Geometric k Shortest Paths*

Sylvester Eriksson-Bique† John Hershberger‡ Valentin Polishchuk§
Bettina Speckmann¶ Subhash Suri‖ Topi Talvitie** Kevin Verbeek‖ Hakan Yıldız‖

Abstract

We consider the problem of computing k shortest paths in a two-dimensional environment with polygonal obstacles, where the jth path, for $1 \leq j \leq k$, is the shortest path in the free space that is also *homotopically* distinct from each of the first $j-1$ paths. In fact, we consider a more general problem: given a source point s, construct a partition of the free space, called the *kth shortest path map* (k-SPM), in which the homotopy of the kth shortest path in a region has the same structure. Our main combinatorial result establishes a tight bound of $\Theta(k^2h + kn)$ on the worst-case complexity of this map. We also describe an $O((k^3h + k^2n)\log(kn))$ time algorithm for constructing the map. In fact, the algorithm constructs the jth map for every $j \leq k$. Finally, we present a simple visibility-based algorithm for computing the k shortest paths between two fixed points. This algorithm runs in $O(m \log n + k)$ time and uses $O(m + k)$ space, where m is the size of the visibility graph. This latter algorithm can be extended to compute k shortest *simple* (non-self-intersecting) paths, taking $O(k^2m(m + kn)\log(kn))$ time.

We invite the reader to play with our applet demonstrating k-SPMs [10].

*B. Speckmann and K. Verbeek were partially supported by the Netherlands' Organisation for Scientific Research (NWO) under project nos. 639.022.707 and 639.023.208. S. Eriksson-Bique was supported as a Graduate Student Fellow by the National Science Foundation grant no. DGE-1342536. S. Eriksson-Bique, V. Polishchuk and T. Talvitie were supported by the Academy of Finland grant 1138520 and University of Helsinki Research Funds. The research of Subhash Suri, Kevin Verbeek and Hakan Yildiz was partially supported by the NSF grant CCF-1161495.

†Courant Institute, NYU. ebs@cims.nyu.edu

‡Mentor Graphics Corporation. john_hershberger@mentor.com

§Communications and Transport Systems, ITN, Linköping University. firstname.lastname@liu.se

¶Dept. of Mathematics and Computer Science, TU Eindhoven. b.speckmann@tue.nl

‖Computer Science, University of California Santa Barbara. [suri|kverbeek|hakan]@cs.ucsb.edu

**Computer Science, University of Helsinki. firstname.lastname@cs.helsinki.fi

1 Introduction

In many applications of mathematical optimization, several "good" solutions are more desirable than a single optimum. This happens because a mathematical model is an imperfect formulation of complex reality, and its various constraints and objectives are only an approximation of the desired goal. Optimization problems are also typically part of a larger system with many interacting parts, where optimal solutions of different parts may be incompatible. In these settings, the system designer must find sub-optimal but high-quality solutions for each part to construct the overall solution. Motivated by these considerations, there is a long history of research on finding k best solutions for discrete optimization problems, including spanning trees and shortest paths in graphs [6, 9, 12, 19].

In this paper, we investigate the fundamental problem of computing k distinct shortest paths among polygonal obstacles in the plane. Because geometric shortest paths live in a continuous (free) space, we need a topological condition on paths to ensure that different paths are non-trivially distinct: otherwise, we can create many nearly identical shortest paths by adding infinitesimal "bumps" to the primary shortest path. The most natural condition is to require paths to have different *homotopy*, where two paths are said to be homotopically equivalent if they can be deformed into each other within the free space of obstacles. Intuitively, two paths are homotopically distinct if they lie on different sides of some obstacle. Multiple shortest paths of distinct homotopies naturally capture the high-level design criteria in geometric environments: e.g., in VLSI design or printed circuit board routing, where obstacles are electronic components, in robot path planning, where obstacles are physical obstructions, in air traffic management, where obstacles model hazardous weather or no-fly zones, etc.

We consider a more general form of the problem: given a source point s, construct a map of the entire free space, partitioning it into equivalence class regions such that the kth shortest path from s to any point in a single region has the same structure. We call this map the *kth shortest path map* (or k-SPM for short). With

the k-SPM, one can compute the kth shortest path to any target easily. The following paragraph describes the key results of our paper.

Our Results We prove that the edges of the k-SPM are $O(k^2h + kn)$ linear or hyperbolic arcs, and give a construction showing that this bound is tight in the worst case (Section 4). We present an $O((k^3h+k^2n)\log(kn))$ time algorithm (Section 5), using the continuous Dijkstra paradigm, for constructing the map. The algorithm computes the jth shortest path map for all $1 \leq j \leq k$. Taking this into account, the algorithm is output sensitive: its running time is $O(\log(kn))$ times the total complexity of the first k shortest path maps. By preprocessing the k-SPM for point-location queries, we can answer kth shortest path queries in $O(\log(kn))$ time; if we want to report (in implicit form) all k shortest paths, the preprocessing time remains the same, but the storage and query time both increase by a factor of k. (If the paths are to be reported explicitly, the complexity of the paths must be added to the query time.) In Section 6, we also present a simpler, albeit asymptotically worse, algorithm for computing the kth shortest path between two fixed points based on the visibility graph. This algorithm runs in $O(m \log n+k)$ time and uses $O(m+k)$ space, where m is the size of the visibility graph. One advantage of this latter algorithm is that it can also be extended to find the kth *simple* (non-self-intersecting) path, taking $O(k^2m(m + kn)\log kn)$ time.

Related work Finding shortest paths is also a central problem in the study of graph algorithms. Apart from finding the shortest path itself, considerable attention has been paid to computing its various alternatives including the second, third, and in general kth shortest path between two nodes in a graph; see, e.g., [6, 9] and references therein. On the other hand, *geometric* kth shortest paths have not been explored before. (One problem for which both the graph and the geometric versions were considered is finding the k smallest spanning trees [4, 5].)

In [16] Mitchell surveys many variations of the geometric shortest path problem; for some recent work see [2, 3]. In addition to computing one shortest path to a single target point, a lot of attention in the literature has been devoted to building shortest path *maps*—structures supporting efficient shortest-path queries. A shortest path map can be viewed as the Voronoi diagram of vertices of the domain, where each vertex is (additively) weighted by the shortest-path distance from the source s [11]. Our study of "kth shortest path maps" benefits from notions introduced by Lee [13] for *higher-order* Voronoi diagrams: when bounding the complexity of the maps in Section 4, we employ Lee's ideas to define "old" and "new" features of the map and to derive relationships between them. Higher-order Voronoi diagrams have been recently reexamined in [1, 14, 15, 17]; in particular, [14] considered geodesic diagrams in polygonal domains. Perhaps unsurprisingly, the complexity of our kth shortest path map differs from that of an order-k geodesic Voronoi diagram; the major difference is that homotopies are irrelevant for Voronoi diagrams, but are central in our work.

2 Preliminaries

We are given a closed polygonal domain P with n vertices and h holes; the holes are also called "obstacles" and the domain is called the "free space." We assume that no three vertices of P are collinear and make other general position assumptions below, as needed. We are also given a source point $s \in P$; unless otherwise stated, all paths will have s as an endpoint. For a point $p \in P$, two paths to p are *homotopically equivalent* if one can be continuously deformed to the other while staying within P. Homotopically equivalent paths form an equivalence class (the *homotopy class*) in the set of s–p paths. A path is *locally shortest* if its length cannot be reduced by an infinitesimal perturbation of the path (i.e., a pulled-taut path).

LEMMA 2.1. ([8]) *A locally shortest path is the unique shortest path of its homotopy class. Furthermore, all bends of a locally shortest path are at vertices of P and turn toward the corresponding obstacles.*

Let $d(p)$ denote the shortest-path (geodesic) distance from s to p. A vertex v of P is a *predecessor* of p if segment $\overline{vp}$ is in free space and $d(p) = d(v)+|\overline{vp}|$. The *shortest path map* of P (or SPM for short) is the partitioning of P such that all points within the same cell of the SPM have the same unique predecessor. The edges of the partition are called *bisectors*; points on bisectors have more than one predecessor. We distinguish between two types of bisectors: *walls* and *windows* (Fig. 1). A bisector is a wall if, for a point p on the bisector, there exist two homotopically different paths to p with length $d(p)$. If there is a unique shortest path to a point p on a bisector, then this bisector is a window; any point p on a window has two predecessors that are collinear with p. We assume that there is a unique shortest path to each vertex of P, and that there are at most three homotopically different shortest paths to each point in P. The former assumption implies that walls are 1-dimensional curves. The endpoints of a wall are either at an obstacle or at a *triple point*, where three walls meet. Windows start at vertices of P and extend until an obstacle or wall is hit. Intuitively, windows can mostly be ignored as far as homotopy types are con-

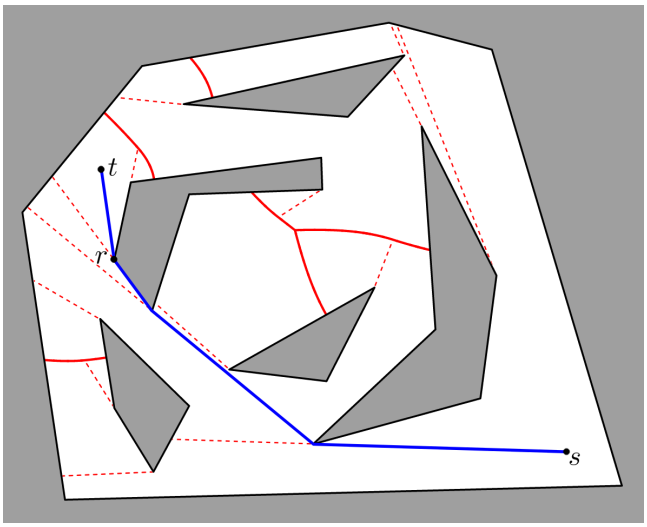


Figure 1: An SPM; bisectors are red (windows are dashed and walls are solid). The shortest s-t path (in blue) reaches t via its predecessor r.

cerned; walls, by contrast, are central to our study. By using standard point location structures on the SPM of P, one can query the shortest path length to any point in P in $O(\log n)$ time and, in addition, report the path in linear output sensitive time [11]. Our goal is to compute a similar structure for kth shortest paths.

We now introduce the subject of our study. For a point $p \in P$, let $H(p)$ denote the set of locally shortest paths from s to p of all possible homotopy types.

DEFINITION 2.1. *A path $\pi \in H(p)$ is a* kth shortest path *(or is a* k-path*) to p if there are exactly $k-1$ shorter paths in $H(p)$.*

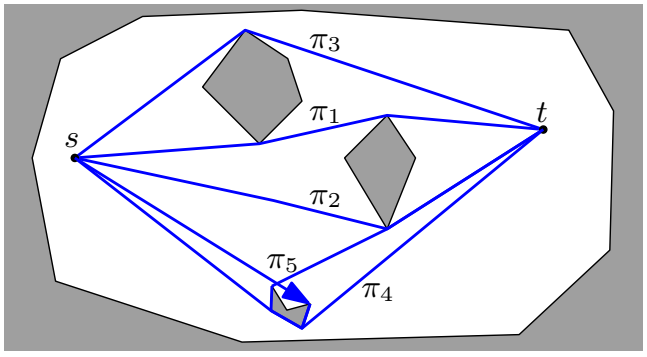


Figure 2: $|\pi_1|<|\pi_2|=|\pi_3|<|\pi_4|<|\pi_5|$. π_1 is the shortest path to t (a 1-path; cf. Def. 2.1), each of π_2 and π_3 is a 2-path, π_4 is a 4-path, π_5 is a 5-path (π_5 is nonsimple—it has a loop going clockwise around the hole).

Figure 2 illustrates the definition. We denote the length of the k-path(s) to p by $d_k(p)$. Notice that, under these definitions, the term 1-path is synonymous with "shortest path" and $d(p) = d_1(p)$.

In order to extend the map concept to k-paths, we first generalize the definition of a predecessor. Let v be an obstacle vertex and i be an integer between 1 and k. For a point p in the plane, the pair (v, i) is a *k-predecessor* of p if the segment $\overline{vp}$ is in free space and $d_k(p) = d_i(v) + |\overline{vp}|$. This implies that a k-path to p can be obtained by concatenating the segment $\overline{vp}$ with the i-path to v. As with the SPM, we assume that each obstacle vertex has a unique i-path for any i, and that there are at most three i-paths in $H(p)$ for each point $p \in P$. Interestingly, for $i > 1$, the former assumption does not follow from a general position assumption. We discuss this issue more in the final version. For the sake of simplicity, we will ignore the issue in this paper and stick to the assumption above.

Observe that, given the k-predecessors of all points in the plane and the i-predecessors of all obstacle vertices for $1 \le i \le k$, one can construct the k-path to any given point p. The *kth shortest path map* (or k-SPM for short) of P is a subdivision of P into cells such that all points within the same cell have the same unique k-predecessor. In order to construct k-paths from the k-SPM, we also assume that it stores the i-predecessors of all vertices, for all $1 \le i \le k$. As with the SPM, one can use standard point location structures to report the k-path length of a query point in $O(\log C_k)$ time, where C_k is the complexity of the k-SPM.

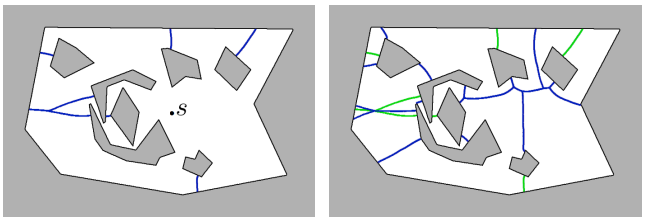
To distinguish the different types of bisectors that form the boundaries of the k-SPM, we generalize the definitions of walls and windows as follows:

DEFINITION 2.2. *A point p is on a k-wall if $H(p)$ contains at least two k-paths.*

DEFINITION 2.3. *A point p is on a k-window if $H(p)$ contains exactly one k-path and p has two k-predecessors.*

Note that the two predecessors of a point p on a k-window must be collinear with p. Furthermore, by the definition of k-paths, a point cannot be on a k-wall and a $(k+1)$-wall at the same time (if a point has two k-paths, then it has no $(k+1)$-path). Similarly to walls in the SPM, k-walls have their endpoints either on obstacles or at triple points, where three k-walls meet. In Section 3, we show that edges of the k-SPM are $(k-1)$-walls, k-walls and k-windows. We also show that our assumption that a k-predecessor is of the form (v, i) with $1 \le i \le k$ is indeed correct.

Figure 3 shows examples of walls in k-SPMs.



(a) walls of 1-SPM (b) walls of 2-SPM

(c) walls of 3-SPM (d) walls of 4-SPM

Figure 3: Walls in k-SPMs.

3 Structural results

Consider a path π from s to some target $t \in P$. This path crosses several walls (1-walls, 2-walls, etc.) in P. We define the *crossing sequence* of π as the sequence of positive integers that represents all the i-walls crossed by this path going back from t to s. That is, if π crosses an i-wall, we add i to the sequence. Although it is not strictly necessary, we generally assume an upper bound on the sequence values (the maximum wall class), so that the sequence is finite. We call a sequence a *k-sequence* if it adheres to the following inductive definition:

- A 1-sequence does not contain 1.
- A k-sequence contains $(k-1)$, the first $(k-1)$ occurs before the first k, and the tail of the sequence after the first $(k-1)$ is a $(k-1)$-sequence.

We need the following property of k-sequences, whose proof appears with other omitted proofs in the final version.

LEMMA 3.1. *A sequence σ cannot be both a k-sequence and an ℓ-sequence if $k \neq \ell$.*

The relation between k-sequences and k-paths is summarized in the following lemma.

LEMMA 3.2. *A locally shortest path π is a k-path if and only if its crossing sequence is a k-sequence.*

Proof. We first show that the crossing sequence of a k-path π is a k-sequence. Let us assume that distances have been scaled so that the length of π is 1. Define $p(x)$ for $0 \leq x \leq 1$ as the point on π such that the distance from t to $p(x)$ along π is x. Let $\gamma(x)$ be the subpath of π from $p(x)$ to t. For any $i \geq 1$, let π_i denote the i-path to t ($\pi = \pi_k$). (We assume that t is not on an i-wall, for any $1 \leq i \leq k$.) The concatenation of π_i and $\gamma(x)$ is a path from s to $p(x)$, via t; let $\pi'_i(x)$ denote the shortest path of this homotopy class (Fig. 4). All paths $\pi'_i(x)$ must have different homotopy classes for different i.

Let $l_i(x)$ be the length of $\pi'_i(x)$; clearly l_i is continuous. By the definition of k-paths, $l_i(0) \leq l_j(0)$ for $i < j$. On the other hand, $l_k(1) = 0$ and $l_i(1) > 0$ for $i \neq k$. Note that as x grows from 0 to 1, $l_k(x)$ decreases not slower than any other $l_i(x)$, $i \neq k$. Thus, the graph of $l_k(x)$ crosses the graphs of all $l_i(x)$ for $i < k$ exactly once, but no other graphs (Fig. 5).

The proof proceeds by induction. A point $p(x)$ is on a j-wall if $l_k(x)$ crosses some other graph at x, and there are exactly $j-1$ graphs that pass below this intersection. Clearly, if $k = 1$, the path π_k cannot cross a 1-wall, since $l_k(x)$ cannot intersect anything. For $k > 1$, the first intersection of $l_k(x)$ must be with a graph $l_i(x)$ with $i < k$, as described above. This means that $p(x)$ must cross a $(k-1)$-wall before crossing a k-wall. After the $(k-1)$-wall at $x = x^*$, the path $\pi'_k(x^*)$ is the $(k-1)$-path to $p(x)$. By induction, the remainder of the crossing sequence must be a $(k-1)$-sequence.

Finally note that a locally shortest path π must be an i-path for some $i \geq 1$. If the crossing sequence of π is a k-sequence, then it cannot be an i-sequence for $i \neq k$ by Lemma 3.1. Thus $i = k$, and π is a k-path.

Lemma 3.2 implies that a k-path from s to t crosses walls "in order": it crosses a 1-wall, then a 2-wall, etc., until it crosses a $(k-1)$-wall, after which it reaches t. Also, any prefix of the k-path is an i-path if it crosses the $(i-1)$-wall and not the i-wall. This property of k-paths inspires the construction of a "parking garage" obtained by "stacking" k copies (or *floors*) of P on top of each

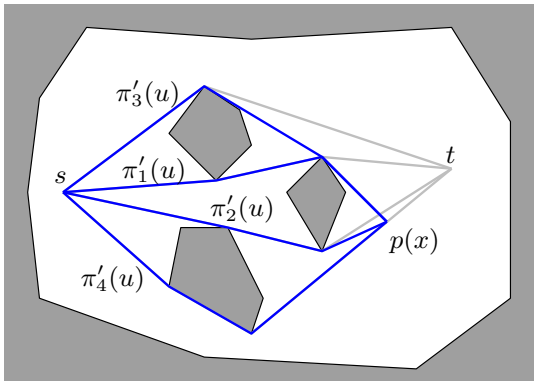


Figure 4: $\pi'_i(x)$ is the shortest path to $p(x)$, homotopically equivalent to s–π_i–t–$p(x)$ ($k = 4$).

other and gluing them along i-walls, for $1 \le i \le k$. To be precise, the *k-garage* is inductively defined as follows:

> The 1-garage is simply P. The $(k+1)$-garage can be obtained by adding a copy of P (the *$(k+1)$-floor*) on top of the k-garage. We cut both the k-floor of the k-garage and the $(k+1)$-floor along k-walls. We then glue one side of a k-wall on the k-floor to the opposite side of the same k-wall on the $(k+1)$-floor, and vice versa, to obtain the $(k+1)$-garage.

The k-garage resembles a covering space of P. However, due to the triple points formed by the i-walls ($i<k$), the k-garage is technically not a covering space, but something that is known as a ramified cover. Nonetheless, each path π in the garage can be projected down to a unique path $\pi^{\downarrow}$ in P. The next lemma relates the k-SPM of P to the SPM of the k-garage.

LEMMA 3.3. *If π is the shortest path in the k-garage from s on the 1-floor to some t on the k-floor, then $\pi^{\downarrow}$ is a k-path to t.*

Proof. We show that the crossing sequence of $\pi^{\downarrow}$ is a k-sequence. Then, by Lemma 3.2, $\pi^{\downarrow}$ is a k-path. We use the property that every tail of a k-sequence is an i-sequence for some $i \le k$. If, going back from t to s, π only goes "down" in the k-garage, then it is easy to see that the crossing sequence of $\pi^{\downarrow}$ is a k-sequence. (Because regions on the i-floor are bounded by $(i-1)$- and i-walls, π enters the i-floor by crossing an i-wall and does not cross any i-wall before it exits the i-floor by crossing an $(i-1)$-wall. Thus the tail of π's crossing sequence that starts from any point on the i-floor is an i-sequence.) For the sake of contradiction, assume that π also goes up in the k-garage. Then there must be a point where π goes up to some i-floor, and then goes monotonically down to the 1-floor. The crossing sequence of the corresponding subpath of $\pi^{\downarrow}$ must be of the form $\sigma = (i-1, \sigma_i)$, where σ_i is an i-sequence. If σ is a j-sequence for $j \neq i$, then σ_i must be a j-sequence, which is not possible by Lemma 3.1. If σ is an i-sequence, then σ_i must be an $(i-1)$-sequence, which again is not possible by Lemma 3.1. Finally note that σ must be a j-sequence for some j, since $\pi^{\downarrow}$ is locally shortest. Thus, π only goes down in the k-garage, and the crossing sequence of $\pi^{\downarrow}$ must be a k-sequence.

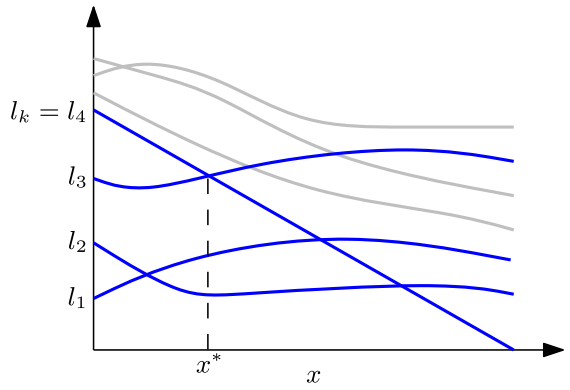


Figure 5: l_k is kth smallest at $x = 0$ and decreases faster than any other l_i ($k = 4$).

Lemma 3.3 directly implies that the SPM on the k-floor of the k-garage is exactly the k-SPM of P. Thus, as claimed before, the edges of the k-SPM consist of $(k-1)$-walls, k-walls, and k-windows. Furthermore, the k-predecessor of a point $p \in P$ must be (v, i) for some i between 1 and k.

4 The complexity of the k-SPM

To obtain an upper bound on the complexity of the k-SPM, we consider a sparser partitioning of P. We define the ($\le k$)-SPM of P as the partitioning induced by only the k-walls of P. Let $H_k(p)$ be the set of the k shortest homotopy classes to $p \in P$. We refer to $H_k(p)$ as the *k-homotopy set* of p. We would like to claim that the set $H_k(p)$ is constant within each cell of the ($\le k$)-SPM. Unfortunately we cannot claim this, since the homotopy classes of paths with different endpoints cannot be compared. To overcome this technicality, we define $H_k(p) \oplus \pi$ as the set of homotopy classes obtained by concatenating each path in $H_k(p)$ with π. If π is a path between p and p', then we can directly compare $H_k(p) \oplus \pi$ and $H_k(p')$.

LEMMA 4.1. *If p and p' lie in the same cell of the ($\le k$)-SPM, and π is a path between p and p' that does not cross a k-wall, then $H_k(p) \oplus \pi = H_k(p')$.*

To keep the notation simple, we simply compare $H_k(p)$ and $H_k(p')$ directly, in which case we really mean that we compare $H_k(p) \oplus \pi$ and $H_k(p')$, where π is the shortest path in P between p and p'. Note that π can cross a k-wall. We need the following property of the ($\le k$)-SPM.

LEMMA 4.2. *The cells of the ($\le k$)-SPM are simply connected.*

We now count the number of k-walls, starting with the case $k = 1$. Let F_1, V_1, and B_1 be the number of faces, triple points, and 1-walls of the (≤ 1)-SPM, respectively. It is easy to see that the (≤ 1)-SPM is simply connected, hence $F_1 = 1$. Now consider the graph G in which each node corresponds to either a

hole (including the outer polygon) or a triple point, and there is an edge between two nodes if there is a 1-wall between the corresponding holes/triple points. Since the ($\leq$1)-SPM is simply connected, G must be a tree. Hence $B_1 = h + V_1$. (The number of polygons bounding P is $h+1$.) Furthermore note that the degree of a triple point in G is three, and every node in G has degree at least one. So, by double counting, $2B_1 \geq 3V_1 + h + 1$ or $V_1 \leq h - 1$. To summarize, $F_1 = 1$, $V_1 \leq h - 1$, and $B_1 = h + V_1$.

To bound the complexity of the ($\leq k$)-SPM for $k > 1$, we relate its features to those of the ($\leq k-1$)-SPM. We consider an in-place transformation of the ($\leq k-1$)-SPM into the ($\leq k$)-SPM. We use lower-case letters $a, b, c, \ldots$ to denote the members of $H_k(p)$. Each k-wall of the ($\leq k$)-SPM locally separates regions of P that differ in exactly one of their k shortest path homotopy classes. Note that a k-wall e of the ($\leq k$)-SPM is not present in the ($\leq k+1$)-SPM: if the k-homotopy sets belonging to the two sides of e are $H \cup a$ and $H \cup b$, with $a \neq b$, then the $(k+1)$-homotopy set of points in the neighborhood of e is uniformly $H \cup \{a, b\}$.

The triple points of the ($\leq k$)-SPM fall into two classes, which we call *new* and *old* (borrowing the terms from [13]). If the three k-homotopy sets in the vicinity of a triple point p are $H \cup a$, $H \cup b$, and $H \cup c$, with a, b, and c all distinct, then p is a new triple point. On the other hand, if the three k-homotopy sets are $H \cup \{a, b\}$, $H \cup \{b, c\}$, and $H \cup \{a, c\}$, with a, b, and c all distinct, then p is an old triple point. These names highlight the difference between what happens in the vicinity of p in the ($\leq k+1$)-SPM. If p is a new triple point in the ($\leq k$)-SPM, then it becomes an old triple point in the ($\leq k+1$)-SPM. The three $(k+1)$-walls incident to p in the ($\leq k+1$)-SPM separate points with $(k+1)$-homotopy sets $(H \cup a) \cup b$ from $(H \cup a) \cup c$, $(H \cup b) \cup a$ from $(H \cup b) \cup c$, and $(H \cup c) \cup a$ from $(H \cup c) \cup b$. If p is an old triple point in the ($\leq k$)-SPM, then the $(k+1)$-homotopy set of points in the neighborhood of e is uniformly $H \cup \{a, b, c\}$, and hence p is in the interior of a face of the ($\leq k+1$)-SPM. See Fig. 6.

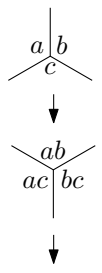


Figure 6: Life cycle of a triple point.

To transform the ($\leq k$)-SPM to the ($\leq k+1$)-SPM, we consider shortest distances to points in each face f of the ($\leq k$)-SPM from its k-walls. The distances from a particular k-wall e are measured according to the homotopy class belonging to the face on the opposite side of e from f. More concretely, let $p \in f$ be a point close to e, and let p' be on the other side of f. Then the shortest paths measured from e use the homotopy class $h_f(e) = H_k(p') \setminus H_k(p)$. For every point $q \in f$, we identify the k-wall e whose homotopy class $h_f(e)$ gives the shortest path to q. Hence $H_{k+1}(q) = H_k(q) \cup h_f(e)$, and this partitions the face f into subfaces, one for each k-wall e, separated by $(k+1)$-walls. To finish the construction of the ($\leq k+1$)-SPM, we erase the k-walls on the boundary of f (recall that their neighborhoods have uniform $(k+1)$-homotopy sets), delete any old triple points whose neighborhoods have uniform $(k+1)$-homotopy sets, and erase any newly added $(k+1)$-walls incident to deleted old triple points on the boundary of f. (These "walls" are actually just windows generated by the triple points; they separate regions with equal $(k+1)$-homotopy sets.)

If a face f of the ($\leq k$)-SPM is bounded by B k-walls, it is initially partitioned into B subfaces. Every pair of subfaces incident to a common old triple point will be merged, so the final number of subfaces is $F' = B - W$, where W is the number of old triple points of the ($\leq k$)-SPM on the boundary of f. Since f is simply connected by Lemma 4.2, and every subface corresponding to a k-wall e must be adjacent to e, the dual graph of the subfaces inside f must be an outerplanar graph. The number of triple points V' added inside f (all of them new) corresponds to the number of (triangular) faces of this outerplanar graph, and hence $0 \leq V' \leq \max(F'-2, 0)$. By Euler's formula, the number of $(k+1)$-walls created inside f (duals to the edges of the outerplanar graph) is $B' = F' - 1 + V'$.

During the iterative construction of the ($\leq k$)-SPM, we count the features at each step. The description above considers what happens within a single face of the ($\leq k$)-SPM during the transformation to the ($\leq k+1$)-SPM. To account for what happens in all the faces simultaneously, we note that each i-wall is shared between two faces, and each triple point is shared between three faces. Let F_i and B_i be the number of faces and i-walls in the ($\leq i$)-SPM. To distinguish between new and old triple points, let V_i and W_i be the number of new and old triple points of the ($\leq i$)-SPM, respectively. (Note that $W_1 = 0$.) If we count just the features added inside faces of ($\leq i$)-SPM, using primed notation, we have

$$\begin{aligned} F'_{i+1} &= 2B_i - 3W_i \\ V'_{i+1} &\leq 2B_i - 3W_i - 2F_i \\ B'_{i+1} &= 2B_i - 3W_i - F_i + V'_{i+1} \\ W'_{i+1} &= 0 \end{aligned}$$

Now let us take into account the deletion of previous i-walls and triple points. All the i-walls and old triple points are deleted between one phase and the next. All new triple points turn into old ones. All subfaces incident to an old triple point merge into one. Thus we

obtain the following recurrence relations, whose solution is given by Lemma 4.3.

$$\begin{aligned} F_{i+1} &= F'_{i+1} - B_i + W_i = B_i - 2W_i \\ V_{i+1} &= V'_{i+1} \le 2B_i - 3W_i - 2F_i \\ B_{i+1} &= B'_{i+1} = 2B_i - 3W_i - F_i + V_{i+1} \\ W_{i+1} &= V_i \end{aligned}$$

$$\begin{aligned} F_1 &= 1 \\ V_1 &\le h - 1 \\ B_1 &= h + V_1 \\ W_1 &= 0 \end{aligned}$$

LEMMA 4.3. *The number of faces, walls, and triple points of the* ($\le k$)*-SPM is* $O(k^2 h)$.

We now return to the complexity of the k-SPM. The number of k-walls and $(k-1)$-walls can be bounded by Lemma 4.3. Each k-wall consists of one or more hyperbolic arcs. Note that the number of hyperbolic arcs for a single k-wall is exactly one more than the number of k-windows that end on the k-wall (and a k-window can end on only one k-wall). Hence it is sufficient to count the number of k-windows. Each k-window is an extension of the edge between a vertex v of P and its i-predecessor for $i \le k$. Thus there can be at most $O(kn)$ k-windows.

THEOREM 4.1. *The k-SPM of a polygonal domain with n vertices and h holes has complexity $O(k^2 h + kn)$.*

Lower Bound. The bound of Theorem 4.1 is in fact tight. Here we describe an example that has $\Omega(k^2 h)$ k-walls and $\Omega(kn)$ k-windows. The full details will be provided in the full version.

Consider the example in Fig. 7, which is constructed so that the shortest paths from s to the vertices p_1, p_2, and p_3 have the same length. Let q be the unique point equidistant from p_1, p_2, p_3. Furthermore, let π_{ij} ($i \in \{1,2,3\}$ and $1 \le j \le k$) be the j-path from s to p_i, and let l_{ij} be the length of π_{ij}. If the obstacle ω_i is small enough, then π_{ij} simply loops around ω_i zero or more times in a clockwise or counterclockwise direction. Hence, for any $\epsilon > 0$, we can ensure that $|l_{ik} - l_{i1}| \le \epsilon$ for $i \in \{1,2,3\}$ by making the obstacles ω_i small enough. Now define q_{abc} as the unique point such that $|q_{abc} - p_1| + l_{1a} = |q_{abc} - p_2| + l_{2b} = |q_{abc} - p_3| + l_{3c}$. This point must exist, since it is the vertex of an additively weighted Voronoi diagram of p_1, p_2, and p_3. If ϵ is chosen small enough, then q_{abc} must lie in the circle in Fig. 7 for $a, b, c \le k$.

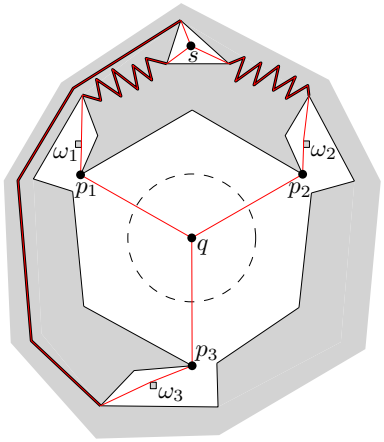


Figure 7: Lower bound gadget.

By construction there are three paths with equal length from s to q_{abc}, and there are exactly $a+b+c-3$ shorter paths from s to q_{abc}. This means that q_{abc} is a triple point of the $(a+b+c-2)$-SPM. Thus, the number of triple points of the k-SPM is exactly the number of triples (a,b,c) with $1 \le a, b, c \le k$ for which $a + b + c - 2 = k$. It is easy to see that there are $\Omega(k^2)$ triples that satisfy these conditions. Note that the gadget has $O(1)$ holes. By connecting $\Theta(h)$ copies of the basic gadget, we get a domain with h holes and $\Omega(k^2 h)$ k-SPM vertices. We can also replace p_3 in one copy by a convex chain of $n' = \Theta(n)$ vertices $v_1, \ldots, v'_n$, such that the line through v_i and v_{i+1} is very close to q for $1 \le i < n'$. This way each vertex v_i contributes k k-windows to the k-SPM. Full details will be provided in the full version.

THEOREM 4.2. *The k-SPM of a polygonal domain with n vertices and h holes can have $\Omega(k^2 h)$ k-walls and $\Omega(kn)$ k-windows.*

5 Computing the k-SPM

We now describe how to compute the k-SPM in $O((k^3 h + k^2 n) \log(kn))$ time. Inspired by the structure of the k-garage and Lemma 3.3, our algorithm iteratively computes the k-SPM for increasing values of k, starting from $k = 1$. Essentially we compute the SPM on the k-garage, one floor at a time. To compute the k-SPM at each iteration, we apply the "continuous Dijkstra" method, which Hershberger and Suri [11] used to compute the shortest path map among polygonal obstacles. We adopt most of the details of the Hershberger–Suri algorithm unchanged, but make a few modifications to support k-SPM computation.

The main idea of the continuous Dijkstra method is to simulate the progress of a wavefront that emerges from the source and expands through the free space with unit speed. If the wavefront reaches a point p at time t, then the shortest path distance between p and

the source is t. At any time, the wavefront consists of circular arc *wavelets*, each expanding from a weighted obstacle vertex called a *generator* (see Fig. 8a). A generator γ is represented as a pair (v, w), where v is an obstacle vertex and w is the shortest path distance from the source to v. For a generator $\gamma = (v, w)$ and a point p such that the segment $\overline{vp}$ is contained in free space, the (weighted) distance between γ and p, denoted $d(p, \gamma)$, is defined as $w + |\overline{vp}|$; it is the length of the shortest path from the source to p that passes through v.

Points in the wavelet corresponding to a generator γ at time t satisfy the equation $d(p, \gamma) = t$. We say that a point p is *claimed* by γ if γ is the generator whose wavelet first reaches p; this implies that the shortest path to p passes through v and has length $d(p, \gamma)$. The points where adjacent wavelets on the wavefront meet trace out the bisectors that form the walls and the windows of the shortest path map. Each bisector separates two cells of the shortest path map, each of which consists of points claimed by a particular generator. The bisector curve separating the regions claimed by two generators γ and γ' satisfies the equation $d(p, \gamma) = d(p, \gamma')$. Because $|vp| - |v'p| = w' - w$, the curve is a hyperbolic arc.

Using the continuous Dijkstra approach, the Hershberger–Suri algorithm computes shortest paths from a single source. It also works for shortest paths from multiple sources with delays. This is summarized in the following lemma, which was proved in [11].

LEMMA 5.1. ([11]) *Given a set of polygonal obstacles with n vertices and a set of $O(n)$ sources with delays, one can compute the corresponding shortest path map in $O(n \log n)$ time.*

To compute the k-SPM, we apply the continuous Dijkstra framework on each floor of the k-garage. Imagine that we start a wavefront expansion from the source. When a wavelet collides with another wavelet during propagation (and thus forms a 1-wall), the portion of the wavelet that is claimed by the other wavelet continues to expand on the 2-floor (see Fig. 8b). Since this portion of the wavelet has passed through a 1-wall, it represents a set of 2-paths, by Lemma 3.3. Any bisectors formed by adjacent wavelets on the 2-floor belong to the 2-SPM. Similarly to the 1-floor, when two wavelets collide on the 2-floor, they form a 2-wall and continue to expand on the 3-floor. We continue to push the colliding wavelets up to higher floors until they reach the k-floor, which will correspond to the k-SPM.

Notice that the wavefront expansion on a single floor is not affected by the expansion on other floors, with the exception of wavelet collisions on the previous floor. We now describe a method that exploits this fact

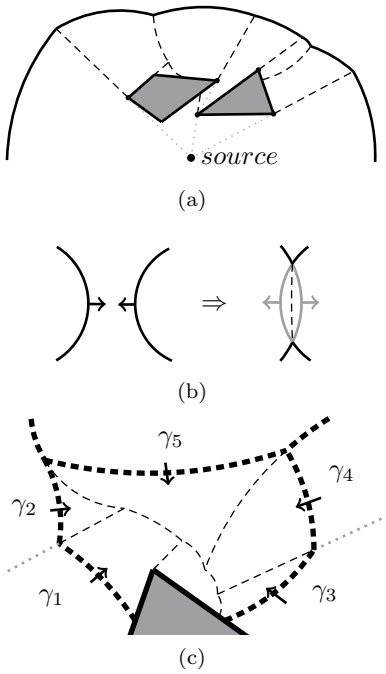


Figure 8: (a) An expanding wavefront. (b) Two colliding wavelets. After the collision, a wall is formed and both wavelets continue to grow on the next floor. (c) A shortest path map is computed by propagating outside generators into the region R.

to compute the k-SPM once the $(k-1)$-SPM has been computed. Thus we can construct the k-SPM by first running the Hershberger–Suri algorithm to compute the 1-SPM and then iteratively applying this step to compute higher floor SPMs.

We compute the k-SPM from the $(k-1)$-SPM as follows. The boundaries of the $(k-1)$-SPM are formed by $(k-1)$-windows, $(k-1)$-walls and $(k-2)$-walls. The $(k-1)$-windows and $(k-2)$-walls do not appear in the k-SPM, so we simply remove them from the map. The $(k-1)$-walls remain in the map and they subdivide the free space into simply connected regions (by Lemma 4.2). To complete the k-SPM, in each such region we compute a special shortest path map whose walls and windows form the k-windows and k-walls of the k-SPM.

The shortest path map computed in each region R is drawn with respect to multiple "restricted" sources with delays, which are determined as follows. Consider a $(k-1)$-wall W bounding R in the $(k-1)$-SPM and let $\gamma = (v, w)$ be the generator that claims the region outside R in the vicinity of W. (It is possible that both sides of W are contained in R. In this case, our description applies to the generators claiming both sides.) Note that W is formed by the collision of the wavelet of γ with another wavelet, and the wavelet of

γ is pushed up to the k-floor inside R. Conceptually, we want to continue expanding the wavelet of γ inside R. To do this, we introduce γ as a source at v with delay w and impose the additional restriction that all paths from γ to the interior of R pass through W.[1] In other words, we do not allow any paths from v that do not pass through W. We create sources in this manner for each $(k-1)$-wall bounding R and draw the shortest path map with respect to these sources (see Fig. 8c).

We can compute the shortest path map inside each region by running a single instance of the Hershberger–Suri algorithm for delayed sources. Our restrictions necessitate some modifications, described in the full version, but with these modifications the algorithm computes the shortest path map in each region bounded by $(k-1)$-walls. Since the paths used to compute the map in each region are k-paths by Lemma 3.3, the walls and windows of the map form the k-walls and k-windows of the k-SPM. This completes the construction of the k-SPM.

THEOREM 5.1. *Given a source point in a polygonal domain with n vertices and h holes, the corresponding k-SPM can be computed in $O((k^3h+k^2n)\log(kn))$ time. If the total complexity of all i-SPMs for $1 \le i \le k$ is M, then the running time is $O(M \log(kn))$.*

6 Visibility-based algorithms

The k-SPM provides an efficient data structure for querying k-paths from a fixed source s. If we are simply interested in the k-path between two fixed points s and t, then it may be inefficient to construct the k-SPM for large values of k. In this section we present a simple visibility-based algorithm to compute the k-path between s and t. For large k, this algorithm is faster than the k-SPM approach. Moreover, this algorithm is relatively easy to implement and may therefore be of more practical interest.

We first compute the visibility graph (VG) of P in $O(n \log n + m)$ time [7, 18], where $m = O(n^2)$ is the size of VG. We also include visibility edges to s and t. The graph contains every locally shortest path from s to t and hence also the k-path to t. However, we cannot simply compute the kth shortest path in VG, since different paths in the graph may be homotopic. We therefore modify VG so that locally shortest paths are in one-to-one correspondence with paths in the modified graph—this ensures that different paths in the graph belong to different homotopy classes by Lemma 2.1. (The same graph is defined in [8] and is called the *canonical graph*. Here we include its construction to argue the running time of computing this graph.) First, we make the graph directed by doubling each edge. Then we expand each vertex v as illustrated in Fig. 9: Draw the two lines supporting the two obstacle edges incident to v; the lines partition the relevant visibility edges at v into two sets A and B (the visibility edges between the lines opposite the obstacle are irrelevant, because they cannot be used by shortest paths). Radially sweep a line through v, initially aligned with one of the obstacle edges, until it is aligned with the other obstacle edge. For each visibility edge e encountered, create a node with an incoming edge if $e \in A$, and an outgoing edge if $e \in B$. Connect all created nodes with a directed path. Also make a copy of this construction with all edges reversed. The expansion of v is connected with other expansions in the obvious way, as dictated by the visibility graph. Finally, remove edges directed toward s and away from t. The constructed graph—which we call the *taut graph* $\vec{G}(P)$—has $O(m)$ vertices and $O(m)$ edges and can be built in $O(m)$ time. Note that, by construction, every path in $\vec{G}(P)$ must be locally shortest and every locally shortest path from s to t exists in $\vec{G}(P)$.

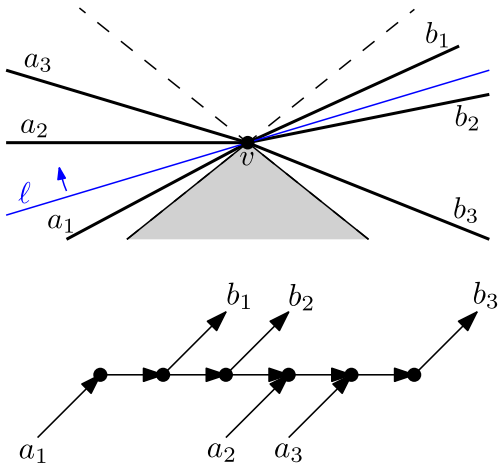


Figure 9: Vertex expansion for the taut graph.

We can now use the algorithm by Eppstein [6] to compute the kth shortest path from s to t in $\vec{G}(P)$, which corresponds to the k-path from s to t in P.

THEOREM 6.1. *The k-path between s and t in P can be computed in $O(m \log n + k)$ time, where m is the size of the visibility graph of P.*

Interestingly, this approach can be extended to compute the kth shortest *simple* path (*simple k-path*) between s and t in polynomial time. Here we define a *simple path* as a path that does not cross itself, although

[1] We also require that the subpath between v and W is a straight line.

repeated vertices and segments are allowed. To compute simple k-paths, we adapt Yen's algorithm [19] for computing simple k-paths in directed graphs (here "simple" means free of repeated nodes). The details are non-trivial and are provided in the full version. We obtain the following result.

THEOREM 6.2. *The simple k-path between s and t can be computed in $O(k^2m(m+kn)\log kn)$ time, where m is the number of edges of the visibility graph of P.*

7 Concluding remarks

We have introduced the k-SPM, a data structure that can efficiently answer k-path queries. We provided a tight bound for the complexity of the k-SPM, and presented an algorithm to compute the k-SPM efficiently. Our algorithm simultaneously computes all the i-SPMs for $i \leq k$. Whether there is a more direct algorithm to compute the k-SPM is an interesting open problem. We also provided a simple visibility-based algorithm to compute k-paths, which may be of practical interest, and is more efficient for large values of k. This latter approach can be extended to compute simple k-paths. Unfortunately, we do not know how to extend the k-SPM to simple k-paths. It seems that simple k-paths lack the useful property that a subpath of a simple k-path is a simple i-path for $i \leq k$. This makes finding a more efficient algorithm to compute simple k-paths a challenging open problem.

Acknowledgments. We thank Yevgeny Schreiber, Niko Kiirala, Jukka Suomela and Joe Mitchell for discussions and anonymous referees for useful comments.

References

[1] C. Bohler, P. Cheilaris, R. Klein, C.-H. Liu, E. Papadopoulou, and M. Zavershynskyi. On the complexity of higher order abstract Voronoi diagrams. In *ICALP (1)*, volume 7965 of *Lecture Notes in Computer Science*, pages 208–219. Springer, 2013.

[2] D. Z. Chen, J. Hershberger, and H. Wang. Computing shortest paths amid convex pseudodisks. *SIAM J. Comput.*, 42(3):1158–1184, 2013.

[3] D. Z. Chen and H. Wang. L_1 shortest path queries among polygonal obstacles in the plane. In *STACS*, volume 20 of *LIPIcs*, pages 293–304. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2013.

[4] D. Eppstein. Finding the k smallest spanning trees. *BIT*, 32(2):237–248, 1992.

[5] D. Eppstein. Tree-weighted neighbors and geometric k smallest spanning trees. *Int. J. Comput. Geometry Appl.*, 4(2):229–238, 1994.

[6] D. Eppstein. Finding the k shortest paths. *SIAM J. Comput.*, 28(2):652–673, 1999.

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

[8] D. Grigoriev and A. Slissenko. Polytime algorithm for the shortest path in a homotopy class amidst semi-algebraic obstacles in the plane. In *Proceedings of the 1998 International Symposium on Symbolic and Algebraic Computation*, ISSAC '98, pages 17–24. ACM, 1998.

[9] J. Hershberger, M. Maxel, and S. Suri. Finding the k shortest simple paths: A new algorithm and its implementation. *ACM Trans. Algorithms*, 3(4):45, 2007.

[10] J. Hershberger, V. Polishchuk, B. Speckmann, and T. Talvitie. Geometric kth shortest paths: The applet. In *Proceedings of the Thirtieth Annual Symposium on Computational Geometry*, SOCG'14, pages 96:96–96:97, New York, NY, USA, 2014. ACM. `http://www.computational-geometry.org/SoCG-videos/socg14video/ksp/index.html`.

[11] J. Hershberger and S. Suri. An optimal algorithm for Euclidean shortest paths in the plane. *SIAM J. Comput.*, 28(6):2215–2256, 1999.

[12] E. L. Lawler. A procedure for computing the K best solutions to discrete optimization problems and its application to the shortest path problem. *Management Science*, 18:401–405, 1972.

[13] D.-T. Lee. On k-nearest neighbor Voronoi diagrams in the plane. *IEEE Trans. Computers*, 31(6):478–487, 1982.

[14] C.-H. Liu and D. T. Lee. Higher-order geodesic Voronoi diagrams in a polygonal domain with holes. In *SODA*, pages 1633–1645. SIAM, 2013.

[15] C.-H. Liu, E. Papadopoulou, and D. T. Lee. The k-nearest-neighbor Voronoi diagram revisited. *Algorithmica*, 2014. To appear.

[16] J. S. B. Mitchell. Geometric shortest paths and network optimization. In J.-R. Sack and J. Urrutia, editors, *Handbook of Computational Geometry*, pages 633–701. Elsevier Science B.V. North-Holland, Amsterdam, 2000.

[17] E. Papadopoulou and M. Zavershynskyi. On higher order Voronoi diagrams of line segments. In *ISAAC*, volume 7676 of *Lecture Notes in Computer Science*, pages 177–186. Springer, 2012.

[18] M. Pocchiola and G. Vegter. Topologically sweeping visibility complexes via pseudotriangulations. *Discrete & Computational Geometry*, 16(4):419–453, 1996.

[19] J. Y. Yen. Finding the K shortest loopless paths in a network. *Management Science*, 17:712–716, 1971.