

# Middlesex University Research Repository

An open access repository of

Middlesex University research

<http://eprints.mdx.ac.uk>

Alt, Helmut, Cabello, Sergio, Giannopoulos, Panos and Knauer, Christian (2017) Minimum cell connection in line segment arrangements. *International Journal of Computational Geometry and Applications*, 27 (3). pp. 159-176. ISSN 1793-6357

Final accepted version (with author's formatting)

This version is available at: <http://eprints.mdx.ac.uk/16045/>

## Copyright:

Middlesex University Research Repository makes the University's research available electronically.

Copyright and moral rights to this work are retained by the author and/or other copyright owners unless otherwise stated. The work is supplied on the understanding that any use for commercial gain is strictly forbidden. A copy may be downloaded for personal, non-commercial, research or study without prior permission and without charge.

Works, including theses and research projects, may not be reproduced in any format or medium, or extensive quotations taken from them, or their content changed in any way, without first obtaining permission in writing from the copyright holder(s). They may not be sold or exploited commercially in any format or medium without the prior written permission of the copyright holder(s).

Full bibliographic details must be given when referring to, or quoting from full items including the author's name, the title of the work, publication details where relevant (place, publisher, date), pagination, and for theses or dissertations the awarding institution, the degree type awarded, and the date of the award.

If you believe that any material held in the repository infringes copyright law, please contact the Repository Team at Middlesex University via the following email address:

[eprints@mdx.ac.uk](mailto:eprints@mdx.ac.uk)

The item will be removed from the repository while any claim is being investigated.

See also repository copyright: re-use policy: <http://eprints.mdx.ac.uk/policies.html#copy>

# Minimum cell connection in line segment arrangements\*

Helmut Alt<sup>†</sup>   Sergio Cabello<sup>‡§</sup>   Panos Giannopoulos<sup>¶</sup>   Christian Knauer<sup>\*\*</sup>

## Abstract

We study the complexity of the following cell connection problems in segment arrangements. Given a set of straight-line segments in the plane and two points  $a$  and  $b$  in different cells of the induced arrangement:

- (i) compute the minimum number of segments one needs to remove so that there is a path connecting  $a$  to  $b$  that does not intersect any of the remaining segments;
- (ii) compute the minimum number of segments one needs to remove so that the arrangement induced by the remaining segments has a single cell.

We show that problems (i) and (ii) are NP-hard and discuss some special, tractable cases. Most notably, we provide a near-linear-time algorithm for a variant of problem (i) where the path connecting  $a$  to  $b$  must stay inside a given polygon  $P$  with a constant number of holes, the segments are contained in  $P$ , and the endpoints of the segments are on the boundary of  $P$ . The approach for this latter result uses homotopy of paths to group the segments into clusters with the property that either all segments in a cluster or none participate in an optimal solution.

## 1 Introduction

In this paper we study the complexity of some natural optimization problems in segment arrangements. Let  $S$  be a set of straight-line segments in  $\mathbb{R}^2$ ,  $\mathcal{A}(S)$  be the arrangement induced by  $S$ , and  $a, b$  be two points not incident to any segment of  $S$  and in different cells of  $\mathcal{A}(S)$ .

In the 2-CELLS-CONNECTION problem we want to compute a set of segments  $S' \subseteq S$  of minimum cardinality with the property that  $a$  and  $b$  belong to the same cell of  $\mathcal{A}(S \setminus S')$ . In other words, we want to compute an  $a$ - $b$  path that *crosses* the minimum number of segments

---

\*A preliminary version of the results contained here were presented at EuroCG 2011 [ACGK11]. Research partially conducted at the 9th McGill - INRIA Barbados Workshop on Computational Geometry, 2010.

<sup>†</sup>Institut für Informatik, Freie Universität Berlin, Takustraße 9, D-14195 Berlin, Germany, [alt@mi.fu-berlin.de](mailto:alt@mi.fu-berlin.de).

<sup>‡</sup>Department of Mathematics, IMFM and FMF, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia, [sergio.cabello@fmf.uni-lj.si](mailto:sergio.cabello@fmf.uni-lj.si).

<sup>§</sup>Research was partially supported by the Slovenian Research Agency, program P1-0297, project J1-4106, and within the EUROCORES Programme EUROGIGA (project GReGAS) of the European Science Foundation.

<sup>¶</sup>School of Science and Technology, Middlesex University, The Burroughs, Hendon, London NW4 4BT, UK, [p.giannopoulos@mdx.ac.uk](mailto:p.giannopoulos@mdx.ac.uk).

<sup>||</sup>Research was partially supported by the German Science Foundation (DFG) under grant Kn 591/3-1.

<sup>\*\*</sup>Institut für Informatik, Universität Bayreuth Universitätsstraße 30 D-95447 Bayreuth Germany, [christian.knauer@uni-bayreuth.de](mailto:christian.knauer@uni-bayreuth.de).



between two specified regions in a polygonal domain has been shown to be polynomial-time solvable by Kloder and Hutchinson [KH07].

The problems we consider can of course be considered for other geometric objects, most notably unit disks. To this end, closely related work was done by Bereg and Kirkpatrick [BK09], who studied the counterpart of 2-CELLS-CONNECTION in arrangements of unit disks and gave a 3-approximation algorithm. There exist polynomial-time algorithms for restricted belt-shaped and simple polygonal domains [KLA07]. The 2-CELLS-CONNECTION for so-called fat regions has been studied at length by Korman et al. [KLSS13], where several positive and negative results are provided. In any case, the complexity of 2-CELLS-CONNECTION for unit (or arbitrary) disks is still unknown.

## 2 Connecting two cells

We show that 2-CELLS-CONNECTION is NP-hard and APX-hard by a reduction from EXACT-MAX-2-SAT, a well studied NP-complete and APX-complete problem (c.f. [Hås01]): Given a propositional CNF formula  $\Phi$  with  $m$  clauses on  $n$  variables and exactly two variables per clause, decide whether there exists a truth assignment that satisfies at least  $k$  clauses, for a given  $k \in \mathbb{N}$ ,  $k \leq m$ . Let  $x_1, \dots, x_n$  be the variables of  $\Phi$ ,  $\ell_i$  be the number of appearances of variable  $x_i$  in  $\Phi$ , and  $\ell = \sum_i \ell_i$ ; since each clause contains exactly 2 variables,  $\ell = 2m$ . The maximum number of satisfiable clauses is denoted by  $\text{opt}(\Phi)$ . Using  $\Phi$  we construct an instance consisting of a set of segments  $S = S(\Phi)$  and two points  $a = a(\Phi)$  and  $b = b(\Phi)$  as follows.

Abusing the terminology slightly, the term *segment* will refer to a set of identical single segments stacked on top of each other. The cardinality of the set is the *weight* of the segment. Either all or none of the single segments in the set can be crossed by a path. There are two different types of segments,  $\tau_1$ , and  $\tau_\infty$ , according to their weight. Segments of type  $\tau_1$  have weight 1 (light or single segments), while segments of type  $\tau_\infty$  have weight  $20m$  (heavy segments). The weight of heavy segments is chosen so that they are never crossed by an optimal  $a$ - $b$  path.

We first provide an informal, high-level description of the construction that uses *curved* segments. Later on, each curved segment will be replaced by a collection of straight-line segments in an appropriate manner. See Fig. 2. We have a rectangle  $R_\infty$  made of heavy segments, with point  $a$  at a lower corner and  $b$  at an upper corner. For each variable  $x_i$ , we add a small vertical segment of type  $\tau_\infty$  in the lower half of  $R_\infty$ . From the segment we place  $\ell_i$  horizontal light segments, denoted by  $R_i$ , going to the right and  $\ell_i$  horizontal light segments, denoted by  $L_i$ , going to the left until they reach the outside of  $R_\infty$ . Roughly speaking, (things are slightly more complicated) an optimal  $a$ - $b$  path will have to choose for each  $x_i$  whether it crosses all segments in  $L_i$ , encoding the assignment  $x_i = \text{T}$ , or all segments in  $R_i$ , encoding the assignment  $x_i = \text{F}$ . Consider a clause like  $x_2 \vee x_n$ , where both literals are positive. We prolong one of the segments of  $L_2$  and one of the segments of  $L_n$  with a curved segment so that they cross again inside  $R_\infty$  (upper half) in such a way that an  $a$ - $b$  path inside  $R_\infty$  must cross one of the prolongations, and one is enough; see Fig. 2, where one of the prolongations passes below  $R_\infty$ . A clause like  $\bar{x}_n \vee x_2$  is represented using prolongations of one segment from  $L_2$  and one segment of  $R_n$ . The other types of clauses are symmetric. For each clause we always prolong different segments; since  $L_i$  and  $R_i$  have  $\ell_i$  segments, there is always some segment that can be prolonged. It will then be possible to argue that the optimal  $a$ - $b$  path

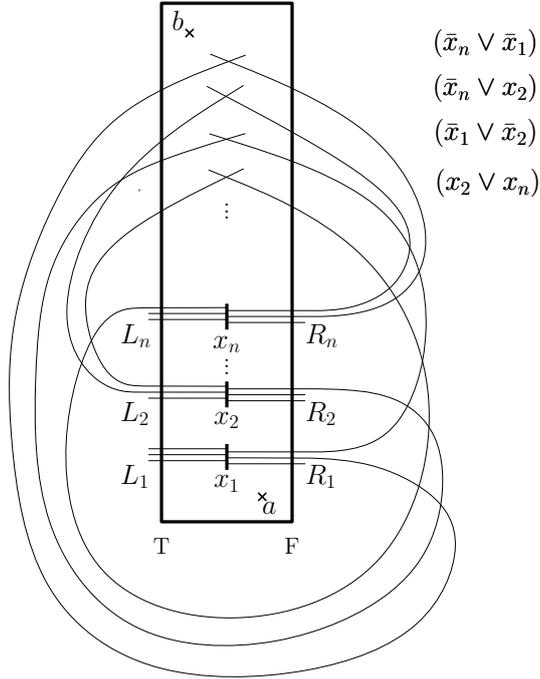


Figure 2: Idea of the construction with curved segments.

has cost  $\ell + (m - \text{opt}(\Phi))$ . We do not provide a careful argument of this here since we will need it later for a most complicated scenario. This finishes the informal description of the idea.

We now describe in detail the construction with straight-line segments. First, we construct a polygon, called the *tunnel*, with heavy boundary segments of type  $\tau_\infty$ ; see Fig. 3(a). The tunnel has a ‘zig-zag’ shape and can be seen as having 8 corridors,  $C_1, \dots, C_8$ . It starts with  $C_1$ , the *main* corridor (at the center of the figure), which contains point  $a$ , then it turns left to  $C_2$ , then right, etc., gradually turning around to  $C_7$  and then to the *end* corridor  $C_8$  (at the top). The latter contains point  $b$ . To facilitate the discussion, we place a point  $b'$  in the tunnel where the transition from  $C_7$  to the end corridor occurs. The tunnel has a total weight of  $21 \cdot 20m = \mathcal{O}(m)$ . The rest of the construction will force any  $a$ - $b$  path of some particular cost (to be given shortly) to stay always in the interior of the tunnel.

Each variable  $x_i$  of  $\Phi$  is represented by a collection of 16 pieces, which form a chain-like structure. Each *piece* is a group of  $\ell_i$  nearly-parallel single segments whose ends are either outside the tunnel or lie on ‘short’ heavy segments of type  $\tau_\infty$  in the interior of the tunnel, referred to as *obstacles*. For each variable, there is one obstacle in each of the corridors  $C_1, C_2, C_7$  and there are two obstacles in each of the corridors  $C_3, C_4, C_5$ , and  $C_6$ . See Fig. 3(a), where we represent each piece by a light gray trapezoid and each obstacle by a bold, short segment. Pieces always contain a part outside the tunnel. The exact description of the structure is cumbersome; we refer the reader to the figures. The obstacle in  $C_2$  contains the extremes of four pieces: two pieces, called  $P_i$ , go to the obstacle in the main corridor, one goes to an obstacle in  $C_3$ , and the fourth piece, which we call  $N_i^l$  goes outside the tunnel. Symmetrically, the obstacle in  $C_7$  contains the extremes of four pieces: two pieces, called  $N_i$ , go to the main corridor, one goes to the corridor  $C_6$ , and one, which we call  $P_i^r$  goes outside

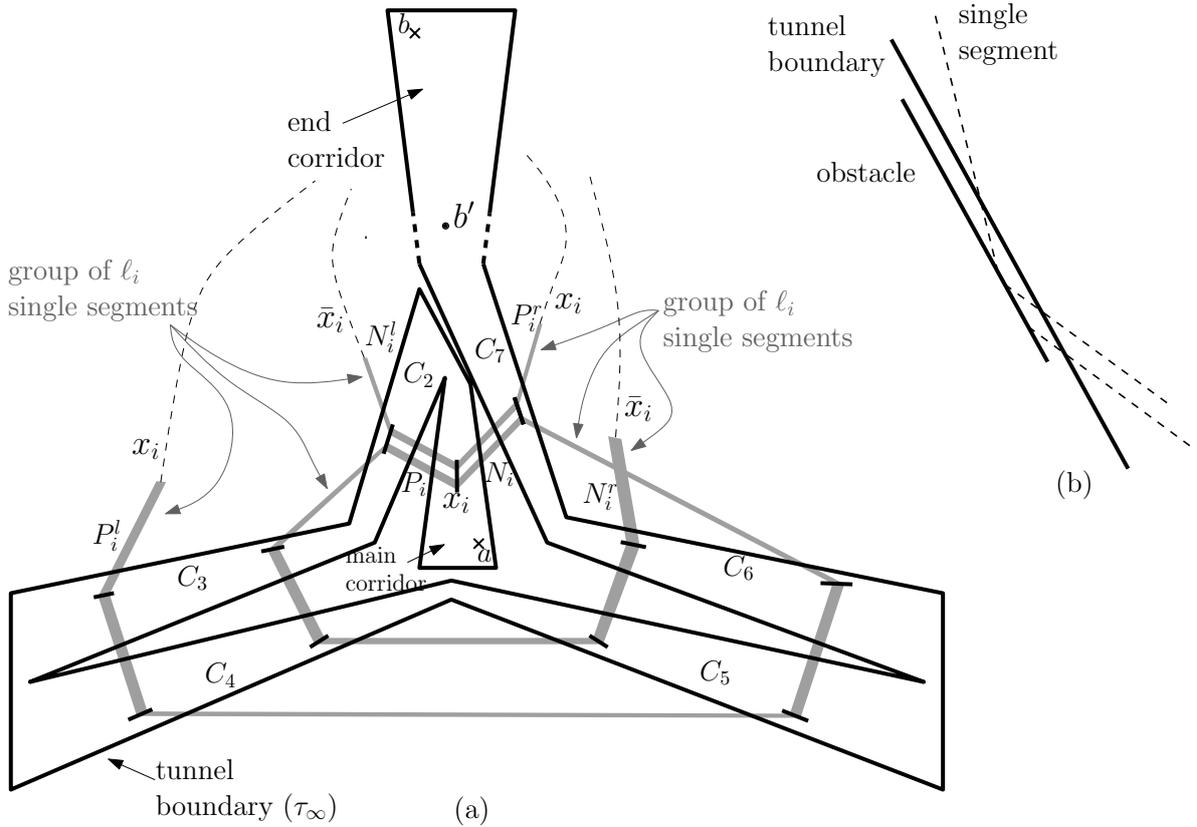


Figure 3: (a) Tunnel and variable chain. Each gray trapezoid represents a piece with  $\ell_i$  parallel segments. (b) Part of a chain piece close to the tunnel.

the tunnel. We add pieces connecting the obstacles in  $C_3$  and  $C_4$ , the obstacles in  $C_4$  and  $C_5$ , and the obstacles in  $C_5$  and  $C_6$ . From the obstacle in  $C_3$  that currently has one piece we add another piece, which we call  $P_i^l$  and whose other extreme is outside the tunnel. From the obstacle in  $C_6$  that currently has one piece we add another piece, which we call  $N_i^r$ , whose other extreme is outside the tunnel.

The obstacles and the pieces of all variables should satisfy some conditions: obstacles should be disjoint, pieces can touch only the obstacles at their extremes, and pieces may cross only outside the tunnel or inside the end corridor of the tunnel. See Fig. 4. Some of the single segments of  $P_i^r$ ,  $P_i^l$ ,  $N_i^r$ ,  $N_i^l$  will be prolonged and rotated slightly to encode the clauses. For this, we will need that the line supporting a segment from  $P_i^r \cup N_i^r$  intersects inside the end corridor the line supporting a segment from  $P_j^l \cup N_j^l$ . This can be achieved by stretching the end corridor sufficiently and placing the obstacles of  $C_2$  and  $C_7$  close to the tunnel boundary; see Fig. 3(b).

For each clause of  $\Phi$  we prolong two segments of  $P_i^r \cup P_i^l \cup N_i^r \cup N_i^l$  as follows; see Fig. 4 for an example of the overall construction, where prolongations are shown by dashed lines. Each segment corresponds to some literal  $x_i$  or  $\bar{x}_i$  in the clause: in the first case the segment comes from either  $P_i^r$  or  $P_i^l$ , while in the second one it comes from either  $N_i^r$  or  $N_i^l$ . For the construction, these choices for each clause can be made arbitrarily, provided that one segment intersects the tunnel from the left side and the other one from the right. These segments are

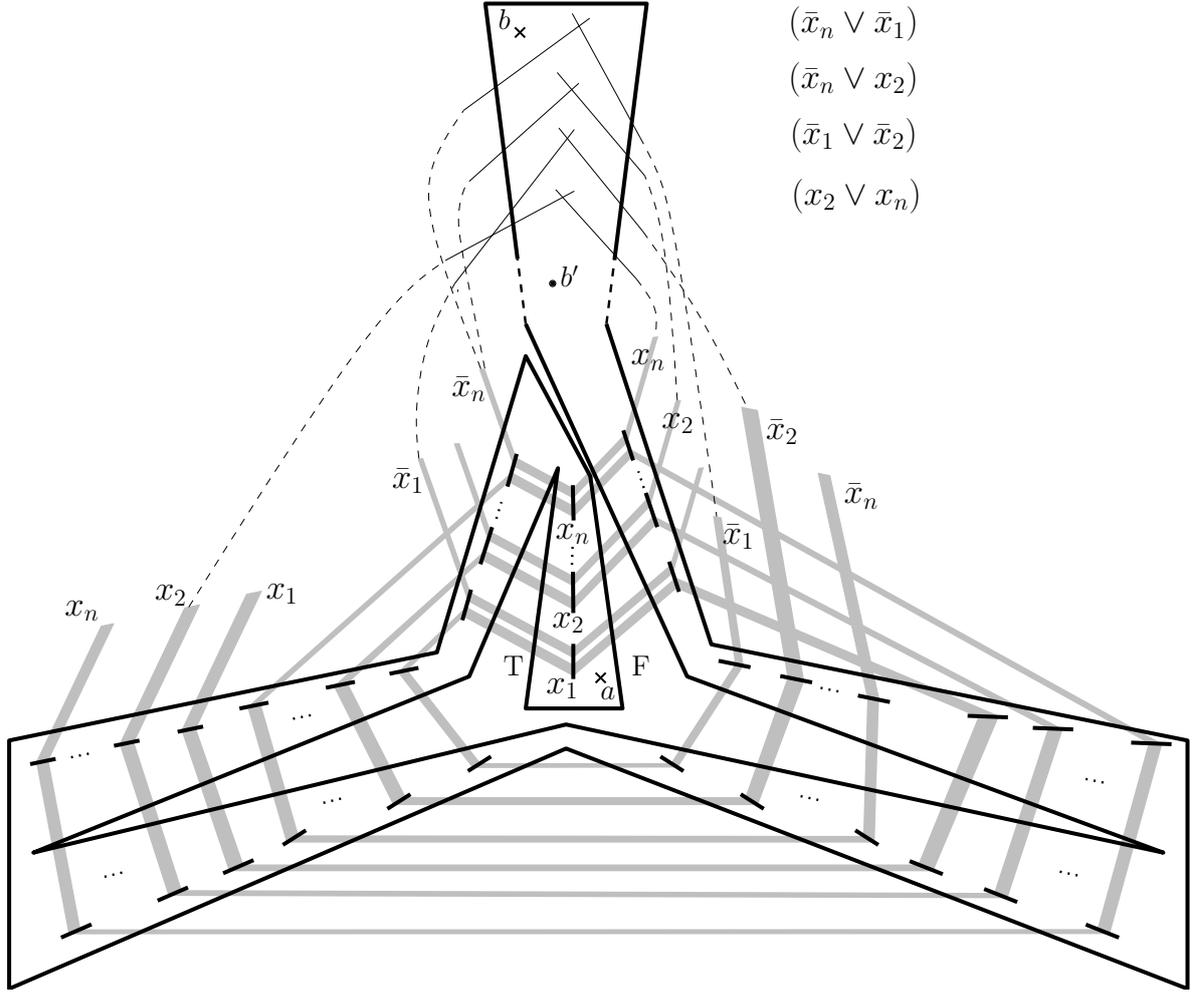


Figure 4: Example of overall construction.

prolonged until their intersection point inside the end corridor. For each clause, two different segments are prolonged. Since the pieces corresponding to variable  $x_i$  have  $\ell_i$  segments, there is always some segment available. Segments corresponding to different clauses may intersect only outside the tunnel; this is ensured by rotating the segments slightly around the endpoint lying in the obstacle. In this way, the end corridor is obstructed by  $m$  pairs of intersecting segments such that any path from the intermediate point  $b'$  to point  $b$  staying inside the tunnel must intersect at least one segment from each pair.

The following lemma establishes the correctness of the reduction.

**Lemma 1.** *There is an  $a$ - $b$  path of cost at most  $8\ell + k$ , where  $1 \leq k \leq m$ , if and only if there is a truth assignment satisfying at least  $(m - k)$  of the clauses.*

*Proof.* We denote by  $S_i$  the set of segments in the pieces corresponding to the variable  $x_i$ . Let  $S_i^T \subset S_i$  be the set constructed in the following way: starting from  $P_i^l$  and moving along the chain structure corresponding to  $x_i$ , we put in  $S_i^T$  the segments in  $P_i^l$ , the segments in the piece connecting  $C_4$  to  $C_5$ , the segments in the piece connecting  $C_6$  to  $C_7$ , the segments in  $P_i^r$ , the segments in the both pieces of  $P_i$ , the segments in the piece connecting  $C_3$  to  $C_4$ ,

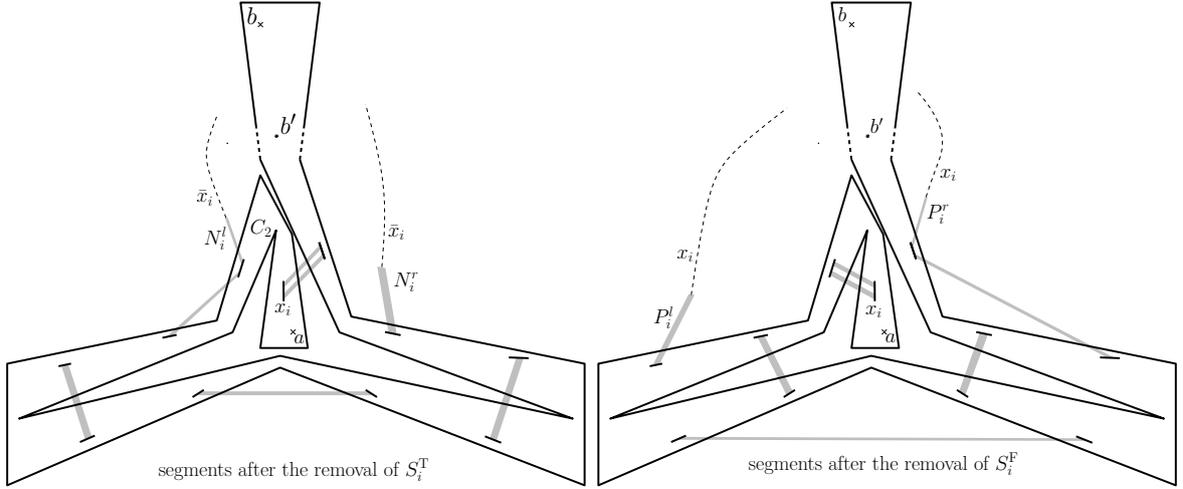


Figure 5: Removal of  $S_i^T$  (left) and  $S_i^F$  (right).

and the segments in the piece connecting  $C_5$  to  $C_6$ ; see Fig. 5(right). Note that  $S_i^T$  contains  $P_i$ ,  $P_i^l$  and  $P_i^r$ . We denote by  $S_i^F$  the segments  $S_i \setminus S_i^T$ . Note that  $S_i^F$  contains  $N_i$ ,  $N_i^l$  and  $N_i^r$ . Each of the sets  $S_i^T$  and  $S_i^F$  contains  $8\ell_i$  segments. Inside the tunnel there is an  $a$ - $b'$  path disjoint from  $S_i^T$  and there is another  $a$ - $b'$  path disjoint from  $S_i^F$ . We also denote by  $T_j$  the two segments used for clause  $C_j$  of  $\Phi$ .

Consider a truth assignment  $\{x_i = b_i\}$ , where each  $b_i \in \{T, F\}$ , satisfying at least  $(m - k)$  clauses. We construct a subset of segments  $S'$  where we include the set  $S_i^{b_i}$ , for each variable  $x_i$ , and a segment of  $T_j$ , for each clause  $C_j$  that is not satisfied by the truth assignment. Since  $|S_i^{b_i}| = 8\ell_i$ , the set  $S'$  contains at most  $8\ell + k$  segments. The removal of  $S'$  leaves the points  $a$  and  $b'$  in the same cell of the arrangement. Equivalently, there is an  $a$ - $b'$  path inside the tunnel that crosses only segments from  $S'$ . If a clause  $C_j$  of  $\Phi$  is satisfied by the truth assignment, then at least one of the segments in  $T_j$  is included in  $S_i^{b_i} \subset S'$ . If a clause  $C_j$  is not satisfied, then one of the segments  $T_j$  is included in  $S'$  by construction. Thus, for each clause  $C_j$  we have  $T_j \cap S' \neq \emptyset$ . It follows that  $b'$  and  $b$  are in the same cell after the removal of  $S'$ .

Conversely, note first that any  $a$ - $b$  path with cost at most  $8\ell + k \leq 16m + m = 17m$  cannot intersect the tunnel boundary or an obstacle because segments of type  $\tau_\infty$  have weight  $20m$ . Let  $S'$  be the set of segments crossed by the path. If  $P_i \subset S'$ , then we define  $b_i = T$ ; otherwise, we define  $b_i = F$ . Note that when  $P_i \not\subset S'$ , then  $N_i \subset S'$  because the  $a$ - $b$  path is inside the tunnel. (However it may be  $N_i \cup P_i \subset S'$ , so the assignment of  $b_i$  is not symmetric.) We next argue that the truth assignment  $\{x_i = b_i\}$  satisfies at least  $(m - k)$  clauses.

Consider the case when  $P_i \subset S'$ . Inspection shows that

$$|S' \cap S_i| \geq 8\ell_i + |S' \cap (N_i^l \cup N_i^r)|.$$

Indeed, after the removal of  $P_i \cup N_i^l \cup N_i^r$  any path from  $a$  to  $b'$  must still cross at least 6 pieces. Similarly, inspection shows that when  $N_i \subset S'$  we have

$$|S' \cap S_i| \geq 8\ell_i + |S' \cap (P_i^l \cup P_i^r)|.$$

Let  $A_i = N_i^l \cup N_i^r$  if  $b_i = T$  and  $A_i = P_i^l \cup P_i^r$  if  $b_i = F$ . The previous cases can be summarized

as

$$|S' \cap S_i| \geq 8\ell_i + |S' \cap A_i|.$$

We further define

$$Y = \bigcup_i (S' \cap A_i).$$

For each clause  $C_j$  we have  $S' \cap T_j \neq \emptyset$  by construction, as otherwise  $a$  and  $b$  cannot be in the same cell of  $S \setminus S'$ . If  $C_j$  is not satisfied by the truth assignment  $\{x_i = b_i\}$ , then it must be  $(S' \cap T_j) \subset S' \cap A_k$  for some variable  $x_k$  in  $C_j$ . This means that  $T_j \cap Y \neq \emptyset$ . Since the sets  $T_j$  are disjoint by construction, the number of unsatisfied clauses is bounded by  $|Y|$ . Using that

$$8\ell + k = |S'| = \sum_{i=1}^n |S' \cap S_i| \geq \sum_{i=1}^n (8\ell_i + |S' \cap A_i|) = 8\ell + \sum_{i=1}^n |S' \cap A_i|,$$

we obtain

$$\sum_{i=1}^n |S' \cap A_i| \leq k.$$

Therefore, the total number of clauses with value F is bounded by

$$|Y| = \sum_i |S' \cap A_i| \leq k. \quad \square$$

The construction can be easily modified by replacing every heavy segment with a set of  $20m$  distinct parallel single segments such that every single segment in  $S$  that originally intersected the heavy segment now intersects all the segments in the new set and such that no three segments have a point in common. We have the following:

**Theorem 2.** *2-CELLS-CONNECTION is NP-hard and APX-hard even when no three segments intersect at a point.*

*Proof.* NP-hardness follows from Lemma 1 and the fact that the reduction produces  $\mathcal{O}(nm)$  segments, whose coordinates can be bounded by a polynomial in  $(n + m)$ . APX-hardness follows from the fact that the reduction is approximation-preserving, as we now show.

First, since there is always an assignment that satisfies at least  $3m/4$  clauses, we have that  $m \leq (4/3)\text{opt}(\Phi)$ . Recall that an optimal  $a$ - $b$  path costs  $8\ell + (m - \text{opt}(\Phi))$ , where  $\ell = 2m$ . A polynomial-time  $c$ -approximation algorithm ( $c > 1$ ) for the problem would give a path that costs at most

$$\begin{aligned} c(8\ell + (m - \text{opt}(\Phi))) &= c(17m - \text{opt}(\Phi)) \\ &= 17m - c\text{opt}(\Phi) + 17(c - 1)m \\ &\leq 17m - c\text{opt}(\Phi) + 17(c - 1)(4/3)\text{opt}(\Phi) \\ &= 17m - \text{opt}(\Phi)(68/3 - (65/3)c) \\ &= 16m + \left[ m - \text{opt}(\Phi)(68/3 - 65c/3) \right] \end{aligned}$$

and, by Lemma 1, a truth assignment that satisfies at least  $\text{opt}(\Phi)(68/3 - 65c/3)$  clauses. However, EXACT-MAX-2-SAT cannot be approximated above  $21/22$  [Hås01], which implies that  $c$  must be larger than  $(68/65 - 63/(22 \cdot 65)) \approx 1.002097\dots$  (A slightly better inapproximability result can be obtained using the better bounds that rely on the unique games conjecture [KKMO07].)  $\square$

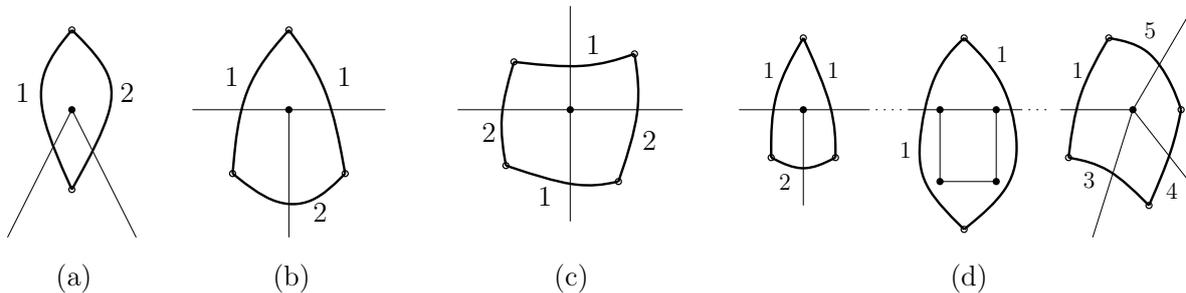


Figure 6: Examples of intersections in  $\mathcal{A}(S)$  and colored edges in  $G$ .

We can reduce 2-CELLS-CONNECTION to the minimum color path problem (MCP): Given a graph  $G$  with colored (or labeled) edges and two of its vertices, find a path between the vertices that uses the minimum possible number of colors. We color the edges of the dual graph  $G$  of  $\mathcal{A}(S)$  as follows: two edges of  $G$  get the same color if and only if their corresponding edges in  $\mathcal{A}(S)$  lie on the same segment of  $S$ . Then, finding an  $a$ - $b$  path of cost  $k$  in  $\mathcal{A}(S)$  amounts to finding a  $k$ -color path in  $G$  between the two cells which  $a, b$  lie in.

However, MCP is NP-hard [BLWZ05] and W[1]-hard [FGI10] (with respect to the solution size) even for planar graphs, it has a polynomial-time  $\mathcal{O}(\sqrt{n})$ -approximation algorithm and is non-approximable within any polylogarithmic factor [HMS07].

### 3 Tractable cases for connecting two cells

We now describe two special cases where 2-CELLS-CONNECTION is tractable. First, we consider the case where the input segments have few crossings, in a sense that is specified below. Then, we return to the special case where we have a polygon and provide an algorithm that takes near-linear time when the number of holes in the polygon is constant.

#### 3.1 Segments crossings.

We say that two segments cross if and only if they intersect at a point that is interior to both segments (a segment crossing). Without loss of generality, we assume that (i) every segment in  $S$  intersects at least two other segments and (ii) both endpoints of a segment are intersection points; otherwise the piece between the endpoint and the last crossing point is removed.

Consider the colored planar dual graph  $G$  of  $\mathcal{A}(S)$  as defined after Theorem 2. A face of  $G$  (except the outer one) corresponds to a point of intersection of some  $r \geq 2$  segments and has  $r$  colors and, depending on the type of intersection, from  $r$  to  $2r$  edges. For example, for  $r = 2$  we can get two multiple edges, a triangle, or a quadrilateral, with two distinct colors. See Fig. 6(a)-(c), where the colors are given as labels.

When any three segments may intersect only at a common endpoint and no two segments cross,  $G$  can only have multiple edges (possible all with the same color), bi-chromatic triangles, and arbitrary large faces where all edges have different colors; See Fig. 6(d) for an example. In this case, since two segments can intersect only at one point, each color induces a connected subgraph of  $G$ , in fact a tree (all but one multiple edges with the same color can be deleted, therefore there can be no monochromatic cycle in  $G$ ). Within such a tree any two vertices

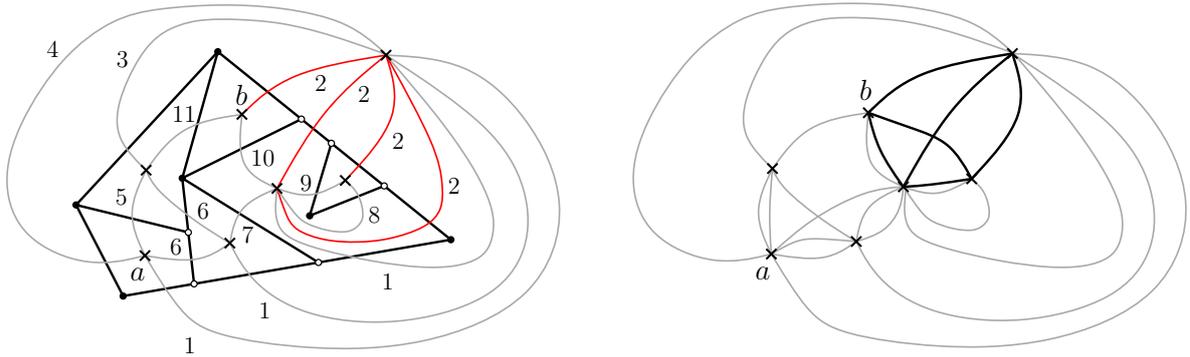


Figure 7: Example of the special case where any three segments may intersect only at a common endpoint and no two segments cross: the colors are given as labels (left) and the graph resulting from the completion of the trees into cliques is shown (right).

are connected by a (monochromatic) path, which means that in  $\mathcal{A}(S)$  the corresponding two cells can be connected by removing only one segment. Therefore, we can complete each tree into a clique by adding all edges missing from the tree and discard the colors; see Fig 7 for an example. Then, 2-CELLS-CONNECTION reduces to a simple shortest path computation between the cells (vertices) containing  $a$  and  $b$  in the resulting uncolored graph. By contrast, note that ALL-CELLS-CONNECTION is still NP-hard for this special case; see Section 4.

Generalizing this, if we allow  $k$  segment crossings, we can easily reduce the problem to  $2^{\mathcal{O}(k)}$  shortest path problems as follows. Let  $C \subseteq S$  be the set of the (at most  $2k$ ) segments participating in these crossings. For a fixed subset  $C'$  of  $C$ , we first contract every edge of  $G$  corresponding to a segment in  $C'$ , effectively putting all segments of  $C'$  into the solution. Then, we delete every edge corresponding to a segment in  $C \setminus C'$  that still participates in a crossing, i.e., we exclude all crossing segments of  $C \setminus C'$  from the solution. In the resulting (possibly disconnected) graph  $G'$ , each of the remaining colors induces again a monochromatic subtree, thus we can compute a shortest path as before and add  $C'$  to the solution. Finally, we return a minimum size solution set over all  $2^{\mathcal{O}(k)}$  possible subsets  $C'$ . Thus, we have just proved the following:

**Theorem 3.** *2-CELLS-CONNECTION is fixed-parameter tractable with respect to the number of segment crossings if any three segments may intersect only at a common endpoint.*

### 3.2 Polygon with holes.

Let  $P$  be a polygon with  $h$  holes and  $S$  be a set of segments lying inside  $P$  with their endpoints on its boundary; see Fig. 1. We use  $n$  as a bound for the number of vertices of  $P$  and segments in  $S$ . We consider the restricted 2-CELLS-CONNECTION problem where the  $a$ - $b$  path may not cross the boundary of  $P$ . This version is also NP-hard by a simple reduction from the general one: place a large polygon enclosing all the segments and add a hole at the endpoint of each segment. We assume for simplicity that  $a$  and  $b$  are in the interior of  $P$ .

A boundary component of  $P$  may be the exterior boundary or the boundary of a hole. For each boundary component  $\beta$  of  $P$ , let  $C_\beta$  be the connected component of  $\mathbb{R}^2 \setminus P$  that has  $\beta$  as boundary, and let  $z_\beta$  be an arbitrary, fixed point in the interior of  $C_\beta$ . If  $\beta$  is the exterior boundary, then  $C_\beta$  is unbounded.

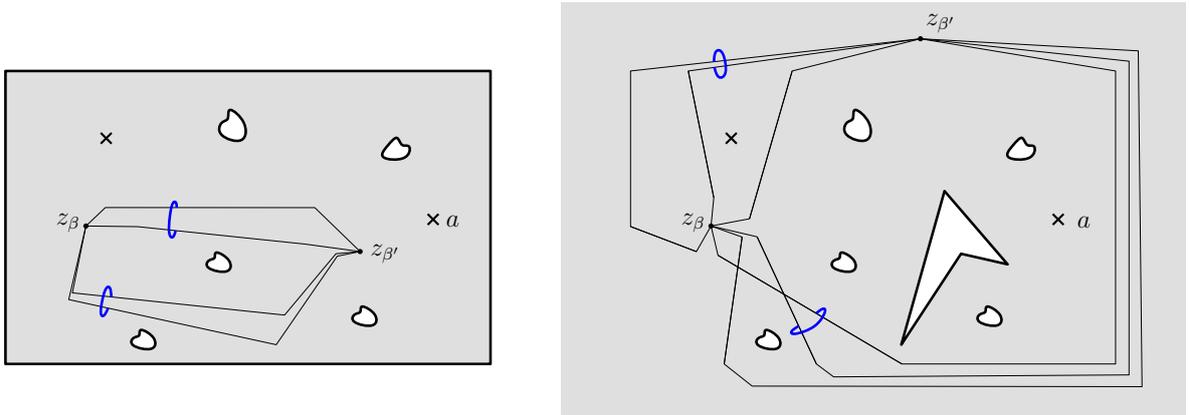


Figure 8: Some of the curves  $\gamma_s$  arising from Fig. 1 and the resulting clusters. In the left case,  $\beta$  and  $\beta'$  are boundaries of holes, while in the right case  $\beta'$  is the exterior boundary. Some curves are drawn with a small perturbation so that their actual common parts are disjoint.

Let  $\beta$  and  $\beta'$  be two boundary components of  $P$ ; it may be that  $\beta = \beta'$ . Let  $S_{\beta, \beta'}$  be the subset of segments from  $S$  with one endpoint in  $\beta$  and another endpoint in  $\beta'$ . We partition  $S_{\beta, \beta'}$  into clusters, as follows. Consider the set  $X_{\beta, \beta'}$  obtained from  $P \setminus \{a, b\}$  by adding  $C_\beta$  and  $C_{\beta'}$ . Note that  $a$  and  $b$  are holes in  $X_{\beta, \beta'}$ . For each segment  $s = \overline{pq} \in S_{\beta, \beta'}$ , with  $p \in \beta$  and  $q \in \beta'$ , we define the following curve  $\gamma_s$ : follow a shortest path in  $C_\beta$  from  $z_\beta$  to  $p$ , then follow  $\overline{pq}$ , and then follow a shortest path in  $C_{\beta'}$  from  $q$  to  $z_{\beta'}$ . See Fig. 8. We say that segments  $s$  and  $s'$  from  $S_{\beta, \beta'}$  are *a-b equivalent* if  $\gamma_s$  and  $\gamma_{s'}$  are homotopic paths in  $X_{\beta, \beta'}$ . Since being homotopic is an equivalence relation (reflexive, symmetric, transitive), being *a-b equivalent* is also an equivalence relation in  $S_{\beta, \beta'}$ . Therefore, we can make equivalence classes, which we call *clusters*. The following two results provide key properties of the clusters.

**Lemma 4.**  *$S_{\beta, \beta'}$  is partitioned into  $\mathcal{O}(h^2)$  clusters. Such partition can be computed in  $\mathcal{O}(hn \log n)$  time.*

*Proof.* Let  $\Gamma_{\beta, \beta'}$  be the set of curves  $\gamma_s$  over all segments  $s \in S_{\beta, \beta'}$ . Note that two curves  $\gamma_s$  and  $\gamma_{s'}$  of  $\Gamma_{\beta, \beta'}$  may cross only once, and they do so along  $s$  and  $s'$ . With a small perturbation of the curves in  $\Gamma_{\beta, \beta'}$  we may assume that  $\gamma_s$  and  $\gamma_{s'}$  are either disjoint or cross at  $s \cap s'$ . (We do not actually use that  $\gamma_s$  contains shortest paths inside  $C_\beta$  and  $C_{\beta'}$  besides for this property of non-crossing curves inside  $C_\beta$  and  $C_{\beta'}$ .)

We now describe a simple criteria using crossing sequences to decide when two segments of  $S_{\beta, \beta'}$  are *a-b equivalent*. We take a set  $\Sigma$  of non-crossing paths in  $X_{\beta, \beta'}$  that have the following property: cutting  $X_{\beta, \beta'}$  along the curves of  $\Sigma$  removes all holes. Such set  $\Sigma$  has a tree-like structure and can be constructed as follows. For each boundary  $\alpha$  of  $P$ , distinct from  $\beta$  and  $\beta'$ , we add to  $\Sigma$  the shortest path in  $P$  between  $a$  and  $\alpha$ . We add to  $\Sigma$  the shortest path in  $P$  between  $a$  and  $b$ . Finally, if  $\beta$  or  $\beta'$  is the exterior boundary of  $P$ , we add to  $\Sigma$  a shortest path from  $a$  to a point that is very far in  $P$  union the the outer face. In total,  $\Sigma$  has  $\mathcal{O}(h)$  polygonal paths in  $X_{\beta, \beta'}$ . Note that the curves in  $\Sigma$  are non-crossing and a small perturbation makes them disjoint, except at the common endpoint  $a$ . See Fig. 9. Each curve  $\sigma \in \Sigma$  is simple and has two sides. We arbitrarily choose one of them as the right side and the other as the left side. We use  $\sigma_1, \dots, \sigma_k$  to denote the curves of  $\Sigma$ .

To each path  $\gamma$  in  $\Gamma_{\beta, \beta'}$  we associate a crossing sequence  $w(\gamma)$  as follows. We start with

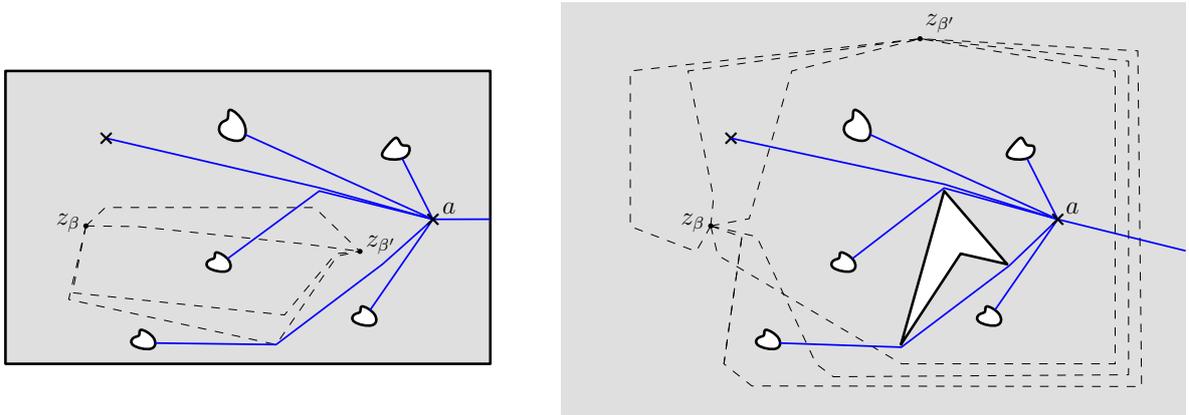


Figure 9: The curves  $\Sigma$  in solid and  $\Gamma_{\beta,\beta'}$  in dashed style for the example of Fig. 8, after a small perturbation.

the empty word and walk along  $\gamma$ . When  $\gamma$  crosses an arc  $\sigma_i \in \Sigma$  from left-to-right we append  $\sigma_i^{\rightarrow}$  to the word, and when  $\gamma$  crosses  $\sigma_i$  from right-to-left we append  $\sigma_i^{\leftarrow}$  to the word. From the crossing sequence  $w(\gamma)$  we can obtain the *reduced crossing sequence*  $w^R(\gamma)$ : we iteratively remove contiguous appearances of  $\sigma_i^{\rightarrow}$  and  $\sigma_i^{\leftarrow}$ , for any  $i$ . For example, from the crossing sequence  $\sigma_1^{\rightarrow}\sigma_2^{\leftarrow}\sigma_3^{\rightarrow}\sigma_3^{\leftarrow}\sigma_2^{\rightarrow}$  we obtain the reduced crossing sequence  $\sigma_1^{\rightarrow}$ . A consequence of using  $\{\sigma_i\}$  to construct the so-called universal cover is the following characterization: the curves  $\gamma_s$  and  $\gamma_{s'}$  are homotopic in  $X_{\beta,\beta'}$  if and only if the curves  $\gamma_s$  and  $\gamma_{s'}$  have the same reduced crossing sequence. See for example [CLMS04] or the lecture notes by Erickson [Eri09]. We conclude that  $s$  and  $s'$  from  $S_{\beta,\beta'}$  are  $a$ - $b$  equivalent if and only if  $w^R(\gamma_s) = w^R(\gamma_{s'})$ .

The union of  $\Sigma$  and  $\Gamma_{\beta,\beta'}$  forms a family of pseudosegments: any two of them crosses at most once. Indeed, by construction different curves can only cross in  $P$ , but inside  $P$  all those curves are shortest paths, and thus can cross at most once. Furthermore, the segments  $\Sigma$  do not cross by construction and the curves of  $\Gamma_{\beta,\beta'}$  have common endpoints. Mount [Mou90, Theorem 1.1] has shown that in such case the curves in  $\Gamma_{\beta,\beta'}$  define at most  $\mathcal{O}(|\Sigma|^2) = \mathcal{O}(h^2)$  distinct crossing sequences. Therefore, there are at most  $\mathcal{O}(h^2)$  homotopy classes defined by the curves in  $\Gamma_{\beta,\beta'}$ , and  $S_{\beta,\beta'}$  defines  $\mathcal{O}(h^2)$  clusters.

The procedure we have described is constructive: we have to compute  $\mathcal{O}(h)$  shortest paths in  $P$  to obtain the curves of  $\Sigma$ , and then, for each segment  $s \in S_{\beta,\beta'}$ , we have to compute the corresponding crossing sequence. Such crossing sequence is already reduced. Note that for computing the crossing sequence of  $\gamma_s$  we never have to construct  $\gamma_s$  itself because all crossings occur along  $s$ . This can be done in  $\mathcal{O}(hn \log n)$  time using algorithms for shortest paths in polygonal domains [HS99] and data structures for ray-shooting among the segments of  $\Sigma$  [CEG<sup>+</sup>94].  $\square$

**Lemma 5.** *For each cluster, either all or none of the segments in the cluster are crossed by a minimum-cost  $a$ - $b$  path.*

*Proof.* Let  $s$  and  $s'$  be two  $a$ - $b$  equivalent segments from  $S_{\beta,\beta'}$ . This implies that  $\gamma_s$  and  $\gamma_{s'}$  are homotopic in  $X_{\beta,\beta'}$ . Therefore, the path  $\gamma$  obtained by concatenating  $\gamma_s$  and the reversal of  $\gamma_{s'}$  is contractible in  $X_{\beta,\beta'}$ .

Let  $\pi$  be a minimum-cost path between  $a$  and  $b$  and assume, for the sake of reaching a contradiction, that  $\pi$  crosses  $s$  but does not cross  $s'$ . We take  $\pi$  that minimizes the total

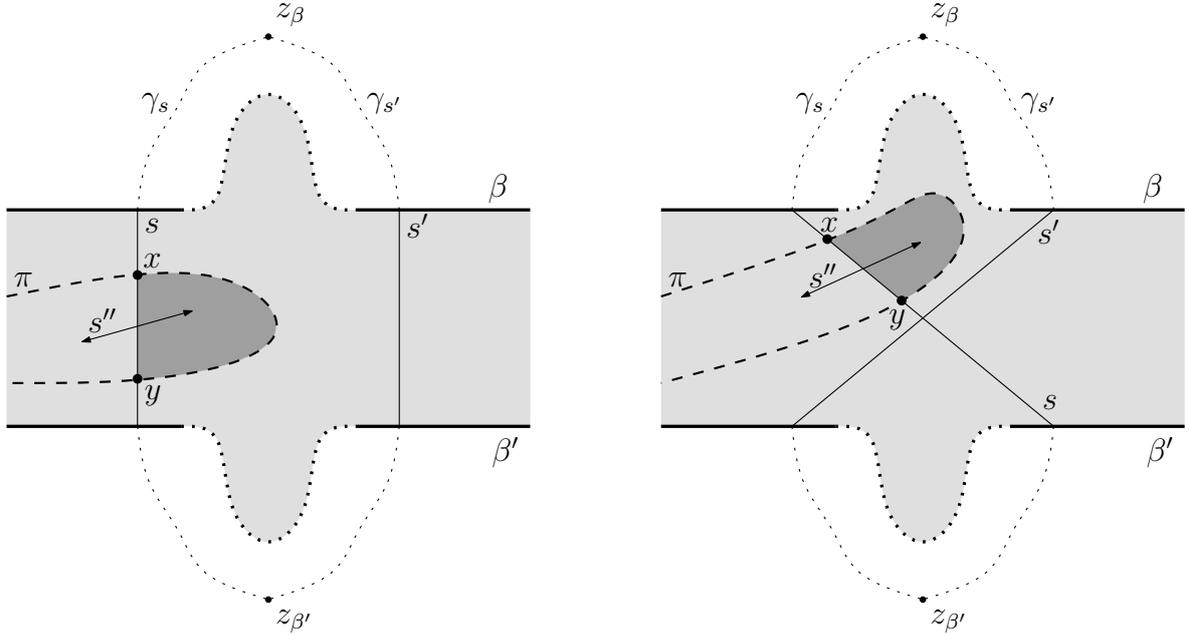


Figure 10: Figure for the proof of Lemma 5. Left: case when  $s$  and  $s'$  are disjoint. Right: case when  $s$  and  $s'$  intersect. In both cases, the darker gray region represents the topological disk defined by  $\pi[x, y]$  and  $\overline{xy}$ .

number of crossings with  $s$ . We may assume that  $\pi$  is simple and disjoint from  $\beta, \beta'$ . We use  $\pi[x, y]$  to denote the subpath of  $\pi$  between points  $x$  and  $y$  of  $\pi$ . We distinguish two cases:

- $s$  and  $s'$  do not intersect. In this case, the curve  $\gamma$  is simple and contractible in  $X_{\beta, \beta'}$ . It follows that  $\gamma$  bounds a topological disk  $D_\gamma$  in  $X_{\beta, \beta'}$ . By hypothesis,  $\pi$  crosses the part of the boundary of  $D_\gamma$  defined by  $s$  but not  $s'$ . Therefore,  $\pi$  must cross at least twice along  $s$ . Let  $x$  and  $y$  be two consecutive crossings of  $\pi$  and  $s$  as we walk along  $\pi$ . See Fig. 10 left. Consider the path  $\pi'$  that replaces  $\pi[x, y]$  by the segment  $\overline{xy}$ . Note that  $\pi[x, y]$  and  $\overline{xy}$  define a topological disk, as it is contained in  $D_\gamma$ . It follows that any segment  $s''$  crossing  $s$  along  $\overline{xy}$  also crosses  $\pi$ . Indeed,  $s''$  has to cross the boundary of the topological disk defined by  $\pi[x, y]$  and  $\overline{xy}$  twice, and cannot cross it at  $\overline{xy}$  again. This means that any segment  $s''$  crossing  $\pi'$  also crosses  $\pi$ . Therefore  $\pi'$  crosses no more segments than  $\pi$  and crosses  $s$  twice less than  $\pi$ . Thus, we reach a contradiction. (If  $\pi'$  is not simple we can take a simple path contained in  $\pi'$ .)
- $s$  and  $s'$  intersect. In this case, the curve  $\gamma$  in  $X_{\beta, \beta'}$  has precisely one crossing. Let  $\gamma'$  and  $\gamma''$  be the two simple loops obtained by splitting  $\gamma$  at its unique crossing. It must be that  $\gamma'$  and  $\gamma''$  are contractible, as otherwise  $\gamma$  would not be contractible. See Fig. 10 right. Therefore, we obtain two topological disks  $D_{\gamma'}$  and  $D_{\gamma''}$ , one bounded by  $\gamma'$  and another by  $\gamma''$ . The path  $\pi$  must cross the boundary of  $D_{\gamma'}$  or  $D_{\gamma''}$ , and the same argument than in the previous item leads to a contradiction.  $\square$

A minimum-cost  $a$ - $b$  path can now be found by testing all possible cluster subsets, that is,  $2^{\mathcal{O}(h^4)}$  possibilities.

**Theorem 6.** *The restricted 2-CELLS-CONNECTION problem in a polygon with  $h$  holes and  $n$  segments can be found in  $2^{\mathcal{O}(h^4)}n$  polylog  $n$  time.*

*Proof.* We classify the segments of  $S$  into  $\mathcal{O}(h^4)$  clusters using Lemma 4. This takes  $\mathcal{O}(h^3n \log n)$  time. Because of Lemma 5, we know that either all or none of the segments in a cluster are crossed by an optimal  $a$ - $b$  path. Each subset of the clusters defines a set of segments  $S'$ , and we can test whether  $S'$  separates  $a$  and  $b$  in  $\mathcal{O}(n \text{ polylog } n)$  time [GSS89, dBDS95].  $\square$

## 4 Connecting all cells

We show that ALL-CELLS-CONNECTION is NP-hard by a reduction from the NP-hard problem of feedback vertex set (FVS) in planar graphs (c.f. [Vaz01]): Given a planar graph  $G$ , find a minimum-size set of vertices  $X$  such that  $G - X$  is acyclic.

First, we subdivide every edge of  $G$  obtaining a planar bipartite graph  $G'$ . It is clear that  $G'$  has a feedback vertex set of size  $k$  if and only if  $G$  has one of size  $k$ . Next, we use the result by de Fraysseix et al. [dFOP91] (see also Hartman et al. [HNZ91]), which states that every planar bipartite graph is the intersection graph of horizontal and vertical segments, where no two of them cross (intersect at a common interior point). Let  $S$  be the set of segments whose intersection graph is  $G'$ ; it can be constructed in polynomial time. Since  $G'$  has no triangles, no three segments of  $S$  intersect at a point. Then, observe that all cells in  $\mathcal{A}(S)$  become connected by removing  $k$  segments if and only if  $G'$  has a feedback vertex set of size  $k$ . Therefore we have:

**Theorem 7.** *ALL-CELLS-CONNECTION is NP-hard even if no three segments intersect at a point and there are no segment crossings.*

It is also easy to see that if no three segments intersect at a point a  $k$ -size solution to ALL-CELLS-CONNECTION corresponds to a  $k$ -size solution of FVS in the intersection graph of the input segments. For general graphs, FVS is fixed-parameter tractable when parameterized with the size of the solution [CFL<sup>+</sup>08], and has a polynomial-time 2-approximation algorithm [Vaz01]. We thus obtain the following:

**Corollary 8.** *When no three segments intersect at a point, ALL-CELLS-CONNECTION is fixed-parameter tractable with respect to the size of the solution and has a polynomial-time 2-approximation algorithm.*

## 5 Conclusion

The main result in this paper is that 2-CELLS-CONNECTION for segments is both NP-hard and APX-hard. There are many interesting open questions, most notably,

- (i) is the problem APX-hard also for the special case of unit-length segments?
- (ii) is the problem NP-hard for unit disks?

## Acknowledgments:

We would like to thank Primož Škraba for bringing to our attention some of the problems studied in this abstract.

## References

- [ACGK11] H. Alt, S. Cabello, P. Giannopoulos, and C. Knauer. On some connection problems in straight-line segment arrangements. In *Abstracts of the 27th European Workshop on Computational Geometry (EuroCG 11)*, pages 27–30, 2011.
- [BK09] S. Bereg and D. G. Kirkpatrick. Approximating barrier resilience in wireless sensor networks. In *Proc. 5th ALGOSENSORS*, volume 5804 of *LNCS*, pages 29–40. Springer, 2009.
- [BLWZ05] H. Broersma, X. Li, G. Woeginger, and S. Zhang. Paths and cycles in colored graphs, p.299. *Australasian J. Combin.*, 31:299–312, 2005.
- [CEG<sup>+</sup>94] B. Chazelle, H. Edelsbrunner, M. Grigni, L. J. Guibas, J. Hershberger, M. Sharir, and J. Snoeyink. Ray shooting in polygons using geodesic triangulations. *Algorithmica*, 12(1):54–68, 1994.
- [CFL<sup>+</sup>08] J. Chen, F. V. Fomin, Y. Liu, S. Lu, and Y. Villanger. Improved algorithms for feedback vertex set problems. *J. Comput. Syst. Sci.*, 74:1188–1198, 2008.
- [CLMS04] S. Cabello, Y. Liu, A. Mantler, and J. Snoeyink. Testing homotopy for paths in the plane. *Discr. & Comp. Geometry*, 31(1):61–81, 2004.
- [dBDS95] M. de Berg, K. Dobrindt, and O. Schwarzkopf. On lazy randomized incremental construction. *Discr. & Comp. Geometry*, 14(3):261–286, 1995.
- [dFOP91] H. de Fraysseix, P. Ossona de Mendez, and J. Pach. Representation of planar graphs by segments. *Intuitive Geometry*, 63:109–117, 1991.
- [Eri09] J. Erickson. Shortest homotopic paths; testing homotopic paths in the plane. Lecture notes available at <http://compgeom.cs.uiuc.edu/~jeffe/teaching/comptop/schedule.html>, 2009.
- [FGI10] M. R. Fellows, J. Guo, and I. Iyad. The parameterized complexity of some minimum label problems. *J. Comput. Syst. Sci.*, 76:727–740, 2010.
- [GSS89] L. J. Guibas, M. Sharir, and S. Sifrony. On the general motion-planning problem with two degrees of freedom. *Discr. & Comp. Geometry*, 4:491–521, 1989.
- [Hås01] J. Håstad. Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001.
- [HMS07] R. Hassin, J. Monnot, and D. Segev. Approximation algorithms and hardness results for labeled connectivity problems. *J. Comb. Optim.*, 14(4):437–453, 2007.
- [HNZ91] I. B. Hartman, I. Newman, and R. Ziv. On grid intersection graphs. *Discrete Mathematics*, 87:41–52, 1991.
- [HS99] J. Hershberger and S. Suri. An optimal algorithm for euclidean shortest paths in the plane. *SIAM J. Comput.*, 28(6):2215–2256, 1999.
- [KH07] S. Kloder and S. Hutchinson. Barrier coverage for variable bounded-range line-of-sight guards. In *Proc. ICRA*, pages 391–396. IEEE, 2007.

- [KKMO07] S. Khot, G. Kindler, E. Mossel, and R. O’Donnell. Optimal inapproximability results for max-cut and other 2-variable csps? *SIAM J. Comput.*, 37(1):319–357, 2007.
- [KLA07] S. Kumar, T.-H. Lai, and A. Arora. Barrier coverage with wireless sensors. *Wireless Networks*, 13(6):817–834, 2007.
- [KLSS13] M. Korman, M. Löffler, R. I. Silveira, and D. Strash. On the complexity of barrier resilience for fat regions. In *Proc. 9th ALGOSENSORS*, volume 8243 of *LNCS*, pages 201–216. Springer, 2013. Full preprint available at <http://arxiv.org/abs/1302.4707>.
- [KYZ14] D. G. Kirkpatrick, B. Yang, and S. Zilles. A polynomial-time algorithm for computing the resilience of arrangements of ray sensors. *Int. J. Comput. Geometry Appl.*, 24(3):225–236, 2014.
- [Mou90] D. M. Mount. The number of shortest paths on the surface of a polyhedron. *SIAM J. Comput.*, 19(4):593–611, 1990.
- [TK12] K.-C. R. Tseng and D. Kirkpatrick. On barrier resilience of sensor networks. In *Proc. 7th ALGOSENSORS*, volume 7111 of *LNCS*, pages 130–144. Springer, 2012.
- [Tse11] K.-C. R. Tseng. Resilience of wireless sensor networks. Master’s thesis, The University Of British Columbia (Vancouver), April 2011.
- [Vaz01] V. V. Vazirani. *Approximation Algorithms*. Springer, 2001.