 we have the following. Note by instead using Corollary 4 and Corollary 2 we can also handle the discrete case.

Theorem 5 *Given curves π and σ of lengths n and m, respectively, along with integers k and ℓ, and a value $\delta > 0$, in $O(nmk^2 4^k \ell^2 4^\ell)$ time one can determine if $\mathsf{d}^{k,\ell}_{\mathcal{PF}}(\pi, \sigma) \leq \delta$, for either the discrete or continuous Fréchet distance. The one sided problem can be solved in $O(nmk^2 4^k)$ by setting if ℓ to 0.*

Recall the MinK-PF problem defined in Section 2, where our goal is to find the minimum value k such that $\mathsf{d}^{k,k}_{\mathcal{PF}}(\pi, \sigma) \leq \delta$. Note that k cannot exceed $\max\{n, m\}$. Thus using Theorem 5, we search for the smallest k such that $\mathsf{d}^{k,k}_{\mathcal{PF}}(\pi, \sigma) \leq \delta$. Observe that the running time in the theorem is exponential in k, and in particular the running time for $k+1$ is a constant factor larger than that for k. Thus if we search for the minimum value, by successively incrementing k by 1, the times of the calls will behave like an increasing geometric series, and thus the overall time will be proportional to the last call.

Corollary 6 *Let κ be the optimal value for the MinK-PF problem. Then the MinK-PF problem can be solved in $O(nm\kappa^4 16^\kappa)$ time, for either the discrete or continuous Fréchet distance. The one-side MinK-PF problem can be solved in $O(nm\kappa^2 4^\kappa)$ time.*

[8] showed how to turn the decision procedure of Theorem 1 into an optimization procedure, using a simple sampling based approach which avoids the more complicated parametric search technique typically used to compute the Fréchet distance. The algorithm is guaranteed to be correct, and while its running time is a random variable (i.e. it is a Las Vegas algorithm), with polynomially high probably it achieves an efficient running time.

We remark that technically the following theorem from [8] was originally stated with only a single start and end vertex in each complex, whereas we require allowing sets of start and end vertices, although it immediately generalizes to this case. Specifically, [8] searches over a set of critical values using a decision procedure. The decision procedure was already generalized in [7] to sets of start and end vertices, as stated above in Theorem 1. Moreover, the critical values remain exactly the same when generalizing to starting and ending vertex sets. (This holds as [8] considered all vertex to vertex pairs as critical events, not simply just the pair of starting vertices and pair of ending vertices.) Thus we have the following.

Theorem 7 ([8], Theorem 6.3) *Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two DAG complexes, of total complexity n, with start and end vertex sets $S_1, T_1 \subseteq V(\mathcal{C}_1)$, $S_2, T_2 \subseteq V(\mathcal{C}_2)$. Then there is an algorithm which computes two curves π_1 and π_2 such that π_1 (resp. π_2) is compliant with $\mathcal{C}_1$ (resp. $\mathcal{C}_2$), starts at a vertex in S_1 (resp. S_2), and ends at a vertex of T_1 (resp. T_2). Moreover, $\mathsf{d}_{\mathcal{F}}(\pi_1, \pi_2)$ is minimum among all such curves. The running time of the algorithm is $O(n^2 \log n)$ with probability at least $1 - 1/n^c$, for any constant $c > 0$.*

Corollary 8 *Let $n = \max\{|\pi|, |\sigma|\}$ and let $\tau = \max\{k, \ell\}$. Then both the one-sided and two-sided versions of the continuous Minδ-PF problem can be solved in $O(n^2 \tau^4 16^\tau (\log n + \tau))$ time, where the time bound holds with probability at least $1 - 1/n^c$, for any constant $c > 0$.*

It is well known that for the discrete Fréchet distance both the decision and optimization problem can be solved in $O(nm)$ time, since rather than propagating reachability in the free space we can propagate the minimum cost to reach the given vertex (which works as the free space is a discrete DAG). Again, in our case even though the input is now two DAG's rather than simply two curves, the product is still just a DAG. Thus we also can solve the optimization problem in the same time as our decision algorithm, yielding the following.

Corollary 9 *Given curves π and σ of lengths n and m, respectively, along with integers k and ℓ, one can solve the discrete Minδ-PF problem in $O(nmk^2 4^k \ell^2 4^\ell)$ time, which becomes $O(nmk^2 4^k)$ time for the one sided case.*

4 Hardness for Weak Permutation Fréchet

We now shift focus to *weak Fréchet distance* and prove NP-hardness for both the discrete and continuous cases $\mathsf{d}^{\mathrm{w}\ k,\ell}_{\mathcal{PDF}}(\pi,\sigma) \le \delta$ in $\mathbb{R}^1$ and $\mathsf{d}^{\mathrm{w}\ k,\ell}_{\mathcal{PF}}(\pi,\sigma) \le \delta$ in $\mathbb{R}^2$ for any constants $k \ge 1$ and $\ell \ge 0$. Our reductions are from 3SAT and closely follow those of [5] and [7]. We consider discrete curves in $\mathbb{R}^1$ first before moving to $\mathbb{R}^2$.

We set δ to 1 and rely on the resulting *Free Space Diagram* (ex. Figure 1), which is built by listing π-values as rows, σ-values as columns, and drawing empty circles where $|\pi_i - \sigma_j| \le \delta$ (free spaces). We use teal to show paths between free spaces, with movement restricted to adjacent spaces (including diagonals). With this, $\mathsf{d}^{\mathrm{w}}_{\mathcal{DF}}(\pi,\sigma) \le 1$ if and only if there is a way to traverse through the teal from the bottom left corner to the top right corner.

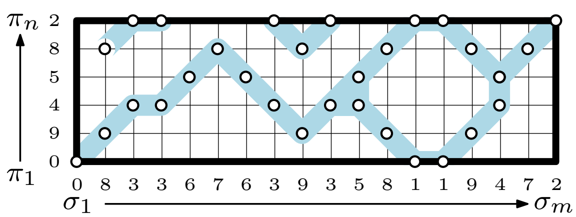

Figure 1: Generic Free Space Diagram.

We can now show the abstract Figure 2, where we embed the 3SAT instance with c clauses into σ with clause gadgets and walls such that $\sigma = \langle\text{wall}\rangle \circ \langle\text{clause } 1\rangle \circ \langle\text{wall}\rangle \circ \cdots \circ \langle\text{wall}\rangle \circ \langle\text{clause } c\rangle \circ \langle\text{wall}\rangle$. We construct π in three sections $\pi = \langle\text{lower}\rangle \circ \langle\text{variable layer}\rangle \circ \langle\text{upper}\rangle$. The interaction between the curves will force the length of π to be traversed for each clause, with 3 options representing satisfying one literal each. The clauses are placed in series, alternating going up or down, such that all clauses must have at least one literal satisfied in order to traverse to the end of σ.

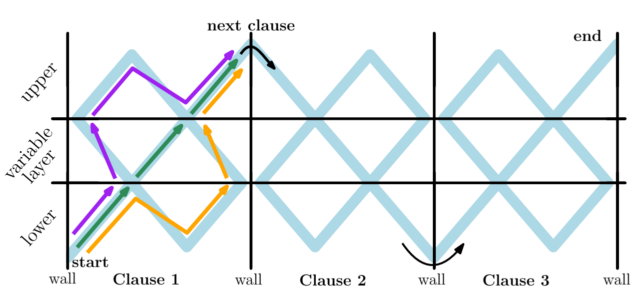

Figure 2: Abstract figure with sections of π and σ labeled. Purple, green, and orange illustrate the 3 ways to get through a clause.

The variable layer holds pairs of values where each pair represents a variable. Whether the pair is permuted or not determines if the variable is set to True or False. We call these permutations *sanctioned* (and all others *unsanctioned*) as these are the only ones that correspond to 3SAT actions. Further, our construction ensures only sanctioned permutations will be useful.

Precise Reduction.
Having seen the general idea, the following is the precise construction for π and σ given a 3SAT instance I with v variables and c clauses, an example of which is shown in Figure 3. Note that we will insert copies of a special point ρ and duplicate others, both of which serve only to discourage unsanctioned permutations.

Let $\rho = \infty$ in $\mathbb{R}^1$ for discrete and $\rho = (0,\infty)$ in $\mathbb{R}^2$ for continuous.[2] Note that when working in $\mathbb{R}^2$, all other points will be of the form $(x, 0)$ for some value x. Additionally let $S^{[x]}$ be the concatenation of x instances of ordered set S (e.g. $\langle 5,1\rangle^{[2]} = \langle 5,1\rangle \circ \langle 5,1\rangle$). For any ordered set S we define S^R as that set in reverse order.

.... **construct** π ....

- Let $\odot$ represent '$\circ\ \langle\rho\rangle^{[2k+1]}\ \circ$'.
- Let $L = \langle 15\rangle \odot \langle 25\rangle \odot \ldots \odot \langle 10v+5\rangle$.
- Let $\hat{L}$ be L but with the value $10j+5$ replaced by $\langle 10j+4, 10j+6\rangle$ for all $1 \le j \le v$.
- Let $\pi = \langle 0\rangle \odot \langle 10\rangle \odot L \odot \langle 10(v+1)\rangle \odot \hat{L}^R \odot \langle 10\rangle \odot L \odot \langle 10(v+1)\rangle \odot \langle 10(v+2)\rangle$.

.... **construct** σ ....

- Let $\circledcirc$ represent '$\circ\ \langle\rho\rangle^{[2\ell+1]}\ \circ$'.
- Let $L = \langle 15\rangle \circledcirc \langle 25\rangle \circledcirc \ldots \circledcirc \langle 10v+5\rangle$.
- Let L_j^+ be obtained from L by replacing the value $10j+5$ with $\langle 10j+4\rangle^{[\ell+1]} \circ \langle 10j+6\rangle^{[\ell+1]}$.
- Let L_j^- be obtained from L by replacing the value $10j+5$ with $\langle 10j+6\rangle^{[\ell+1]} \circ \langle 10j+4\rangle^{[\ell+1]}$.
- Let σ^i represent clause i of I which contains variables $\mathcal{X}_{j_1}$, $\mathcal{X}_{j_2}$, and $\mathcal{X}_{j_3}$ and therefore $\sigma^i = \langle 10\rangle \circledcirc L^{\pm}_{j_1} \circledcirc \langle 10(v+1)\rangle \circledcirc (L^{\pm}_{j_2})^R \circledcirc \langle 10\rangle \circledcirc L^{\pm}_{j_3} \circledcirc \langle 10(v+1)\rangle$, where $L^{\pm}_{j_i} = L^{+}_{j_i}$ if $\mathcal{X}_{j_i}$ appears as a positive literal and $L^{\pm}_{j_i} = L^{-}_{j_i}$ if $\mathcal{X}_{j_i}$ appears as a negated literal.
- Let $\sigma = \langle 0\rangle \circledcirc \sigma^1 \circledcirc \langle 10(v+2)\rangle \circledcirc (\sigma^2)^R \circledcirc \langle 0\rangle \circledcirc \ldots \circledcirc \sigma^c \circledcirc \langle 10(v+2)\rangle$ (if c is even, duplicate one clause so the total number of clauses is odd).

From this reduction comes the following two theorems which we spend the remainder of the paper proving.

Theorem 10 *Determining if* $\mathsf{d}^{\mathrm{w}\ k,\ell}_{\mathcal{PDF}}(\pi,\sigma) \le \delta$ *in* $\mathbb{R}^1$ *for any constants* $k \ge 1$ *and* $\ell \ge 0$ *is NP-hard.*

[2] ∞ can be replaced with a sufficiently large finite value far away from the other values as explained in [7].

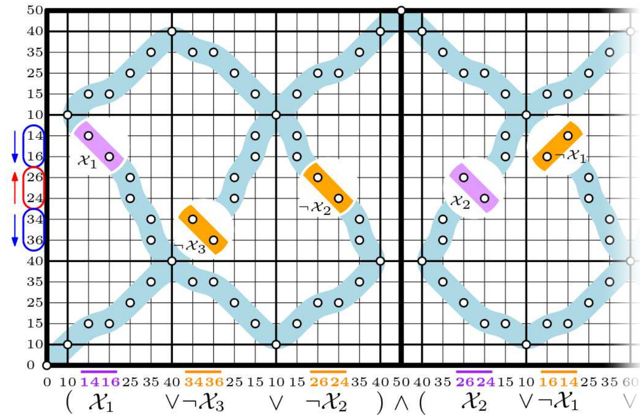

Figure 3: An example where $\mathcal{X}_2$ is set to False and other variables to True as seen in circled pairs on the left. ρ values are not shown, and $\ell = 0$. Figure faded out after second literal of second clause.

Theorem 11 *Determining if* $\mathsf{d}^{\mathrm{w}\,k,\ell}_{\mathcal{PF}}(\pi,\sigma) \le \delta$ *in* $\mathbb{R}^2$ *for any constants* $k \ge 1$ *and* $\ell \ge 0$ *is NP-hard .*

Proof of correctness for Theorem 10.
We now argue correctness. First, we argue that if only sanctioned permutations are made, the distance is $\le \delta$ if and only if I is satisfiable. Second, we argue that unsanctioned permutations are futile and will not result in the distance being $\le \delta$ unless it would have been possible without them.

So suppose that only sanctioned permutations are allowed. Consider the variable layer, where the jth variable $\mathcal{X}_j$ is represented by consecutive rows (i.e. a sanctioned permutation pair of vertices from π) with value $10j + 6$ followed by value $10j + 4$, which we can either leave in that order or permute (see Figure 3). We argue that leaving this pair unpermuted corresponds to setting the variable to True and permuting it corresponds to setting the variable to False. Specifically, consider the variable layer restricted to the columns of the ith clause, represented by σ^i. If $\mathcal{X}_j$ does not appear in this clause, then because the corresponding value $10j + 5$ in σ^i is within distance $\delta = 1$ of both $10j + 4$ and $10j + 6$, the three paths through the variable layer (again see Figure 2) will be unobstructed at the rows corresponding to $\mathcal{X}_j$, regardless of whether we permute this sanctioned permutation pair or not.

Conversely, if $\mathcal{X}_j$ is in the ith clause then rather than $10j + 5$ in σ^i we have the values $10j + 4$ and $10j + 6$ (possibly with repetition if $\ell > 0$). The portion of σ^i corresponding to $\mathcal{X}_j$ is either ascending or descending, depending on the parity of i and whether $\mathcal{X}_j$ is the first, second, or third variable in the clause. For simplicity, assume it is descending (the ascending case is symmetric). In this case, by construction, the pair of values $10j+4$ and $10j+6$ will appear in decreasing order if $\mathcal{X}_j$ occurs as a positive literal in clause i, and in increasing order if $\mathcal{X}_j$ occurs as a negated literal in clause i. So suppose $\mathcal{X}_j$ appears a positive literal, and consider the sanctioned permutation pair on π corresponding to $\mathcal{X}_j$. As the values in the variable layer are also descending (before any permutation occurs), in order to pass through the rows of this sanctioned permutation pair along the path corresponding to positive literal $\mathcal{X}_j$ in this clause, we *must not* permute the sanction pair (i.e. the rows and columns must both agree to descend in value for this pair). Conversely, if $\mathcal{X}_j$ appears as a negated literal then $10j + 4$ and $10j + 6$ will appear in increasing order on σ^i and so we *must* permute the sanctioned permutation pair to match if we want to pass along the corresponding path. As at least one of three paths through the variable layer for this clause must be passable, this corresponds to setting the variables in such a way that at least one of the literals in the clause is True, as required for the 3SAT instance I to be satisfiable. Therefore, if only sanctioned permutations are allowed, $\mathsf{d}^{\mathrm{w}\,k,\ell}_{\mathcal{PDF}}(\pi,\sigma) \le \delta$ if and only if I is satisfiable.

To complete the proof we now argue that even if unsanctioned permutations are allowed, they will not lower the Fréchet distance, and hence are futile. Specifically, by inserting a sufficiently large number of ρ points between adjacent pairs of non-ρ points, we insured that k,ℓ-permutations could not change the ordering of the those pairs. On π the only time ρ was not inserted was between points in pairs whose permutation was sanctioned. On σ the only time ρ was not inserted was between points in pairs that enforce a literal.[3] However, for each such pair $\langle x, y\rangle$ we duplicated x and y both $\ell+1$ times. This means regardless of any ℓ-permutation, the first copy of x must come before the first of y and the last copy of x before the last of y. This is sufficient since when the path of this literal is broken due to incorrect variable assignment, it is because the free space of both the first x and the last y become disconnected from the rest of the path.[4] If this disconnect exists before permutation, it will remain as the first x and last y remain. Finally, observe that a ρ point on one curve must map to a ρ point on the other curve, but the initial construction inserted ρ points between all pairs that were neither sanctioned on π nor representing a literal on σ, and we inserted enough such that there must remain at least one ρ between these pairs even after k,l-permutation. Therefore futility is established.

Proof of correctness for Theorem 11.
As was done in [5] and [7], we can extend the argument to the continuous Fréchet distance by raising to $\mathbb{R}^2$, us-

[3]Inserting ρ between these would have given π issues traversing its sanctioned permutation pairs

[4]Note that permutation does not change the free spaces, just their order. Thus we refer to incorrect assignment as 'disconnecting' the free spaces rather than making them 'no longer free'.

ing $\rho = (0, \infty)$, and converting $x \to (x, 0)$ for any other point x in the above reduction. We now argue this limits continuous to have Fréchet distance $\leq \delta$ if and only if discrete would have as well.

Consider a pair of traversals, one on each curve, such that their Fréchet distance is finite. Then on both curves, these traversals must simultaneously start at the first vertex, and go the next vertex (which is ρ) which is always possible in both continuous and discrete. The same is true for traversing between the second to last vertex (which is ρ) and the last on both curves. We break the remaining traversals of both curves into sub-traversals that start at a ρ, traverse some potentially different number of non-ρ vertex(es), and end at a ρ. Let Ξ_π (resp. Ξ_σ) be the non-ρ vertices traversed in an arbitrary one of these sub-traversals on π (resp. σ) where vertices may appear multiple times due to using weak Fréchet distance. Recall because the Fréchet distance is finite, that a ρ on the traversal of one curve must map to a ρ on the other. This means that a sub-traversal between ρ vertices on one curve must map to such a sub-traversal on the other, and one can argue that whether such sub-traversals are within Fréchet distance δ is equivalent to whether the corresponding Ξ_π and Ξ_σ are within Fréchet distance δ. (We can assume Ξ_π and Ξ_σ are non-empty, since if they were both empty we did not make any actual progress in traversing either curve. If Ξ_π was empty and $\Xi_\sigma = \langle x \dots \rangle$ then for the Fréchet distance to be $\leq \delta$, the sub-traversal of π must be approaching a vertex y where $||x - y|| \leq \delta$, and so one can argue y could have been included.)

The above implies it suffices to show that $\mathsf{d}_{\mathcal{DF}}(\Xi_\pi, \Xi_\sigma) \leq \delta \iff \mathsf{d}_{\mathcal{F}}(\Xi_\pi, \Xi_\sigma) \leq \delta$ for all Ξ_π and Ξ_σ. There are two types of Ξ_π: i) individual points, and ii) those including at least two vertices in a pair representing a variable assignment. Likewise there are two types of Ξ_σ: i) individual points, and ii) those including at least two vertices in a pair (duplicated if $\ell > 0$) which enforces a literal.

If either Ξ_π or Ξ_σ is a single point, then the discrete and continuous Fréchet distances are $\leq \delta$ iff. all points on the other curve are within δ of it. This shows the equivalence for all the cases except the distance between type ii of Ξ_π and type ii of Ξ_σ, so consider this case. Ξ_π cannot contain vertices from more than one variable, and the same is true for Ξ_σ w.r.t literals. Additionally, if they do not refer to the same variable/literal, or they refer to the same variable/literal but the variable is not set in such a way that the literal is satisfied, then the the Fréchet distance is $> \delta$ since the first vertices are too far from one another. Now observe that if both Ξ_π and Ξ_σ refer to the same variable/literal and the variable is set in such a way to satisfy the literal, then the Fréchet distance $\leq \delta$ if and only if they start and end with the same value, regardless of discrete or continuous.

References

[1] Paul Accisano and Alper Üngör. Approximate matching of curves to point sets. In *Proceedings of the 26th Canadian Conference on Computational Geometry (CCCG)*. Carleton University, Ottawa, Canada, 2014. URL: `http://www.cccg.ca/proceedings/2014/papers/paper65.pdf`.

[2] Paul Accisano and Alper Üngör. Hardness results on curve/point set matching with fréchet distance, 2012. `arXiv:1211.2030`.

[3] Karl Bringmann. Why walking the dog takes time: Frechet distance has no strongly subquadratic algorithms unless seth fails. In *Proceedings of the 2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, FOCS '14, page 661–670, USA, 2014. IEEE Computer Society. `doi:10.1109/FOCS.2014.76`.

[4] Karl Bringmann and Wolfgang Mulzer. Approximability of the discrete Fréchet distance. *JoCG*, 7(2):46–76, 2016.

[5] Kevin Buchin, Maarten Löffler, Tim Ophelders, Aleksandr Popov, Jérôme Urhausen, and Kevin Verbeek. Computing the Fréchet distance between uncertain curves in one dimension. *Comput. Geom.*, 109:101923, 2023. `doi:10.1016/j.comgeo.2022.101923`.

[6] Maike Buchin and Bernhard Kilgus. Fréchet distance between two point sets. *Computational Geometry*, 102:101842, 2022. URL: `https://www.sciencedirect.com/science/article/pii/S0925772121000985`, `doi:10.1016/j.comgeo.2021.101842`.

[7] Emily Fox, Amir Nayyeri, Jonathan James Perry, and Benjamin Raichel. Fréchet edit distance, 2024. `arXiv:2403.12878`.

[8] Sariel Har-Peled and Benjamin Raichel. The Fréchet distance revisited and extended. *ACM Trans. Algorithms*, 10(1):3:1–3:22, 2014. `doi:10.1145/2532646`.

[9] Anil Maheshwari, Jörg-Rüdiger Sack, Kaveh Shahbaz, and Hamid Zarrabi-Zadeh. Staying close to a curve. In *Proceedings of the 23rd Annual Canadian Conference on Computational Geometry (CCCG)*, 2011. URL: `http://www.cccg.ca/proceedings/2011/papers/paper97.pdf`.

[10] Tim Wylie and Binhai Zhu. Following a curve with the discrete fréchet distance. *Theoretical Computer Science*, 556:34–44, 2014. Combinatorial Optimization and Applications. URL: `https://www.sciencedirect.com/science/article/pii/S0304397514004629`, `doi:10.1016/j.tcs.2014.06.026`.