Uncertain Curve Simplification

Kevin Buchin
Department of Mathematics and Computer Science, TU Eindhoven, The Netherlands

Maarten Löffler
Department of Information and Computing Sciences, Utrecht University, The Netherlands

Aleksandr Popov
Department of Mathematics and Computer Science, TU Eindhoven, The Netherlands

Marcel Roeloffzen
Department of Mathematics and Computer Science, TU Eindhoven, The Netherlands

Abstract

We study the problem of polygonal curve simplification under uncertainty, where instead of a sequence of exact points, each uncertain point is represented by a region which contains the (unknown) true location of the vertex. The regions we consider are disks, line segments, convex polygons, and discrete sets of points. We are interested in finding the shortest subsequence of uncertain points such that no matter what the true location of each uncertain point is, the resulting polygonal curve is a valid simplification of the original polygonal curve under the Hausdorff or the Fréchet distance. For both these distance measures, we present polynomial-time algorithms for this problem.

2012 ACM Subject Classification Theory of computation → Computational geometry

Keywords and phrases Curves, Uncertainty, Simplification, Fréchet Distance, Hausdorff Distance

Digital Object Identifier 10.4230/LIPIcs.MFCS.2021.26

Related Version *Full Version*: https://arxiv.org/abs/2103.09223 [12]

Funding *Maarten Löffler*: Partially supported by the Dutch Research Council (NWO) under project no. 614.001.504.
Aleksandr Popov: Supported by the Dutch Research Council (NWO) under project no. 612.001.801.
Marcel Roeloffzen: Supported by the Dutch Research Council (NWO) under project no. 628.011.005.

1 Introduction

In this paper, we study the topic of curve simplification under uncertainty. There are many classical algorithms dealing with curve simplification with different distance metrics; however, it is typically assumed that the locations of points making up the curves are known precisely, which often does not suit real-life data. An example highlighting the necessity of taking uncertainty into account comes with GPS data, where each measured location is inherently imprecise, and the real location is likely to be within a certain distance from the measurement. This imprecision can be modelled as a disk (or some other shape if the GPS signals are blocked or reflected by rocks, buildings, etc.). Curve simplification is used to reduce the noise-to-signal ratio in the trajectory data before applying other algorithms or when storing large amounts of data. In both cases modelling uncertainty could reduce the error introduced by simplifying imprecise curves while maintaining a short, efficient representation of the data.

There is a large volume of foundational work on curve simplification [4], including work on vertex-constrained simplification, such as the algorithms by Ramer and by Douglas and Peucker [17, 34] using the Hausdorff distance, by Agarwal et al. [3] using the Fréchet distance, by Imai and Iri [23] using either, and various improvements and related approaches [7, 8, 10, 16, 21, 22, 31, 36]. The Imai–Iri algorithm involves computing the *shortcut graph,* which captures all the possible simplifications of a curve, and then finding a path through the graph

46th International Symposium on Mathematical Foundations of Computer Science (MFCS 2021).
Editors: Filippo Bonchi and Simon J. Puglisi; Article No. 26; pp. 26:1–26:22
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

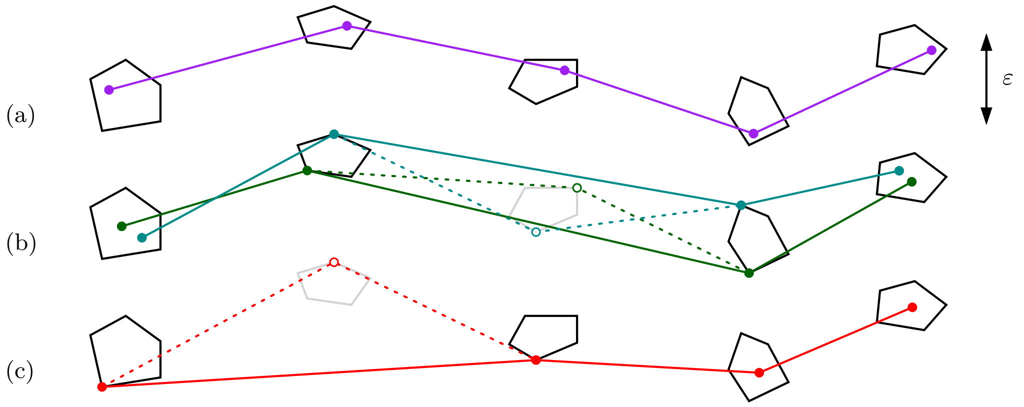

Figure 1 (a) An uncertain curve modelled with convex polygons and a realisation. (b) A valid simplification under the Hausdorff distance with the threshold ε: for every realisation, the subsequence is within distance ε from the full sequence. (c) An invalid simplification: there is a realisation for which the subsequence is not within distance ε from the full sequence.

with minimal edge count from the start to the end node, yielding the shortest simplification. We adapt this approach to the setting with uncertainty. It seems natural to apply disk stabbing to test shortcuts [22]; we discuss why this does not work in our setting in Section 3.

There are recent advances in the study of uncertainty in computational geometry, with work on optimising various measures on uncertain points [24, 25, 26, 28, 30], triangulations [11, 29, 37], visibility in uncertain polygons [15], and other problems [1, 2, 18, 19, 20, 27, 32, 35]. There is work by Ahn et al. [5], and, more recently, by Buchin et al. [9, 33] on various minimisation and maximisation variants of curve similarity with the Fréchet distance under uncertainty, and other work combining trajectory analysis and uncertainty [6, 13, 14]. To our knowledge, there is no previous work studying curve simplification under uncertainty.

We use the locational model for uncertainty: we know that each point exists, but not its exact location. It can be modelled as a discrete set of points, of which one is the true location; this model uses *indecisive* points. We also use *imprecise* points, modelled as compact continuous sets, such as disks, line segments, or convex polygons; the true location is one unknown point from the set. An *uncertain curve* is a sequence of uncertain points of the same kind. A *realisation* of an uncertain curve is a polygonal curve obtained by taking one point from each uncertain point. We solve the following problem (see Figure 1): *given an uncertain curve as a sequence of n uncertain points, find the shortest subsequence of the points such that for any realisation of the curve, the realisation restricted to the subsequence is a valid simplification.* We give efficient algorithms for this problem for the Hausdorff and the Fréchet distance. They run in $\mathcal{O}(n^3)$ time for uncertainty modelled with disks or line segments and in $\mathcal{O}(n^3k^3)$ time for convex polygons and indecisive points with k vertices.

2 Preliminaries

Denote[1] $[n] \stackrel{\text{def}}{=} \{1, 2, \ldots, n\}$ for any $n \in \mathbb{N}^{>0}$. Given two points $p, q \in \mathbb{R}^2$, denote their Euclidean distance with $\|p - q\|$.

[1] We use := and =: to denote assignment, $\stackrel{\text{def}}{=}$ for equivalent quantities in definitions or to point out equality by earlier definition, and = in other contexts. We also use $\equiv$, but its usage is always explained.

Denote a *sequence* of points in $\mathbb{R}^2$ with $\pi = \langle p_1, \dots, p_n \rangle$. For only two points $p, q \in \mathbb{R}^2$, we also write pq instead of $\langle p, q \rangle$. Denote a subsequence of a sequence π from index i to j with $\pi[i : j] = \langle p_i, p_{i+1}, \dots, p_j \rangle$. This notation can also be applied if we interpret π as a *polygonal curve* on n vertices (of *length* n). It is defined by linearly interpolating between the successive points in the sequence and can be seen as a continuous function, for $i \in [n-1]$ and $\alpha \in [0, 1]$: $\pi(i + \alpha) = (1 - \alpha)p_i + \alpha p_{i+1}$.

We also introduce the notation for the order of points along a curve. Let $p \coloneqq \pi(a)$ and $q \coloneqq \pi(b)$ for $a, b \in [1, n]$. Then $p \prec q$ iff $a < b$, $p \preccurlyeq q$ iff $a \leq b$, and $p \equiv q$ iff $a = b$. Note that we can have $p = q$ for $a \neq b$ if the curve intersects itself.

Finally, given points $p, q, r \in \mathbb{R}^2$, define the distance from p to the segment qr as $d(p, qr) \overset{\text{def}}{=} \min_{t \in qr} \|p - t\|$.

An *uncertainty region* $U \subset \mathbb{R}^2$ describes a possible location of a true point: it has to be inside the region, but there is no information as to where exactly. We use several uncertainty models, so the regions U are of different shape. An *indecisive point* is a form of an uncertain point where the uncertainty region is represented as a discrete set of points, and the true point is one of them: $U = \{p^1, \dots, p^k\}$, with $k \in \mathbb{N}^{>0}$ and $p^i \in \mathbb{R}^2$ for all $i \in [k]$. *Imprecise points* are modelled with uncertainty regions that are compact continuous sets. In particular, we consider *disks* and *polygonal closed convex sets.* We denote a disk with the centre $c \in \mathbb{R}^2$ and the radius $r \in \mathbb{R}^{\geq 0}$ as $D(c, r)$. Formally, $D(c, r) \overset{\text{def}}{=} \{p \in \mathbb{R}^2 \mid \|p - c\| \leq r\}$. Define a *polygonal closed convex set (PCCS)* as a closed convex set with bounded area that can be described as the intersection of a *finite* number of closed half-spaces. Note that this definition includes both convex polygons and line segments (in 2D). Given a PCCS U, let $V(U)$ denote the set of vertices of U, i.e. vertices of a convex polygon or endpoints of a line segment.

We call a sequence of uncertainty regions an *uncertain curve:* $\mathcal{U} = \langle U_1, \dots, U_n \rangle$. If we pick a point from each uncertainty region of $\mathcal{U}$, we get a polygonal curve π that we call a *realisation* of $\mathcal{U}$ and denote it with $\pi \Subset \mathcal{U}$. That is, if for some $n \in \mathbb{N}^{>0}$ we have $\pi = \langle p_1, \dots, p_n \rangle$ and $\mathcal{U} = \langle U_1, \dots, U_n \rangle$, then $\pi \Subset \mathcal{U}$ if and only if $p_i \in U_i$ for all $i \in [n]$.

Suppose we are given a polygonal curve $\pi = \langle p_1, \dots, p_n \rangle$, a threshold $\varepsilon \in \mathbb{R}^{>0}$, and a curve built on the subsequence of vertices of π for some set $I = \{i_1, \dots, i_\ell\} \subseteq [n]$, i.e. $\sigma = \langle p_{i_1}, \dots, p_{i_\ell} \rangle$ with $i_j < i_{j+1}$ for all $j \in [\ell - 1]$ and $\ell \leq n$. We call σ an *ε-simplification* of π if for each segment $\langle p_{i_j}, p_{i_{j+1}} \rangle$, we have $\delta(\langle p_{i_j}, p_{i_{j+1}} \rangle, \pi[i_j : i_{j+1}]) \leq \varepsilon$, where δ denotes some distance measure, e.g. the Hausdorff or the Fréchet distance.

The *Hausdorff distance* between two sets $P, Q \subset \mathbb{R}^2$ is defined as

$$d_{\mathrm{H}}(P, Q) \overset{\text{def}}{=} \max \left\{ \sup_{p \in P} \inf_{q \in Q} \|p - q\|, \sup_{q \in Q} \inf_{p \in P} \|p - q\| \right\}.$$

For two polygonal curves π and σ in $\mathbb{R}^2$, since π and σ are closed and bounded, we get

$$d_{\mathrm{H}}(\pi, \sigma) = \max \left\{ \max_{p \in \pi} \min_{q \in \sigma} \|p - q\|, \max_{q \in \sigma} \min_{p \in \pi} \|p - q\| \right\}.$$

The *Fréchet distance* is often described through an analogy with a person and a dog walking along their respective curves without backtracking, where the Fréchet distance is the shortest leash needed for such a walk. Formally, consider a set of *reparametrisations* Φ_ℓ of length ℓ, defined as continuous non-decreasing surjective functions $\phi : [0, 1] \to [1, \ell]$. Given two polygonal curves π and σ of lengths m and n, respectively, we can define the Fréchet distance as

$$d_{\mathrm{F}}(\pi, \sigma) \overset{\text{def}}{=} \inf_{\alpha \in \Phi_m, \beta \in \Phi_n} \max_{t \in [0,1]} \|\pi(\alpha(t)) - \sigma(\beta(t))\|.$$

We refer to the pair of reparametrisations as an *alignment.* We often consider the Fréchet distance between a curve $\pi = \langle p_1, \dots, p_n \rangle$ and a line segment p_1p_n, for some $n \in \mathbb{N}^{\geq 3}$. In this setting, the alignment can be described in a more intuitive way; see also Figure 2. It can be described as a sequence of locations on the line segment with which the vertices of the curve are aligned, $\langle s_2, \dots, s_{n-1} \rangle$, where $s_i \in [1, 2]$ for all $i \in \{2, \dots, n-1\}$ and $s_i \leq s_{i+1}$ for all $i \in \{2, \dots, n-2\}$. To see that, assign $s_1 \coloneqq 1$ and $s_n \coloneqq 2$ and construct a helper reparametrisation $\phi : [0, 1] \to [1, n]$, defined as $\phi(t) = (n-1) \cdot t + 1$ for any $t \in [0, 1]$. Construct another reparametrisation $\psi : [1, n] \to [1, 2]$, defined as

$$\psi(t) = \begin{cases} s_{\lfloor t \rfloor} \cdot (1 - t + \lfloor t \rfloor) + s_{\lfloor t \rfloor + 1} \cdot (t - \lfloor t \rfloor) & \text{if } t \in [1, n), \\ s_n & \text{if } t = n. \end{cases}$$

Note that ϕ and $\psi \circ \phi$ satisfy the definition of reparametrisations for π and p_1p_n, respectively.

We also define an *alignment* between a curve and a line segment for the Hausdorff distance (see Figure 2). It represents the map from the curve to the line segment, where each point on the curve is mapped to the closest point on the line segment. It is given by a sequence $\langle s_1, \dots, s_n \rangle$, where $s_i \in [1, 2]$ for all $i \in [n]$, such that $p_1p_n(s_i) = \operatorname{argmin}_{p' \in p_1p_n} \|p' - p_i\|$. In other words, $p_1p_n(s_i)$ is the closest point to p_i for all $i \in [n]$; as we discuss in Appendix A.1, the Hausdorff distance is realised as the distance between p_i and $p_1p_n(s_i)$ for some $i \in [n]$. Therefore, establishing such an alignment and checking that $\|p_1p_n(s_i) - p_i\| \leq \varepsilon$ for all $i \in [n]$ allows us to check that $d_{\mathrm{H}}(\pi, p_1p_n) \leq \varepsilon$ for some $\varepsilon \in \mathbb{R}^{>0}$.

We are discussing the following problem: given an uncertain curve $\mathcal{U} = \langle U_1, \dots, U_n \rangle$ with $n \in \mathbb{N}^{\geq 3}$ and $U_i \subset \mathbb{R}^2$ for all $i \in [n]$, and the threshold $\varepsilon \in \mathbb{R}^{>0}$, find a minimal-length subsequence $\mathcal{U}' = \langle U_{i_1}, \dots, U_{i_\ell} \rangle$ of $\mathcal{U}$ with $\ell \leq n$, such that for any realisation $\pi \Subset \mathcal{U}$, the corresponding realisation $\pi' \Subset \mathcal{U}'$ forms an ε-simplification of π under some distance measure δ. We solve this problem for the Hausdorff and the Fréchet distance for uncertainty modelled with indecisive points, line segments, disks, and convex polygons.

3 Overview of the Approach

We first present the summary of our approach. On the highest level, we use the *shortcut graph.* Each uncertain point of a curve $\mathcal{U} = \langle U_1, \dots, U_n \rangle$ corresponds to a vertex. An edge connects two vertices i and j if and only if the distance between any realisation of $\mathcal{U}[i : j]$ and the corresponding line segment from U_i to U_j is below the threshold. The path with the fewest edges from vertex 1 to vertex n then corresponds to the simplification using fewest uncertain points. So, we construct the shortcut graph and find the shortest path between two vertices. The key idea is that we find shortcuts that are valid for *all* realisations, so any sequence of shortcuts can be chosen.

In order to construct the shortcut graph, we need to check whether an edge should be added to the graph, i.e. whether a shortcut is *valid.* It is natural to think that shortcut testing can be solved by disk stabbing with disks of suitable radius, as in the work by Guibas et al. [22]. The idea would then be, given the distance threshold ε, to replace the uncertainty regions with the intersection of ε-disks over all the points of a region; this way we would e.g. replace disks of radius r by disks of radius $\varepsilon - r$, and then check if a shortcut stabs these regions. However, except for disks, this approach does not work – the reader can see this by trying to apply the method on an uncertainty region shaped as a long line segment (or a skinny convex polygon) that is parallel to the potential shortcut line segment. The intersection of ε-disks may be empty, while clearly one can create an alignment for both the Hausdorff and the Fréchet distance. For disks the approach is more suitable; however, when

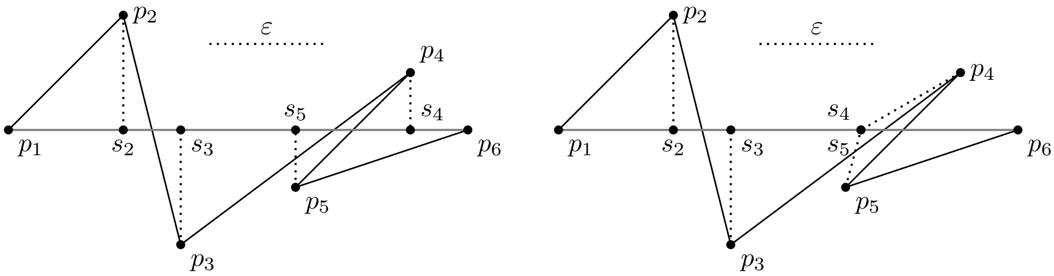

Figure 2 Left: Alignment for the Hausdorff distance. Right: Alignment for the Fréchet distance. In both cases, the alignment is described as the sequence $\langle s_1 \coloneqq p_1, s_2, s_3, s_4, s_5, s_6 \coloneqq p_6\rangle$.

testing a shortcut, the first and the last disk of a shortcut fulfil a different function than the intermediate disks. This means that we can rephrase the problem for the intermediate disks of a shortcut as disk stabbing, but not for the first and the last disk, as the quantifiers in the problem are different. Furthermore, the work by Guibas et al. [22] does not provide running time guarantees for disks of different radii, and the initialisation in their approach is not applicable in our setting with no restriction on disk intersections. So, we need to use a different approach to test shortcuts.

The approach is different for the Hausdorff and the Fréchet distance and for each uncertainty model. For the first and the last uncertain point of the shortcut, we state in Section 5 that there are several critical pairs of realisations that need to be tested explicitly, and then for any other pair of realisations, we know that the distance is also below the threshold. Testing each pair corresponds to finding the distance between a precise line segment and any realisation of an uncertain curve; we discuss this in Section 4 and show the procedures to do this in detail in Appendix A.

▶ **Theorem 1.** *We can find the shortest vertex-constrained simplification of an uncertain curve, such that for any realisation the simplification is valid, both for the Hausdorff and the Fréchet distance, in time $\mathcal{O}(n^3)$ for uncertainty modelled with disks and line segments, and in time $\mathcal{O}(n^3k^3)$ for uncertainty modelled with indecisive points and convex polygons, where k is the number of options or vertices and n is the length of the curve.*

4 Shortcut Testing: Intermediate

Here we discuss testing a shortcut with the first and the last points fixed, i.e. we want to check $\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p_1, \pi(n) \equiv p_n} \delta(\pi, p_1p_n) \le \varepsilon$ for $\delta \coloneqq d_\mathrm{H}$ and $\delta \coloneqq d_\mathrm{F}$. We can do so in linear time in all the models; here we show the intuitive explanation, and we treat this topic in detail in Appendix A. We solve the following problem.

▶ **Problem 2.** Given an uncertain curve $\mathcal{U} = \langle U_1, \dots, U_n\rangle$ on $n \in \mathbb{N}^{\ge 3}$ uncertain points in $\mathbb{R}^2$, as well as realisations $p_1 \in U_1$, $p_n \in U_n$, check if the largest Hausdorff or Fréchet distance between $\mathcal{U}$ and its one-segment simplification is below a threshold $\varepsilon \in \mathbb{R}^{>0}$ for any realisation with the fixed start and end points, i.e. for $\delta \coloneqq d_\mathrm{H}$ or $\delta \coloneqq d_\mathrm{F}$, verify

$$\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p_1, \pi(n) \equiv p_n} \delta(\pi, p_1p_n) \le \varepsilon\,.$$

Hausdorff distance. It is a well-known fact that the Hausdorff distance between the curve and the line segment that simplifies that curve is the largest distance from a vertex of the curve to the line segment, so $d_{\mathrm{H}}(\pi, \langle \pi(1), \pi(n) \rangle) = \max_{i \in [n]} d(\pi(i), \langle \pi(1), \pi(n) \rangle)$ for a polygonal curve π of length n. (See Figure 2.) We can use the same idea in the uncertain setting; however, for indecisive curves, we can choose any realisation for each intermediate point, so we need to test all of them, so we need the largest distance from any realisation of any indecisive point to the line segment. Then for indecisive points, given a curve $\mathcal{U} = \langle U_1, \dots, U_n \rangle$ with $U_i = \{p_i^1, \dots, p_i^k\}$ for all $i \in [n]$, we have $\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p_1, \pi(n) \equiv p_n} \delta(\pi, p_1 p_n) = \max_{i \in [n]} \max_{j \in [k]} d(p_i^j, p_1 p_n)$. For disks, line segments, and convex polygons, the key point is the same: all of the realisations of the intermediate points need to be close enough to the given line segment. For disks, we can simply check the furthest points, which are one radius further away from the line segment than the disk centre. For line segments and convex polygons, it suffices to test all the vertices.

Fréchet distance. For the Fréchet distance, there is also an intuitive procedure in the precise setting [22, Lemma 8]. We can align each vertex from the curve with the earliest possible point in the line segment. Each next point cannot be aligned before the previous points, so choosing the earliest possible alignment point maximises the possibilities for the remainder of the curve. (See Figure 2: s_4 is as close as possible to p_1.) We use the same approach in the uncertain setting; however, for indecisive points, as any realisation of a point is possible, we need to choose the realisation that pushes the earliest alignment forward the most, as this is the most restrictive realisation for the remainder of the curve. In more detail, we iteratively find the value for s_i. Given s_{i-1}, we find the earliest t_i^j along the segment for each realisation p_i^j of U_i, such that $\|t_i^j - p_i^j\| \le \varepsilon$ and $s_{i-1} \preccurlyeq t_i^j$. Then we pick $s_i \coloneqq \max_{j \in [k]} t_i^j$, in terms of $\preccurlyeq$. We continue this procedure until the end of the segment, starting with $s_1 \coloneqq p_1$ and assigning $s_n \coloneqq p_n$. In one direction, the sequence of s_i we find corresponds to a possible realisation; in the other direction, we can see that for any $i \in \{2, \dots, n-1\}$, we have $s_{i-1} \preccurlyeq t_i^j \preccurlyeq s_i$ for all $j \in [k]$; so for any other realisation the alignment is in order, as well. We can show for line segments and convex polygons that we again only need to focus on the vertices. For disks, we instead reframe the problem as that of disk stabbing. Instead of testing closeness from all points of some $D(c, r)$ to the line segment, we can check if the line segment stabs $D(c, \varepsilon - r)$ for the threshold ε. Then the correct alignment order corresponds to picking points inside disks in order. Again, choosing the earliest possible one is key.

5 Shortcut Testing: All Points

In the previous section, we have covered testing a shortcut, given that the first and the last points are fixed. Here we remove the restriction on the endpoints.

▶ **Problem 3.** Given an uncertain curve $\mathcal{U} = \langle U_1, \dots, U_n \rangle$ on $n \in \mathbb{N}^{\ge 3}$ uncertain points in $\mathbb{R}^2$, check if the largest Hausdorff or Fréchet distance between $\mathcal{U}$ and its one-segment simplification is below a threshold $\varepsilon \in \mathbb{R}^{>0}$ for any realisation, i.e. for $\delta \coloneqq d_{\mathrm{H}}$ or $\delta \coloneqq d_{\mathrm{F}}$, verify $\max_{\pi \Subset \mathcal{U}} \delta(\pi, \langle \pi(1), \pi(n) \rangle) \le \varepsilon$.

For the indecisive points, we can simply check all pairs from $U_1 \times U_n$; this is quite easy to show.

▶ **Lemma 4.** *Given $n, k \in \mathbb{N}^{>0}$, $n \ge 3$, and $\delta \coloneqq d_{\mathrm{H}}$ or $\delta \coloneqq d_{\mathrm{F}}$, for any indecisive curve $\mathcal{U} = \langle U_1, \dots, U_n \rangle$ with $U_i = \{p_i^1, \dots, p_i^k\}$ for all $i \in [n]$ and $p_i^j \in \mathbb{R}^2$ for all $i \in [n]$, $j \in [k]$, we have*

$$\max_{\pi \Subset \mathcal{U}} \delta(\pi, \langle \pi(1), \pi(n) \rangle) = \max_{a \in [k]} \max_{b \in [k]} \max_{\sigma \Subset \mathcal{U}, \sigma(1) \equiv p_1^a, \sigma(n) \equiv p_n^b} \delta(\sigma, p_1^a p_n^b)\,.$$

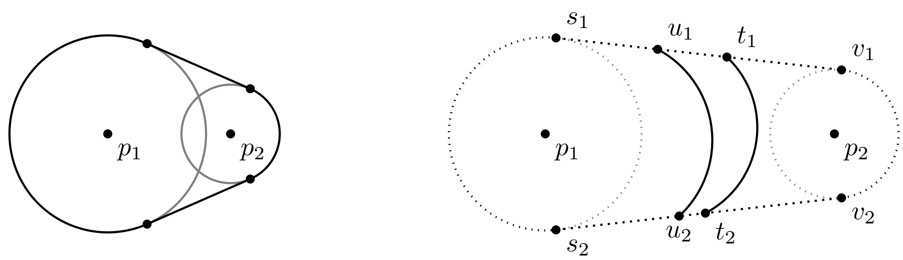

Figure 3 Left: Illustration for Observation 5. The convex hull of the disks is highlighted in black. The order in which the outer tangents touch the disks is the same. Right: Illustration for Definition 6. Here O_1 (t_1 to t_2) is to the right of O_2 (u_1 to u_2).

Proof. We can derive

$$\begin{aligned}
&\max_{\pi \Subset \mathcal{U}} \delta(\pi, \langle \pi(1), \pi(n) \rangle) \\
&\quad \{\text{Def. } \Subset\} \\
&= \max_{p_1 \in U_1, \ldots, p_n \in U_n} \delta(\langle p_1, \ldots, p_n \rangle, p_1 p_n) \\
&= \max_{p_1 \in U_1} \max_{p_n \in U_n} \max_{p_2 \in U_2, \ldots, p_{n-1} \in U_{n-1}} \delta(\langle p_1, \ldots, p_n \rangle, p_1 p_n) \\
&\quad \{\text{Def. } \Subset\} \\
&= \max_{p_1 \in U_1} \max_{p_n \in U_n} \max_{\sigma \Subset \mathcal{U}, \sigma(1) \equiv p_1, \sigma(n) \equiv p_n} \delta(\sigma, p_1 p_n) \\
&= \max_{a \in [k]} \max_{b \in [k]} \max_{\sigma \Subset \mathcal{U}, \sigma(1) \equiv p_1^a, \sigma(n) \equiv p_n^b} \delta(\sigma, p_1^a p_n^b)\,,
\end{aligned}$$

as was to be shown. ◀

That is to say, for either Hausdorff or Fréchet distance we can simply test the shortcut using the corresponding procedure from Lemma 16 or Lemma 20, and do so for each combination of the start and end points. We can then test an indecisive shortcut of length n overall in time $\mathcal{O}(k^2 \cdot nk) = \mathcal{O}(nk^3)$.

We now proceed to show the approach for disks and polygonal closed convex sets. The procedure is the same for the Hausdorff and the Fréchet distance, but differs between disks and PCCSs, since disks have some convenient special properties.

5.1 Disks

▶ **Observation 5.** *Suppose we are given two non-degenerate disks $D_1 \coloneqq D(p_1, r_1)$ and $D_2 \coloneqq D(p_2, r_2)$ with $D_1 \nsubseteq D_2$ and $D_2 \nsubseteq D_1$. We make the following observations.*

- *There are exactly two* outer tangents *to the disks, and the convex hull of $D_1 \cup D_2$ consists of an arc from D_1, an arc from D_2, and the outer tangents.*
- *Assume the lines of the outer tangents intersect. When viewed from the intersection point, the order in which the tangents touch the disks is the same, i.e. either both first touch D_1 and then D_2, or the other way around. If the lines are parallel, the same statement holds when viewed from points on the tangent lines at infinity. (See Figure 3.)*

To see that the second point is true, note that the distance from the intersection point to the tangent points of a disk is the same for both tangent lines. These observations mean that we can restrict our attention to the area bounded by the outer tangents and define an ordering in the resulting strip.

▶ **Definition 6.** *Given two distinct non-degenerate disks $D_1 \coloneqq D(p_1, r_1)$ and $D_2 \coloneqq D(p_2, r_2)$, consider a strip defined by the lines that form the outer tangents to the disks. Assume we have two circular arcs O_1, O_2 that intersect both tangents and lie inside the strip. Define s_1 and v_1 to be the points where one of the tangents touches D_1 and D_2, respectively, and let t_1 and u_1 be the points where O_1 and O_2 intersect that tangent, respectively. Define the order on the tangents from D_1 to D_2, so $s_1 \prec v_1$. Define points s_2, t_2, u_2, v_2 similarly for the other tangent. We say that O_2 is* to the right *of O_1 if either $t_i = u_i$ for $i \in \{1, 2\}$ and the radius of O_1 is larger than that of O_2; or if otherwise $t_i \preccurlyeq u_i$ for $i \in \{1, 2\}$ and O_1 and O_2 do not properly intersect. We say that O_2 is* to the left *of O_1 if either $t_i = u_i$ for $i \in \{1, 2\}$ and the radius of O_1 is smaller than that of O_2; or if otherwise $u_i \preccurlyeq t_i$ for $i \in \{1, 2\}$ and O_1 and O_2 do not properly intersect. (See Figure 3 for a visual interpretation.)*

We state the main result: it suffices to check the tangents to the first and the last disk and the order of the intermediate disks.

▶ **Lemma 7.** *Given $n \in \mathbb{N}^{\geq 3}$, for any imprecise curve modelled with disks $\mathcal{U} = \langle U_1, \dots, U_n \rangle$ with $U_i = D(c_i, r_i)$ for all $i \in [n]$ and $c_i \in \mathbb{R}^2$, $r_i \in \mathbb{R}^{\geq 0}$ for all $i \in [n]$, and assuming $U_1 \neq U_n$, we have with $\delta \in \{d_\mathrm{H}, d_\mathrm{F}\}$ that $\max_{\pi \Subset \mathcal{U}} \delta(\pi, \langle \pi(1), \pi(n) \rangle) \leq \varepsilon$ if and only if both of the following are true:*

- $$\max\left\{ \max_{\pi \Subset \mathcal{U}, \pi(1) \equiv s, \pi(n) \equiv t} \delta(\pi, st), \max_{\pi \Subset \mathcal{U}, \pi(1) \equiv u, \pi(n) \equiv v} \delta(\pi, uv) \right\} \leq \varepsilon\,,$$

  where $s, u \in U_1$, $t, v \in U_n$, and st and uv are the outer tangents to $U_1 \cup U_n$;
- *for each $i \in \{2, \dots, n-1\}$, the right arc of the disk $D(c_i, \varepsilon - r_i)$ bounded by the intersection points with the tangent lines is to the right of the right arc of U_1 and the left arc of the disk $D(c_i, \varepsilon - r_i)$ is to the left of the left arc of U_n.*

Proof. We first prove the claim for $\delta = d_\mathrm{H}$. Assume the right side of the lemma statement holds. First of all, as we have $\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv s, \pi(n) \equiv t} d_\mathrm{H}(\pi, st) \leq \varepsilon$, we know that for all $i \in \{2, \dots, n-1\}$, we have $d(c_i, st) + r_i \leq \varepsilon$, or $d(c_i, st) \leq \varepsilon - r_i$, so st stabs each disk $D(c_i, \varepsilon - r_i)$ (see Lemma 18 in Appendix A.1). We can draw a similar conclusion for uv. Therefore, each disk $D(c_i, \varepsilon - r_i)$ crosses the entire strip bounded by the tangent lines, with the intersection points splitting it into the left and the right circular arcs. We can thus apply Definition 6 to these arcs, as stated in the lemma.

First suppose that the disks U_1 and U_n do not intersect. Then for any line segment from U_1 to U_n and any disk $D' \coloneqq D(c_i, \varepsilon - r_i)$, we exit D' after exiting U_1 and enter D' before entering U_n. Hence, for any line pq with $p \in U_1$ and $q \in U_n$ and any $i \in \{2, \dots, n-1\}$, we can find a point $w \in pq \cap D'$; this means that indeed $\max_{w' \in U_i} d(w', pq) \leq \varepsilon$ (see Lemma 22 in Appendix A.2). As this holds for all disks and any choice of p and q, we conclude that $\max_{\pi \Subset \mathcal{U}} d_\mathrm{H}(\pi, \langle \pi(1), \pi(n) \rangle) \leq \varepsilon$.

Now assume that the disks U_1 and U_n intersect. If we consider the line segments pq with $p \in U_1$, $q \in U_n$, we end up in the previous case if either $p \notin U_1 \cap U_n$ or $q \notin U_1 \cap U_n$. So assume that the segment pq lies entirely in the intersection $U_1 \cap U_n$. However, it can be seen that for each disk $D' \coloneqq D(c_i, \varepsilon - r_i)$, the left boundary of the intersection is to the right of the left boundary of the disk, and the right boundary of the intersection is to the left of the right boundary of the disk; hence, $pq \subset U_1 \cap U_n \subseteq D'$. Therefore, we have $\max_{w' \in U_i} d(w', pq) \leq \varepsilon$, and so also in this case $\max_{\pi \Subset \mathcal{U}} d_\mathrm{H}(\pi, \langle \pi(1), \pi(n) \rangle) \leq \varepsilon$.

We now assume that the right side of the lemma statement is false and show that then $\max_{\pi \Subset \mathcal{U}} d_\mathrm{H}(\pi, \langle \pi(1), \pi(n) \rangle) > \varepsilon$. If $\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv s, \pi(n) \equiv t} d_\mathrm{H}(\pi, st) > \varepsilon$, then immediately $\max_{\pi \Subset \mathcal{U}} d_\mathrm{H}(\pi, \langle \pi(1), \pi(n) \rangle) > \varepsilon$. Same holds for uv. So, assume those statements hold; then it must be that for at least one intermediate disk the arcs do not lie to the left or to the

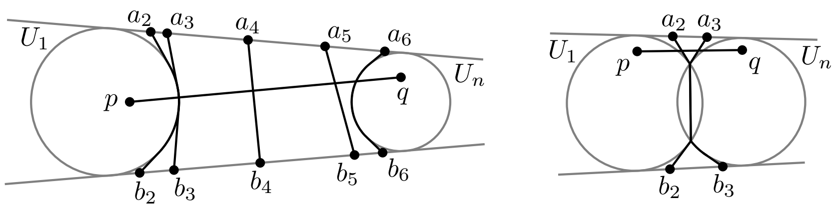

Figure 4 Having established the alignments along the two tangents, we can connect them to create a sequence of paths.

right of the arcs of the respective disks. Assume this is disk i, so the disk $D' := D(c_i, \varepsilon - r_i)$. W.l.o.g. assume that the right arc of the disk does not lie entirely to the right of the right arc of U_1. The argument for the left arc w.r.t. U_n is symmetric.

There must be at least one point p' on the right arc of U_1 that lies outside of D'. Assume for now that U_1 and U_n are disjoint. Then a line segment $p'q$ for any $q \in U_n$ does not stab D', so $\max_{w' \in U_i} d(w', pq) > \varepsilon$, and so $\max_{\pi \Subset \mathcal{U}} d_{\mathrm{H}}(\pi, \langle \pi(1), \pi(n) \rangle) > \varepsilon$. If U_1 and U_n intersect, then either p' is outside of the intersection and of D' and there is a point $q \in U_n$ such that $p'q$ does not stab D'; or we can pick the degenerate line segment $p'p'$, as $p' \in U_1 \cap U_n$, and so $p'p'$ also does not stab D'. In either case, we conclude that $\max_{\pi \Subset \mathcal{U}} d_{\mathrm{H}}(\pi, \langle \pi(1), \pi(n) \rangle) > \varepsilon$.

For $\delta = d_{\mathrm{F}}$, first assume that $\max_{\pi \Subset \mathcal{U}} d_{\mathrm{F}}(\pi, \langle \pi(1), \pi(n) \rangle) \leq \varepsilon$. As $d_{\mathrm{F}}(\pi, \sigma) \geq d_{\mathrm{H}}(\pi, \sigma)$ for any curves π, σ, this also means that $\max_{\pi \Subset \mathcal{U}} d_{\mathrm{H}}(\pi, \langle \pi(1), \pi(n) \rangle) \leq \varepsilon$. Furthermore, immediately we get that $\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv s, \pi(n) \equiv t} d_{\mathrm{F}}(\pi, st) \leq \varepsilon$, and the same for uv, which yields the right side of the lemma as shown above.

Now assume that the right side holds. As for the Hausdorff distance, we know that the disks cross the entire strip and that Definition 6 applies. It remains to show that for any line segment pq with $p \in U_1$, $q \in U_n$, there is a valid alignment that maintains the correct ordering and bottleneck distance, assuming it exists for every realisation for st and uv. Consider a valid alignment established for st and uv, so the sequence of points a_i on st and b_i on uv that are mapped to U_i. We can always find such points for each individual U_i (see Lemma 21 in Appendix A.2), and as we know that the Fréchet distance is below the threshold for st and uv, there is such a valid alignment, i.e. we know that $a_i \preccurlyeq a_{i+1}$ and $b_i \preccurlyeq b_{i+1}$ for all $i \in [n-1]$.

For the rest of the proof, the rough idea is as follows. We can create paths from a_i to b_i so that every segment pq with $p \in U_1$ and $q \in U_n$ crosses these paths in the correct order, thus proving that a Fréchet alignment exists. When U_1 and U_n are disjoint, these paths are simply geodesic paths within the region bounded by the two tangents and the U_1 and U_n. If they intersect, we can instead create these paths by connecting a_i to the top intersection point of the disks and b_i to the bottom intersection point, as in Figure 4. Note that when the two disks intersect and the segment pq goes through the intersection, it may not cross the paths at all; however, every point in the intersection is close enough to all intermediate disks. We now discuss this idea in more detail.

First suppose that the disks U_1 and U_n do not intersect. Consider the region R bounded by the outer tangents and the disk arcs that are not part of the convex hull of $U_1 \cup U_n$. We connect, for each $i \in \{2, \ldots, n-1\}$, a_i to b_i with a geodesic shortest path in R, as in Figure 4. We claim that for any line segment pq defined above, the intersection points of the shortest paths with the segment give a valid alignment, yielding $\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p, \pi(n) \equiv q} d_{\mathrm{F}}(\pi, pq) \leq \varepsilon$. As the choice of pq was arbitrary, this will complete the proof.

To show that the alignment is valid, we need to show that the order is correct and that the distances fall below the threshold. First consider the case where the geodesic shortest path for point i does not touch the boundary formed by arcs of region R. In this case, it is simply a line segment $a_i b_i$. Note that by definition $a_i, b_i \in D(c_i, \varepsilon - r_i)$; as disks are convex, also $a_i b_i \subset D(c_i, \varepsilon - r_i)$; thus, the intersection point p'_i of pq with $a_i b_i$ is in $D(c_i, \varepsilon - r_i)$, so by Lemma 21, $\max_{w \in U_i} \|p'_i - w\| \leq \varepsilon$. Furthermore, note that $a_i \preccurlyeq a_{i+1}$ and $b_i \preccurlyeq b_{i+1}$; thus, the line segments $a_i b_i$ and $a_{i+1} b_{i+1}$ cannot cross, so also $p'_i \preccurlyeq p'_{i+1}$.

Now w.l.o.g. consider the case where the geodesic shortest path for point i touches the arc of U_1. The geodesic shortest paths do not cross: on the path from a_i (or b_i) to the arc they form a tangent to the arc, thus for $a_i \preccurlyeq a_{i+1}$ the tangent point for a_i comes before that of a_{i+1} when going along the arc from s to u. So, just as in the previous case, these line segments cannot cross. Having reached the arc, both shortest paths will follow it, as otherwise the path would not be a shortest path; thus, the arcs do not cross, either. Finally, a path from the previous case does not touch any path that touches the arc boundary of R by definition. Finally, note that the condition that we have established on the right arcs of disks being to the right of the right arc of U_1 (and symmetric for the left arcs and U_n) means that the geodesic shortest paths that touch the arc boundary of R stay within the respective disks $D(c_i, \varepsilon - r_i)$. Thus, we have established that for all i we have $p'_i \preccurlyeq p'_{i+1}$ and $\max_{w \in U_i} \|p'_i - w\| \leq \varepsilon$, concluding the proof for disjoint U_1 and U_n.

Finally, consider the case where U_1 intersects U_n. Above we used geodesic paths within the region R. However, when U_1 intersects U_n, R consists of two disconnected regions. Observe that one region contains a_i and the other contains b_i. To connect a_i with b_i we use the geodesic from a_i to the intersection point of the two inner boundaries of U_1 and U_n that is in the same region of R, the geodesic from b_i to the other intersection point of the inner boundaries, and join these two by a line segment between the intersection points. Any line segment from a point in U_1 to a point in U_n crosses these paths in order, just like in the previous case. If the line segment goes through the intersection, note that any point in the intersection is close enough to all the intermediate objects, as the intersection is the subset of each disk. So, any point in the intersection can be chosen to establish the trivially in-order alignment to all the intermediate objects. ◀

It is worth noting that the case of $U_1 = U_n$ is similar to how we treat the intersection $U_1 \cap U_n$; however, our definition for the ordering between two disks does not apply. So, if $U_1 = U_n$, then $\max_{\pi \Subset \mathcal{U}} \delta(\pi, \langle \pi(1), \pi(n) \rangle) \leq \varepsilon$ if and only if $U_1 \subseteq D(c_i, \varepsilon - r_i)$ for all $i \in \{2, \ldots, n-1\}$.

5.2 Non-intersecting PCCSs

Suppose the regions are modelled by convex polygons. Consider first the case where the interiors of U_1 and U_n do not intersect, so at most they share a boundary segment.

▶ **Observation 8.** *Given an uncertain curve modelled by convex polygons $\mathcal{U} = \langle U_1, \ldots, U_n \rangle$ with the interiors of U_1 and U_n not intersecting, note:*

- *There are two* outer tangents *to the polygons U_1 and U_n, and the convex hull of $U_1 \cup U_2$ consists of a convex chain from U_1, a convex chain from U_n, and the outer tangents.*
- *Let C_i be the convex chain from U_i that is not part of the convex hull for $i \in \{1, n\}$. Then for $\delta \coloneqq d_\mathrm{H}$ or $\delta \coloneqq d_\mathrm{F}$,*

$$\max_{\pi \Subset \mathcal{U}} \delta(\pi, \langle \pi(1), \pi(n) \rangle) \leq \varepsilon \iff \max_{\pi \Subset \mathcal{U}, \pi(1) \in C_1, \pi(n) \in C_n} \delta(\pi, \langle \pi(1), \pi(n) \rangle) \leq \varepsilon \,.$$

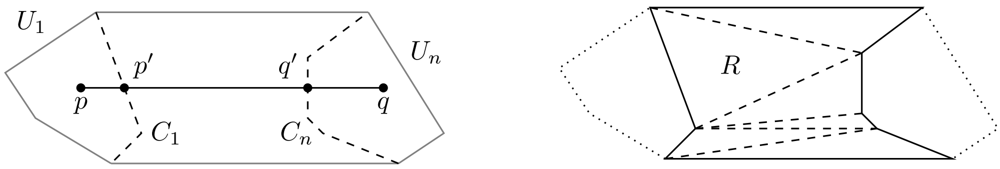

Figure 5 Left: Illustration for Observation 8. The convex hull of the polygons is shown in grey. The dotted chains are C_1 and C_n. Any line segment pq with $p \in U_1$ and $q \in U_n$ crosses C_1 and C_n. Right: Illustration for the procedure. The region R is triangulated.

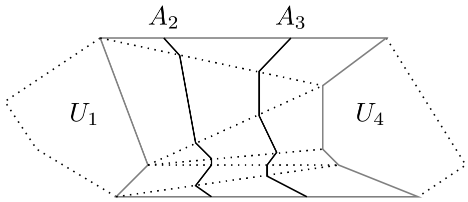

Figure 6 An example set of curves $A = \{A_2, A_3\}$ discussed in Lemmas 9 and 10.

To see that the second observation is true, note that one direction is trivial. In the other direction, note that any line segment pq with $p \in U_1$, $q \in U_n$ crosses both C_1 and C_n, say, at $p' \in C_1$ and $q' \in C_n$. We know that there is a valid alignment for $p'q'$, both for the Hausdorff and the Fréchet distance; we can then use this alignment for pq. See Figure 5.

We claim that we can check $\max_{\pi \Subset \mathcal{U}} d_{\mathrm{H}}(\pi, \langle \pi(1), \pi(n) \rangle) \leq \varepsilon$ using the following procedure (see Figure 5).

1. Triangulate the region R bounded by two convex chains C_1 and C_n and the outer tangents.
2. For each line segment st of the triangulation with $s \in C_1$, $t \in C_n$, and for either $\delta := d_{\mathrm{H}}$ or $\delta := d_{\mathrm{F}}$, check that $\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv s, \pi(n) \equiv t} \delta(\pi, st) \leq \varepsilon$.

First of all, observe that we can compute a triangulation, and that every triangle has two points from one convex chain and one point from the other chain. If all three points were from the same chain, then the triangle would lie outside of R. Now consider some line segment pq with $p \in C_1$, $q \in C_n$. To complete the argument, it remains to show that the checks in step 2 mean that also $\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p, \pi(n) \equiv q} \delta(\pi, pq) \leq \varepsilon$. Observe that the triangles span across the region R, so when going from one tangent to the other within R we cross all the triangles. Therefore, we can number the edges of the triangles that go from C_1 to C_n, in the order of occurrence on such a path, from 1 to k. Denote the alignment established on line $j \in [k]$ with the sequence of a_i^j, for $i \in \{2, \ldots, n-1\}$; this alignment can be established both for $\delta := d_{\mathrm{H}}$ and $\delta := d_{\mathrm{F}}$. We can establish polygonal curves $A_i := \langle a_i^1, \ldots, a_i^k \rangle$; they all stay within R (see Figure 6). We claim that for any line segment pq defined above, we can establish a valid alignment from intersection points of pq and A_i. We do this separately for the Fréchet and the Hausdorff distance.

▶ **Lemma 9.** *Given a set of curves $A := \{A_2, \ldots, A_{n-1}\}$ in R described above for $\delta := d_{\mathrm{H}}$ and a line segment pq with $p \in C_1$, $q \in C_n$, we have* $\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p, \pi(n) \equiv q} d_{\mathrm{H}}(\pi, pq) \leq \varepsilon$.

Proof. Note that pq crosses each A_i at least once. We can take any one crossing for each i and establish the alignment. Consider such a crossing point p_i'. It falls in some triangle bounded by a segment from either C_1 or C_n and two line segments that contain points a_i^j and a_i^{j+1} for some $j \in [k-1]$. We know, using Lemma 19, that $\max_{w \in U_i} \|a_i^j - w\| \leq \varepsilon$ and $\max_{w \in U_i} \|a_i^{j+1} - w\| \leq \varepsilon$. Consider any point $w' \in U_i$. Then, using Lemma 14 with $c := d := w'$, we find that $\|w' - p_i'\| \leq \varepsilon$. Therefore, also $\max_{w \in U_i} \|p_i' - w\| \leq \varepsilon$; using Lemma 19, we conclude that indeed $\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p, \pi(n) \equiv q} d_{\mathrm{H}}(\pi, pq) \leq \varepsilon$. ◀

For the Fréchet distance, we can use the same argument to show closeness; however, we need more care to establish the correct order for the alignment.

▶ **Lemma 10.** *Given a set of curves $A \coloneqq \{A_2, \dots, A_{n-1}\}$ in R described above for $\delta \coloneqq d_\mathrm{F}$ and a line segment pq with $p \in C_1$, $q \in C_n$, we have* $\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p, \pi(n) \equiv q} d_\mathrm{F}(\pi, pq) \leq \varepsilon$.

Proof. Compared to Lemma 9, instead of taking any intersection point of pq with each A_i, we take the *last* intersection point. First, we need to show that curves A_i and A_{i+1} do not cross for any $i \in [n-1]$. Each curve A_i crosses each triangle once, so it suffices to show that a segment $a_i^j a_i^{j+1}$ does not cross $a_{i+1}^j a_{i+1}^{j+1}$. Indeed, as $a_i^j \preccurlyeq a_{i+1}^j$ and $a_i^{j+1} \preccurlyeq a_{i+1}^{j+1}$, these line segments cannot cross.

Now consider, for each $i \in \{2, \dots, n-1\}$. the polygon P_i bounded by C_1, A_i, and the corresponding segments of the outer tangents. With the previous statement, it is easy to see that $P_2 \subseteq P_3 \subseteq \dots \subseteq P_{n-1}$. Assume this is not the case, so some $P_i \not\subseteq P_{i+1}$. Then there is a point $z \in P_i$, but $z \notin P_{i+1}$. The point z falls into some triangle with lines j and $j+1$. In this triangle, it means that z is between C_1 and $a_i^j a_i^{j+1}$, but not between C_1 and $a_{i+1}^j a_{i+1}^{j+1}$. However, as these segments do not cross, this would imply that $a_{i+1}^j \prec a_i^j$, but then the check in step 2 would not pass for line j.

Consider the points at which the line segment pq leaves the polygons P_i for the last time. From the definition it is obvious that $p \in P_i$ for all $i \in \{2, \dots, n-1\}$, so this is well-defined. Clearly, due to the subset relationship, the order of such points p'_i is correct, i.e. $p'_i \preccurlyeq p'_{i+1}$. Furthermore, each such $p'_i \in A_i$, so using the arguments of Lemma 9 we can show that also the distances are below ε. Thus, we conclude that indeed $\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p, \pi(n) \equiv q} d_\mathrm{F}(\pi, pq) \leq \varepsilon$. ◀

The proofs of Lemmas 9 and 10 show us how to solve the problem for two convex polygons with non-intersecting interiors. We can also use them directly for the case of line segments that do not intersect except at endpoints.

▶ **Lemma 11.** *Given $n \in \mathbb{N}^{\geq 3}$, for any imprecise curve modelled with line segments $\mathcal{U} = \langle U_1, \dots, U_n \rangle$ with $U_i = p_i^1 p_i^2 \subset \mathbb{R}^2$ for all $i \in [n]$, given a threshold $\varepsilon \in \mathbb{R}^{>0}$, and given that $U_1 \cap U_n \subset \{p_1^1, p_1^2\}$, and assuming that the triangles $p_1^1 p_n^1 p_1^2$ and $p_1^2 p_n^1 p_n^2$ form a triangulation of the convex hull of $U_1 \cup U_n$, we have $\max_{\pi \Subset \mathcal{U}} \delta(\pi, \langle \pi(1), \pi(n) \rangle) \leq \varepsilon$ if and only if*

$$\max \Big\{ \max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p_1^1, \pi(n) \equiv p_n^1} \delta(\pi, p_1^1 p_n^1), \max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p_1^2, \pi(n) \equiv p_n^1} \delta(\pi, p_1^2 p_n^1), \max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p_1^2, \pi(n) \equiv p_n^2} \delta(\pi, p_1^2 p_n^2) \Big\} \leq \varepsilon\,.$$

We should note that in this particular case it is not necessary to use a triangulation, so we can get rid of the second term; also in the previous proofs a convex partition could work instead, but a triangulation is easier to define.

5.3 Intersecting PCCSs

We now discuss the situation where the interiors of U_1 and U_n intersect, or where line segments U_1 and U_n cross. The argument is the same for $\delta \coloneqq d_\mathrm{H}$ and $\delta \coloneqq d_\mathrm{F}$.

Line segments. Assume line segments $U_1 \stackrel{\text{def}}{=} p_1^1 p_1^2$ and $U_n \stackrel{\text{def}}{=} p_n^1 p_n^2$ cross; call their intersection point s. Then we can use Lemma 11 separately on pairs of $\{p_1^1 s, s p_1^2\} \times \{p_n^1 s, s p_n^2\}$. These pairs cover the entire set of realisations pq with $p \in U_1$, $q \in U_n$, completing the checks.

▶ **Lemma 12.** *Given $n \in \mathbb{N}^{\geq 3}$, for any imprecise curve modelled with line segments $\mathcal{U} = \langle U_1, \dots, U_n \rangle$ with $U_i = p_i^1 p_i^2 \subset \mathbb{R}^2$ for all $i \in [n]$, given a threshold $\varepsilon \in \mathbb{R}^{>0}$, we can check for both $\delta \coloneqq d_\mathrm{H}$ and $\delta \coloneqq d_\mathrm{F}$, using procedures above, that $\max_{\pi \Subset \mathcal{U}} \delta(\pi, \langle \pi(1), \pi(n) \rangle) \leq \varepsilon$.*

Convex polygons. Convex polygons whose interiors intersect can be partitioned along the intersection lines, so into a convex polygon $R \coloneqq U_1 \cap U_n$ and two sets of polygons $\mathcal{P}_1 \coloneqq \{P_1^1, \ldots, P_1^k\}$ and $\mathcal{P}_n \coloneqq \{P_n^1, \ldots, P_n^\ell\}$ for some $k, \ell \in \mathbb{N}^{>0}$. Just as for line segments, we can look at pairs from $\mathcal{P}_1 \times \mathcal{P}_n$ separately. The pairs where R is involved are treated later. Consider some $(P, Q) \in \mathcal{P}_1 \times \mathcal{P}_n$. Note that P and Q are convex polygons with a convex cut-out, so the boundary forms a convex chain, followed by a concave chain. We need to compute some convex polygons P' and Q' with non-intersecting interiors that are equivalent to P and Q, so that we can apply the approaches from Section 5.2.

We claim that we can simply take the convex hull of P and Q to obtain P' and Q'. Clearly, the resulting polygons will be convex. Also, the concave chains of P are bounded by points s and t and are replaced with the line segment st; same happens for Q with point u and v. The points s, t, u, v are points of intersection of original polygons U_1 and U_n, so they lie on the boundary of R, and their order along that boundary can only be s, t, u, v or s, t, v, u. Thus, it cannot happen that st crosses uv, and it cannot be that uv is in the interior of the convex hull of P, as otherwise R would not be convex. Hence, the interiors of P' and Q' cannot intersect, so they satisfy the necessary conditions.

Finally, we need to show that the solution for (P', Q') is equivalent to that for (P, Q). One direction is trivial, as $P \subseteq P'$ and $Q \subseteq Q'$; for the other direction, consider any line segment that leaves P through the concave chain. In our approach, we test the lines starting in s and t; the established alignments are connected into paths. The paths A_i do not cross st. So, any alignment in the region of $\mathrm{CH}(P \cup Q) \setminus (P \cup Q)$ can also be made in the region $\mathrm{CH}(P' \cup Q') \setminus (P' \cup Q')$. So, this approach yields valid solutions for all pairs not involving R.

Now consider the pair (R, R). A curve may now consist of a single point, so the approach for the Fréchet and the Hausdorff distance is the same: all the points of U_i need to be close enough to all the points of R. To check that, observe that the pair of points $p \in U_i$ and $q \in R$ that has maximal distance has the property that p is an extreme point of U_i in direction qp and q is an extreme point of R in direction pq. So, it suffices, starting at the rightmost point of U_i and leftmost point of R in some coordinate system, to then rotate clockwise around both regions keeping track of the distance between tangent points. Note that only vertices need to be considered, as the extremal point cannot lie on an edge.

Finally, any other pair that involves R is covered by the stronger case of (R, R): for any line we can align every intermediate object with any point in R. To elaborate, the cases above are not truly a case distinction, as *all* of these combinations should hold; so given a line segment for a pair (P, R) or (R, Q) for some $P \in \mathcal{P}_1$, $Q \in \mathcal{P}_n$, we can pick any point of the segment that lies inside R to establish the alignment, deferring to the stronger previous case (R, R). Also observe that some line segments covered by the case (P, Q) with $P \in \mathcal{P}_1$, $Q \in \mathcal{P}_n$ may go through R; this does not impose any unnecessary constraints, so it does not matter that the cases can overlap.

▶ **Lemma 13.** *Given $n \in \mathbb{N}^{\geq 3}$, for any imprecise curve modelled with convex polygons $\mathcal{U} = \langle U_1, \ldots, U_n \rangle$ with $U_i \subset \mathbb{R}^2$ for all $i \in [n]$ and $V(U_i) = \{p_i^1, \ldots, p_i^k\}$ for all $i \in [n]$, $k \in \mathbb{N}^{>0}$, given a threshold $\varepsilon \in \mathbb{R}^{>0}$, we can check for $\delta \coloneqq d_{\mathrm{H}}$ and $\delta \coloneqq d_{\mathrm{F}}$, using procedures above, that $\max_{\pi \Subset \mathcal{U}} \delta(\pi, \langle \pi(1), \pi(n) \rangle) \leq \varepsilon$.*

6 Combining Steps

In the previous sections, we have shown how to check if a shortcut of length $n \geq 3$ is valid under the Hausdorff or the Fréchet distance, for indecisive points, disks, line segments, and convex polygons. It is easy to see that a shortcut of length $n = 2$ is always valid. Therefore,

Table 1 Running time of our approach in each setting. For indecisive points, k is the number of options per point. For convex polygons, k is the number of vertices.

	Indecisive	Disks	Line segments	Convex polygons
Hausdorff distance	$\mathcal{O}(n^3k^3)$	$\mathcal{O}(n^3)$	$\mathcal{O}(n^3)$	$\mathcal{O}(n^3k^3)$
Fréchet distance	$\mathcal{O}(n^3k^3)$	$\mathcal{O}(n^3)$	$\mathcal{O}(n^3)$	$\mathcal{O}(n^3k^3)$

we can use the previously described procedures to construct a shortcut graph; any path in such a graph from the vertex 1 to vertex n corresponds to a valid simplification, so the shortest path gives us the result we need.

▶ **Theorem 1.** *We can find the shortest vertex-constrained simplification of an uncertain curve, such that for any realisation the simplification is valid, both for the Hausdorff and the Fréchet distance, in time $\mathcal{O}(n^3)$ for uncertainty modelled with disks and line segments, and in time $\mathcal{O}(n^3k^3)$ for uncertainty modelled with indecisive points and convex polygons, where k is the number of options or vertices and n is the length of the curve.*

Proof. Correctness of the approaches has been shown before. We now analyse the running time, also shown in Table 1. For the running time, observe that we need $\mathcal{O}(n^2T)$ time in any setting, due to the shortcut graph construction.

For indecisive points, when testing a shortcut we do $\mathcal{O}(nk)$-time testing for $\mathcal{O}(k^2)$ combinations of starting and ending points, where k is the number of options per point. For disks, we do a linear number of constant-time checks and two linear-time checks, getting $T \in \mathcal{O}(n)$. For line segments, we also do two (three) linear-time checks per part; two line segments can be split into at most two parts each, so we repeat the process four times. Either way, we get $T \in \mathcal{O}(n)$.

Finally, for convex polygons, assume the complexity of each polygon is at most k. Assume the partitioning resulting from two intersecting polygons yields ℓ_1 and ℓ_2 parts for the first and the second polygon, respectively. Denote the two polygons P and Q and the resulting parts with $P_1, \dots, P_{\ell_1}$ and $Q_1, \dots, Q_{\ell_2}$, respectively. Suppose part P_i has complexity k_i and part Q_j has complexity k'_j, so $|V(P_i)| = k_i$ and $|V(Q_j)| = k'_j$ for some $i \in [\ell_1]$, $j \in [\ell_2]$. We know that every vertex of the original polygons occurs in a constant number of parts, so $\sum_{i=1}^{\ell_1} k_i \in \mathcal{O}(k)$ and $\sum_{j=1}^{\ell_2} k'_j \in \mathcal{O}(k)$; we also know $\ell_1 + \ell_2 \in \mathcal{O}(k)$. We consider all pairs from P and Q, and for each pair we triangulate and do the checks on the triangulation. The triangulation can be done in time $\mathcal{O}((k_i + k'_j) \cdot \log(k_i + k'_j))$, yielding $\mathcal{O}(k_i + k'_j)$ lines, each of which is tested in time $\mathcal{O}(nk)$. The testing dominates, so we need $\mathcal{O}((k_i + k'_j) \cdot nk)$ time. We are interested in

$$\sum_{i=1}^{\ell_1} \sum_{j=1}^{\ell_2} \mathcal{O}((k_i + k'_j) \cdot nk) = \mathcal{O}(nk) \cdot \sum_{i=1}^{\ell_1} \sum_{j=1}^{\ell_2} \mathcal{O}(k_i + k'_j) = \mathcal{O}(nk^3)\,.$$

So, $T \in \mathcal{O}(nk^3)$ both for the Fréchet and the Hausdorff distance. ◀

7 Conclusion

We have presented approaches for finding the optimal simplification of an uncertain curve under various uncertainty models for the Hausdorff and the Fréchet distance. To recap, we can use Lemmas 7, 12, and 13 and the procedure for indecisive points to test a single shortcut. Constructing a shortcut graph yields the solution. In future work, it would be interesting to see, similarly to the precise simplification approaches, if an improvement in the running time is possible to subcubic time, or whether one can show a conditional lower bound [8]. It would also be interesting to consider what uncertainty means in the context of global simplification; our approach does not seem easily transferable.

References

1. Pankaj K. Agarwal, Boris Aronov, Sariel Har-Peled, Jeff M. Phillips, Ke Yi, and Wuzhou Zhang. Nearest-neighbor searching under uncertainty II. *ACM Transactions on Algorithms*, 13(1):3:1–3:25, 2016. doi:10.1145/2955098.
2. Pankaj K. Agarwal, Alon Efrat, Swaminathan Sankararaman, and Wuzhou Zhang. Nearest-neighbor searching under uncertainty I. *Discrete & Computational Geometry*, 58(3):705–745, 2017. doi:10.1007/s00454-017-9903-x.
3. Pankaj K. Agarwal, Sariel Har-Peled, Nabil H. Mustafa, and Yusu Wang. Near-linear time approximation algorithms for curve simplification. *Algorithmica*, 42(3):203–219, 2005. doi:10.1007/s00453-005-1165-y.
4. Pankaj K. Agarwal and Kasturi R. Varadarajan. Efficient algorithms for approximating polygonal chains. *Discrete & Computational Geometry*, 23(2):273–291, 2000. doi:10.1007/PL00009500.
5. Hee-Kap Ahn, Christian Knauer, Marc Scherfenberg, Lena Schlipf, and Antoine Vigneron. Computing the discrete Fréchet distance with imprecise input. *International Journal of Computational Geometry & Applications*, 22(01):27–44, 2012. doi:10.1142/S0218195912600023.
6. Sander P. A. Alewijnse, Kevin Buchin, Maike Buchin, Stef Sijben, and Michel A. Westenberg. Model-based segmentation and classification of trajectories. *Algorithmica*, 80(8):2422–2452, 2018. doi:10.1007/s00453-017-0329-x.
7. Gill Barequet, Danny Z. Chen, Ovidiu Daescu, Michael T. Goodrich, and Jack S. Snoeyink. Efficiently approximating polygonal paths in three and higher dimensions. *Algorithmica*, 33(2):150–167, 2002. doi:10.1007/s00453-001-0096-5.
8. Karl Bringmann and Bhaskar Ray Chaudhury. Polyline simplification has cubic complexity. In *35th International Symposium on Computational Geometry (SoCG 2019)*, volume 129 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 18:1–18:16, Dagstuhl, Germany, 2019. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.SoCG.2019.18.
9. Kevin Buchin, Chenglin Fan, Maarten Löffler, Aleksandr Popov, Benjamin Raichel, and Marcel Roeloffzen. Fréchet distance for uncertain curves. In *47th International Colloquium on Automata, Languages, and Programming (ICALP 2020)*, volume 168 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 20:1–20:20, Dagstuhl, Germany, 2020. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ICALP.2020.20.
10. Kevin Buchin, Maximilian Konzack, and Wim Reddingius. Progressive simplification of polygonal curves. *Computational Geometry: Theory & Applications*, 88:101620:1–101620:18, 2020. doi:10.1016/j.comgeo.2020.101620.
11. Kevin Buchin, Maarten Löffler, Pat Morin, and Wolfgang Mulzer. Preprocessing imprecise points for Delaunay triangulation: Simplified and extended. *Algorithmica*, 61(3):674–693, 2011. doi:10.1007/s00453-010-9430-0.
12. Kevin Buchin, Maarten Löffler, Aleksandr Popov, and Marcel Roeloffzen. Uncertain curve simplification, 2021. arXiv:2103.09223.
13. Kevin Buchin, Stef Sijben, T. Jean Marie Arseneau, and Erik P. Willems. Detecting movement patterns using Brownian bridges. In *Proceedings of the 20th International Conference on Advances in Geographic Information Systems (SIGSPATIAL '12)*, pages 119–128, New York, NY, USA, 2012. ACM. doi:10.1145/2424321.2424338.
14. Maike Buchin and Stef Sijben. Discrete Fréchet distance for uncertain points, 2016. Presented at EuroCG 2016, Lugano, Switzerland. URL: http://www.eurocg2016.usi.ch/sites/default/files/paper_72.pdf [cited 2019-07-10].

15 Leizhen Cai and Mark Keil. Computing visibility information in an inaccurate simple polygon. *International Journal of Computational Geometry & Applications*, 7:515–538, 1997. doi:10.1142/S0218195997000326.

16 W. S. Chan and Francis Chin. Approximation of polygonal curves with minimum number of line segments or minimum error. *International Journal of Computational Geometry & Applications*, 6(1):59–77, 1996. doi:10.1142/S0218195996000058.

17 David H. Douglas and Thomas K. Peucker. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 10(2):112–122, 1973. doi:10.3138/FM57-6770-U75U-7727.

18 Anne Driemel, Herman Haverkort, Maarten Löffler, and Rodrigo I. Silveira. Flow computations on imprecise terrains. *Journal of Computational Geometry*, 4(1):38–78, 2013. doi:10.20382/jocg.v4i1a3.

19 William Evans, David Kirkpatrick, Maarten Löffler, and Frank Staals. Competitive query strategies for minimising the ply of the potential locations of moving points. In *Proceedings of the 29th Annual Symposium on Computational Geometry (SoCG '13')*, pages 155–164, New York, NY, USA, 2013. ACM. doi:10.1145/2462356.2462395.

20 Chris Gray, Frank Kammer, Maarten Löffler, and Rodrigo I. Silveira. Removing local extrema from imprecise terrains. *Computational Geometry: Theory & Applications*, 45(7):334–349, 2012. doi:10.1016/j.comgeo.2012.02.002.

21 Joachim Gudmundsson, Jyrki Katajainen, Damian Merrick, Cahya Ong, and Thomas Wolle. Compressing spatio-temporal trajectories. *Computational Geometry: Theory & Applications*, 42(9):825–841, 2009. doi:10.1016/j.comgeo.2009.02.002.

22 Leonidas J. Guibas, John E. Hershberger, Joseph S. B. Mitchell, and Jack S. Snoeyink. Approximating polygons and subdivisions with minimum-link paths. *International Journal of Computational Geometry & Applications*, 3(4):383–415, 1993. doi:10.1142/S0218195993000257.

23 Hiroshi Imai and Masao Iri. Computational-geometric methods for polygonal approximations of a curve. *Computer Vision, Graphics, and Image Processing*, 36(1):31–41, 1986. doi:10.1016/S0734-189X(86)80027-5.

24 Allan Jørgensen, Jeff M. Phillips, and Maarten Löffler. Geometric computations on indecisive points. In *Algorithms and Data Structures (WADS 2011)*, volume 6844 of *Lecture Notes in Computer Science (LNCS)*, pages 536–547, Berlin, Germany, 2011. Springer Berlin Heidelberg. doi:10.1007/978-3-642-22300-6_45.

25 Christian Knauer, Maarten Löffler, Marc Scherfenberg, and Thomas Wolle. The directed Hausdorff distance between imprecise point sets. *Theoretical Computer Science*, 412(32):4173–4186, 2011. doi:10.1016/j.tcs.2011.01.039.

26 Maarten Löffler. *Data Imprecision in Computational Geometry.* PhD thesis, Universiteit Utrecht, 2009. URL: https://dspace.library.uu.nl/bitstream/handle/1874/36022/loffler.pdf [cited 2019-06-15].

27 Maarten Löffler and Wolfgang Mulzer. Unions of onions: Preprocessing imprecise points for fast onion decomposition. *Journal of Computational Geometry*, 5(1):1–13, 2014. doi:10.20382/jocg.v5i1a1.

28 Maarten Löffler and Jeff M. Phillips. Shape fitting on point sets with probability distributions. In *Algorithms – ESA 2009*, number 5757 in Lecture Notes in Computer Science (LNCS), pages 313–324, Berlin, Germany, 2009. Springer Berlin Heidelberg. doi:10.1007/978-3-642-04128-0_29.

29 Maarten Löffler and Jack S. Snoeyink. Delaunay triangulations of imprecise points in linear time after preprocessing. *Computational Geometry: Theory & Applications*, 43(3):234–242, 2010. doi:10.1016/j.comgeo.2008.12.007.

30 Maarten Löffler and Marc van Kreveld. Largest and smallest tours and convex hulls for imprecise points. In *Algorithm Theory – SWAT 2006*, volume 4059 of *Lecture Notes in Computer Science (LNCS)*, pages 375–387, Berlin, Germany, 2006. Springer Berlin Heidelberg. doi:10.1007/11785293_35.

31 Avraham Melkman and Joseph O'Rourke. On polygonal chain approximation. In Godfried T. Toussaint, editor, *Computational Morphology*, volume 6 of *Machine Intelligence and Pattern Recognition*, pages 87–95. Elsevier Science Publishers, 1988. doi:10.1016/B978-0-444-70467-2.50012-6.

32 Jian Pei, Bin Jiang, Xuemin Lin, and Yidong Yuan. Probabilistic skylines on uncertain data. In *Proceedings of the 33rd International Conference on Very Large Data Bases*, pages 15–26, Los Angeles, CA, USA, 2007. VLDB Endowment. doi:10.5555/1325851.1325858.

33 Aleksandr Popov. Similarity of uncertain trajectories. Master's thesis, Eindhoven University of Technology, 2019. URL: https://research.tue.nl/en/studentTheses/similarity-of-uncertain-trajectories [cited 2019-12-18].

34 Urs Ramer. An iterative procedure for the polygonal approximation of plane curves. *Computer Graphics and Image Processing*, 1(3):244–256, 1972. doi:10.1016/S0146-664X(72)80017-0.

35 Subhash Suri, Kevin Verbeek, and Hakan Yıldız. On the most likely convex hull of uncertain points. In *Algorithms – ESA 2013*, volume 8125 of *Lecture Notes in Computer Science (LNCS)*, pages 791–802, Berlin, Germany, 2013. Springer Berlin Heidelberg. doi:10.1007/978-3-642-40450-4_67.

36 Mees van de Kerkhof, Irina Kostitsyna, Maarten Löffler, Majid Mirzanezhad, and Carola Wenk. Global curve simplification. In *27th Annual European Symposium on Algorithms (ESA 2019)*, volume 144 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 67:1–67:14, Dagstuhl, Germany, 2019. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ESA.2019.67.

37 Marc van Kreveld, Maarten Löffler, and Joseph S. B. Mitchell. Preprocessing imprecise points and splitting triangulations. *SIAM Journal on Computing*, 39(7):2990–3000, 2010. doi:10.1137/090753620.

A Shortcut Testing: Intermediate Points

In this appendix, we discuss testing a single shortcut where we fix the realisations of the first and the last uncertain point. We provided intuitive explanations in Section 4 and discuss the details here. We state some basic facts about the Hausdorff and the Fréchet distance in the precise setting (see the full version [12] for the proofs) and use them to design simple algorithms for testing shortcuts in the uncertain settings.

A.1 Hausdorff Distance

We start by recalling some useful facts about the Hausdorff distance in the precise setting.

▶ **Lemma 14.** *Given points $a, b, c, d \in \mathbb{R}^2$ forming segments ab and cd, the largest distance from one segment to the other is $\max_{p \in ab} d(p, cd) = \max\{d(a, cd), d(b, cd)\}$.*

▶ **Lemma 15.** *Given $n \in \mathbb{N}^{>0}$, for any precise curve $\pi = \langle p_1, \ldots, p_n \rangle$ with $p_i \in \mathbb{R}^2$ for all $i \in [n]$, we have $d_{\mathrm{H}}(\pi, p_1p_n) = \max_{i \in [n]} d(p_i, p_1p_n)$.*

Indecisive points. We generalise the setting to include imprecision. We first claim that the straightforward setting with indecisive points permits an easy solution using Lemma 15; the proof can be found in the full version [12].

▶ **Lemma 16.** *Given $n, k \in \mathbb{N}^{>0}$, for any indecisive curve $\mathcal{U} = \langle U_1, \ldots, U_n \rangle$ with $n \geq 3$, $U_i = \{p_i^1, \ldots, p_i^k\}$ for all $i \in [n]$ and $p_i^j \in \mathbb{R}^2$ for all $i \in [n]$, $j \in [k]$, and given some $p_1 \in U_1$ and $p_n \in U_n$, we have $\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p_1, \pi(n) \equiv p_n} d_{\mathrm{H}}(\pi, p_1p_n) = \max_{i \in \{2, \ldots, n-1\}} \max_{j \in [k]} d(p_i^j, p_1p_n)$.*

Note that this means that when the start and end realisations are fixed, we can test that a shortcut is valid using the lemma above in time $\mathcal{O}(nk)$ for a shortcut of length n.

Disks. We proceed to present the way to test shortcuts for fixed realisations of the first and the last points when the imprecision is modelled using disks. In the next arguments the following form of a triangle inequality is useful (again, see the full version [12] for details).

▶ **Lemma 17.** *For any $p, q \in \mathbb{R}^2$ and a line segment ab on $a, b \in \mathbb{R}^2$, $d(p, ab) \leq \|p - q\| + d(q, ab)$.*

▶ **Lemma 18.** *Given $n \in \mathbb{N}^{\geq 3}$, for any imprecise curve modelled with disks $\mathcal{U} = \langle U_1, \dots, U_n \rangle$ with $U_i = D(c_i, r_i)$ for all $i \in [n]$ and $c_i \in \mathbb{R}^2$, $r_i \in \mathbb{R}^{\geq 0}$ for all $i \in [n]$, and given some $p_1 \in U_1$ and $p_n \in U_n$, we have*

$$\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p_1, \pi(n) \equiv p_n} d_{\mathrm{H}}(\pi, p_1 p_n) = \max_{i \in \{2, \dots, n-1\}} \left(d(c_i, p_1 p_n) + r_i \right) .$$

Proof. Assume the setting of the lemma. We derive

$$\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p_1, \pi(n) \equiv p_n} d_{\mathrm{H}}(\pi, p_1 p_n) = \{\text{Lemma 15}\} \max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p_1, \pi(n) \equiv p_n} \max_{i \in [n]} d(\pi(i), p_1 p_n)$$
$$= \{\text{Def. } \Subset,\ d(p_1, p_1 p_n) = d(p_n, p_1 p_n) = 0\} \max_{i \in \{2, \dots, n-1\}} \max_{p \in U_i} d(p, p_1 p_n) .$$

It remains to show that $\max_{p \in U_i} d(p, p_1 p_n) = d(c_i, p_1 p_n) + r_i$ for any $i \in \{2, \dots, n-1\}$.

Pick $p' \coloneqq \operatorname{argmax}_{p \in U_i} d(p, p_1 p_n)$. Note that by Lemma 17, $d(p', p_1 p_n) \leq \|p' - c_i\| + d(c_i, p_1 p_n)$. Furthermore, as $p' \in U_i$, by definition of U_i we have $\|p' - c_i\| \leq r_i$. Thus, $\max_{p \in U_i} d(p, p_1 p_n) \leq d(c_i, p_1 p_n) + r_i$, and we need to show the other direction.

Now pick a point $q' \coloneqq \operatorname{argmin}_{q \in p_1 p_n} \|q - c_i\|$, so that $d(c_i, p_1 p_n) = \|q' - c_i\|$. Draw the line through c_i and q' and pick the point p' on that line on the boundary of U_i on the opposite side of q w.r.t. c_i. Clearly, $\|p' - c_i\| = r_i$ and $q' = \operatorname{argmin}_{q \in p_1 p_n} \|q - p'\|$. Thus,

$$d(p', p_1 p_n) = \|p' - q'\| = \|q' - c_i\| + \|p' - c_i\| = d(c_i, p_1 p_n) + r_i .$$

Note that $p' \in U_i$, so we conclude $\max_{p \in U_i} d(p, p_1 p_n) \geq d(c_i, p_1 p_n) + r_i$. ◀

This lemma allows us to test a shortcut in time $\mathcal{O}(n)$ for a shortcut of length n.

Polygonal closed convex sets.

▶ **Lemma 19.** *Given $n, k \in \mathbb{N}^{>0}$, $n \geq 3$, for any imprecise curve modelled with PCCSs $\mathcal{U} = \langle U_1, \dots, U_n \rangle$ with $U_i \subset \mathbb{R}^2$ and $V(U_i) = \{p_i^1, \dots, p_i^k\}$ for all $i \in [n]$, and given some $p_1 \in U_1$ and $p_n \in U_n$, we have*

$$\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p_1, \pi(n) \equiv p_n} d_{\mathrm{H}}(\pi, p_1 p_n) = \max_{i \in \{2, \dots, n-1\}} \max_{v \in V(U_i)} d(v, p_1 p_n) .$$

Proof. Assume the setting of the lemma. As before, derive

$$\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p_1, \pi(n) \equiv p_n} d_{\mathrm{H}}(\pi, p_1 p_n) = \{\text{Lemma 15}\} \max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p_1, \pi(n) \equiv p_n} \max_{i \in [n]} d(\pi(i), p_1 p_n)$$
$$= \{\text{Def. } \Subset,\ d(p_1, p_1 p_n) = d(p_n, p_1 p_n) = 0\} \max_{i \in \{2, \dots, n-1\}} \max_{p \in U_i} d(p, p_1 p_n) .$$

To show that the claim holds, it remains to show that for any PCCS U and a line segment ab, $\max_{p \in U} d(p, ab) = \max_{v \in V(U)} d(v, ab)$. Firstly, as $V(U) \subset U$, we immediately have $\max_{p \in U} d(p, ab) \geq \max_{v \in V(U)} d(v, ab)$. Consider any $p \in U$. We show that there is some $v \in V(U)$ such that $d(v, ab) \geq d(p, ab)$, completing the proof, with a case distinction on p.

Algorithm 1 Testing a shortcut on an indecisive curve with the Fréchet distance.

Require: $\mathcal{U} = \langle U_1, \dots, U_n \rangle$, $n, k \in \mathbb{N}^{>0}$, $\forall i \in [n] : U_i = \{p_i^1, \dots, p_i^k\}$, $\forall i \in [n], j \in [k] : p_i^j \in \mathbb{R}^2$, $\varepsilon \in \mathbb{R}^{>0}$, $p_1 \in U_1$, $p_n \in U_n$

function CheckFréchetInd$(\mathcal{U}, p_1, p_n, n, k, \varepsilon)$
  $s_1 := 1$
  for $i \in \{2, \dots, n-1\}$ **do**
    $T_i := \emptyset$
    for $j \in [k]$ **do**
      $S_i^j := \{t \in [s_{i-1}, 2] \mid \|p_i^j - p_1p_n(t)\| \leq \varepsilon\}$
      if $S_i^j = \emptyset$ **then**
        return False
      $T_i := T_i \cup \min S_i^j$
    $s_i := \max T_i$
  return True

- $p \in V(U)$. Then pick $v := p$, and we are done.
- $p \notin V(U)$, but p is on the boundary of U. Consider the vertices $v, w \in V(U)$ with $p \in vw$. Using Lemma 14, we note $\max_{q \in vw} d(q, ab) = \max\{d(v, ab), d(w, ab)\}$. W.l.o.g. suppose $d(v, ab) \geq d(w, ab)$. Then for v indeed we have $d(v, ab) \geq d(p, ab)$.
- p is in the interior of U (cannot occur for line segments). Find the point $q' := \operatorname{argmin}_{q \in ab} \|p - q\|$, so $d(p, ab) = \|p - q'\|$. Draw the line through p and q'; let p' be the point on that line on the boundary of U on the opposite side of q' w.r.t. p. Clearly, $q' = \operatorname{argmin}_{q \in ab} \|p' - q\|$, so $d(p', ab) > d(p, ab)$. Then we can find a vertex $v \in V(U)$ as in the previous cases, yielding $d(v, ab) \geq d(p', ab) > d(p, ab)$.

This covers all the cases, so the statement holds. ◀

As before, this lemma gives us a simple way to test the shortcut with fixed realisations of the first and the last points in time $\mathcal{O}(nk)$ for a shortcut of length n and PCCSs with k vertices.

A.2 Fréchet Distance

We omit results for the Fréchet distance in the precise setting here; see the full version [12].

Indecisive points. The idea is that in the precise case we can always align greedily as we move along the line segment. In this case, we also need to find the realisation for each indecisive point that makes for the 'worst' greedy choice.

▶ **Lemma 20.** *Given $n, k \in \mathbb{N}^{>0}$ and $\varepsilon \in \mathbb{R}^{>0}$, for any indecisive trajectory $\mathcal{U} = \langle U_1, \dots, U_n \rangle$ with $U_i = \{p_i^1, \dots, p_i^k\}$ for all $i \in [n]$ and $p_i^j \in \mathbb{R}^2$ for all $i \in [n]$, $j \in [k]$, and given some $p_1 \in U_1$ and $p_n \in U_n$, we have, using Algorithm 1,*

$$\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p_1, \pi(n) \equiv p_n} d_{\mathrm{F}}(\pi, p_1p_n) \leq \varepsilon \iff \textit{CheckFréchetInd}(\mathcal{U}, p_1, p_n, n, k, \varepsilon) = \text{True}\,.$$

Proof. First, assume that $\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p_1, \pi(n) \equiv p_n} d_{\mathrm{F}}(\pi, p_1p_n) \leq \varepsilon$. In the algorithm, we compute some set S_i^j for each p_i^j and then pick one value from it and add it to T_i; from T_i we then pick a single value as s_i. So, $s_i \in S_i^j$ for some $j_i \in [k]$, on every iteration $i \in \{2, \dots, n-1\}$. Consider a realisation $\pi \Subset \mathcal{U}$ with $\pi(1) \equiv p_1$, $\pi(n) \equiv p_n$, and $\pi(i) \equiv p_i^{j_i}$ for every $i \in \{2, \dots, n-1\}$, where j_i is chosen as the value corresponding to s_i. Then we know $d_{\mathrm{F}}(\pi, p_1p_n) \leq \varepsilon$. So, there is an alignment that can be given as a sequence of n positions, $t_i \in [1, 2]$, such that $\|\pi(i) - p_1p_n(t_i)\| \leq \varepsilon$ and $t_i \leq t_{i+1}$ for all i. The alignment is established by interpolating between the consecutive points on the curves (see Section 2).

We now show by induction that $s_i \le t_i$ for all i. For $i = 2$, we get, for the chosen j_2, $s_2 \coloneqq \min\{t \in [1,2] \mid \|p_2^{j_2} - p_1p_n(t)\| \le \varepsilon\}$. As we have $t_2 \in \{t \in [1,2] \mid \|p_2^{j_i} - p_1p_n(t)\| \le \varepsilon\}$, we get $s_2 \le t_2$. Now assume the statement holds for some i, then for $i+1$ we get $s_{i+1} \coloneqq \min\{t \in [s_i, 2] \mid \|p_{i+1}^{j_{i+1}} - p_1p_n(t)\| \le \varepsilon\}$; we can rephrase this so that

$$s_{i+1} \stackrel{\text{def}}{=} \min\left(\{t \in [1,2] \mid \|p_{i+1}^{j_{i+1}} - p_1p_n(t)\| \le \varepsilon\} \cap [s_i, 2]\right).$$

So, there are two options.

- $s_{i+1} = s_i$. Then we know $s_{i+1} = s_i \le t_i \le t_{i+1}$.
- $s_{i+1} > s_i$. Then we can use the same argument as for $i = 2$ to find that $s_{i+1} \le t_{i+1}$.

Now we know that for every i, $t_i \in S_i^{j_i}$ for the choice of j_i described above. Therefore, for any $p_{i+1}^{j_{i+1}}$ there is always a realisation prefix such that any valid alignment has $t_{i+1} \ge s_i$; as we know that there is a valid alignment for every realisation, we conclude that every S_i^j is non-empty. Thus, the algorithm returns True.

Now assume that the algorithm returns True. Consider any realisation $\pi \Subset \mathcal{U}$. We claim that there is a valid alignment, described with a sequence of $t_i \in [1,2]$ for $i \in \{2, \dots, n-1\}$, such that $s_{i-1} \le t_i \le s_i$ and $\|p_1p_n(t_i) - \pi(i)\| \le \varepsilon$. Denote the realisation $\pi \stackrel{\text{def}}{=} \langle p_1, p_2^{j_2}, p_3^{j_3}, \dots, p_{n-1}^{j_{n-1}}, p_n\rangle$, so the sequence $\langle j_2, \dots, j_{n-1}\rangle$ describes the choices of the realisation. Consider the set $S_i^{j_i}$ for any $i \in \{2, \dots, n-1\}$. We know that it is non-empty, otherwise the algorithm would have returned False. We claim that we can pick $t_i = \min S_i^{j_i}$ for every i. By definition, $S_i^{j_i} \subseteq [1,2]$ and $\|p_1p_n(t_i) - \pi(i)\| \le \varepsilon$. We also trivially get that $s_{i-1} \le t_i$. Finally, note that $t_i \in T_i$, and $s_i \coloneqq \max T_i$, so $t_i \le s_i$.

This argument shows that $t_i \le t_{i+1}$ for every i, and that $\|p_1p_n(t_i) - \pi(i)\| \le \varepsilon$. Therefore, $d_{\mathrm{F}}(\pi, p_1p_n) \le \varepsilon$. As this works for any realisation with $\pi(1) \equiv p_1$ and $\pi(n) \equiv p_n$, we conclude $\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p_1, \pi(n) \equiv p_n} d_{\mathrm{F}}(\pi, p_1p_n) \le \varepsilon$. ◀

Disks. To show the generalisation to disks, it is helpful to reframe the problem as that of disk stabbing for appropriate disks. We state some useful facts first (see the full version [12]).

▶ **Lemma 21.** *Given a disk $D_1 \coloneqq D(c, r)$ with $c \in \mathbb{R}^2$, $r \in \mathbb{R}^{\ge 0}$, a threshold $\varepsilon \in \mathbb{R}^{>0}$, and a point $p \in \mathbb{R}^2$, define $D_2 \coloneqq D(c, \varepsilon - r)$. We have $\max_{p' \in D_1} \|p - p'\| \le \varepsilon \iff p \in D_2$.*

▶ **Lemma 22.** *Given a disk $D_1 \coloneqq D(c, r)$ with $c \in \mathbb{R}^2$, $r \in \mathbb{R}^{\ge 0}$, $\varepsilon \in \mathbb{R}^{>0}$, and a line segment pq with $p, q \in \mathbb{R}^2$, define $D_2 \coloneqq D(c, \varepsilon - r)$. Then $\max_{p' \in D_1} d(p', pq) \le \varepsilon \iff pq \cap D_2 \ne \emptyset$.*

Algorithm 2 Testing a shortcut on an imprecise curve modelled with disks with the Fréchet distance.

Require: $\mathcal{U} = \langle U_1, \dots, U_n\rangle$, $n \in \mathbb{N}^{>0}$, $\forall i \in [n] : U_i = D(c_i, r_i)$, $\forall i \in [n] : c_i \in \mathbb{R}^2, r_i \in \mathbb{R}^{\ge 0}$, $\varepsilon \in \mathbb{R}^{>0}$, $p_1 \in U_1$, $p_n \in U_n$

```
function CheckFréchetDisks(U, p_1, p_n, n, ε)
    s_1 := 1
    for i ∈ {2, ..., n − 1} do
        S_i := {t ∈ [s_{i−1}, 2] | ‖c_i − p_1p_n(t)‖ ≤ ε − r_i}
        if S_i = ∅ then
            return False
        s_i := min S_i
    return True
```

▶ **Lemma 23.** *Given $n \in \mathbb{N}^{>0}$ and $\varepsilon \in \mathbb{R}^{>0}$, for any imprecise curve modelled with disks $\mathcal{U} = \langle U_1, \dots, U_n \rangle$ with $U_i = D(c_i, r_i)$ for all $i \in [n]$ and $c_i \in \mathbb{R}^2$, $r_i \in \mathbb{R}^{\geq 0}$ for all $i \in [n]$, and given some $p_1 \in U_1$ and $p_n \in U_n$, we have, using Algorithm 2,*

$$\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p_1, \pi(n) \equiv p_n} d_{\mathrm{F}}(\pi, p_1 p_n) \leq \varepsilon \iff \textsc{CheckFréchetDisks}(\mathcal{U}, p_1, p_n, n, \varepsilon) = \mathrm{True}\,.$$

Proof. We use Lemma 22 to change the problem: rather than establishing an alignment that comes in the correct order and satisfies the distance constraints, we do disk stabbing and pick the stabbing points in the correct order. So, we have $\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p_1, \pi(n) \equiv p_n} d_{\mathrm{F}}(\pi, p_1 p_n) \leq \varepsilon$ if and only if there exists a sequence of points $p'_i \in p_1 p_n \cap D(c_i, \varepsilon - r_i)$ for all $i \in \{2, \dots, n-1\}$ such that $p'_i \preccurlyeq p'_{i+1}$ along $p_1 p_n$ for all $i \in \{2, \dots, n-2\}$. We show that this is exactly what Algorithm 2 computes in the full version [12]. ◀

Algorithm 3 Testing a shortcut on an imprecise curve modelled with PCCSs with the Fréchet distance.

Require: $\mathcal{U} = \langle U_1, \dots, U_n \rangle$, $n, k \in \mathbb{N}^{>0}$, $\forall i \in [n] : U_i$ is a PCCS, $V(U_i) = \{p_i^1, \dots, p_i^k\}$, $\forall i \in [n], j \in [k] : p_i^j \in \mathbb{R}^2$, $\varepsilon \in \mathbb{R}^{>0}$, $p_1 \in U_1$, $p_n \in U_n$

function $\textsc{CheckFréchetPCCS}(\mathcal{U}, p_1, p_n, n, k, \varepsilon)$
  $s_1 := 1$
  for $i \in \{2, \dots, n-1\}$ **do**
    $T_i := \emptyset$
    for $j \in [k]$ **do**
      $S_i^j := \{t \in [s_{i-1}, 2] \mid \|p_i^j - p_1 p_n(t)\| \leq \varepsilon\}$
      if $S_i^j = \emptyset$ **then**
        return False
      $T_i := T_i \cup \min S_i^j$
    $s_i := \max T_i$
  return True

Polygonal closed convex sets.

▶ **Lemma 24.** *Given $n, k \in \mathbb{N}^{>0}$ and $\varepsilon \in \mathbb{R}^{>0}$, for any imprecise curve modelled with PCCSs $\mathcal{U} = \langle U_1, \dots, U_n \rangle$ with $U_i \subset \mathbb{R}^2$ and $V(U_i) = \{p_i^1, \dots, p_i^k\}$ for all $i \in [n]$, and given some $p_1 \in U_1$ and $p_n \in U_n$, we have, using Algorithm 3,*

$$\max_{\pi \Subset \mathcal{U}, \pi(1) \equiv p_1, \pi(n) \equiv p_n} d_{\mathrm{F}}(\pi, p_1 p_n) \leq \varepsilon \iff \textsc{CheckFréchetPCCS}(\mathcal{U}, p_1, p_n, n, k, \varepsilon) = \mathrm{True}\,.$$

Proof. As we have shown in Lemma 19, it suffices to test the vertices of a PCCS to establish that the distance from every point to the line segment is below the threshold. It remains to show that the extreme alignment (in terms of ordering) for the Fréchet distance is also achieved at a vertex. This case then becomes identical to the indecisive points case, so we can use Lemma 20 to show correctness.

Consider an arbitrary point $t \in U_i$ and let s be the earliest point in the ε-disk around t that is on pq. Clearly, if t is in the interior of U_i, then we can take any t' on the line through t parallel to pq and get the corresponding s' with $s \prec s'$. So, assume t is on the boundary of U_i. Suppose that $t \in uv$ with $u, v \in V(U_i)$. Rotate and translate the coordinate plane so that pq lies on the x-axis. Derive the equation for the line containing uv, say, $y' = kx' + b$. First consider $k = 0$, so the line containing uv is parallel to the line containing pq. In this case, clearly, moving along uv in the direction coinciding with the direction from p to q

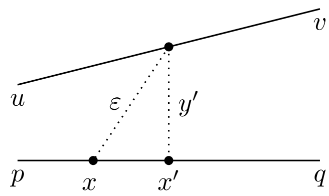

Figure 7 Illustration for the computation in Lemma 24.

increases the x-coordinate of point of interest, so moving to a vertex is optimal. Now assume $k > 0$. If $k < 0$, reflect the coordinate plane about $y = 0$. Geometrically, it is easy to see (Figure 7) that the coordinate of interest can be expressed as

$$x = x' - \sqrt{\varepsilon^2 - y'^2} = \frac{y' - b}{k} - \sqrt{\varepsilon^2 - y'^2} .$$

We want to maximise x by picking the appropriate y'. We take the derivative: $\mathrm{d}x/\mathrm{d}y' = 1/k + y'/\sqrt{\varepsilon^2 - y'^2}$. We equate it to 0 to find the critical point of the function, $y'_0 = -\varepsilon/\sqrt{k^2 + 1}$. We can check that for $y' < y'_0$, the value of the derivative is negative, and for $y' > y'_0$ it is positive, so at $y' = y'_0$ we achieve a local minimum. There are no other critical points, so to maximise x, we want to move away from the local minimum as far as possible. As we are limited to the line segment uv, the maximum is achieved at an endpoint. ◀