Approximating Corridors and Tours via Restriction and Relaxation Techniques

ARTURO GONZALEZ-GUTIERREZ and TEOFILO F. GONZALEZ

University of California, Santa Barbara

Abstract. Given a rectangular boundary partitioned into rectangles, the Minimum-Length Corridor (MLC-R) problem consists of finding a corridor of least total length. A corridor is a set of connected line segments, each of which must lie along the line segments that form the rectangular boundary and/or the boundary of the rectangles, and must include at least one point from the boundary of every rectangle and from the rectangular boundary. The MLC-R problem is known to be NP-hard. We present the first polynomial-time constant ratio approximation algorithm for the MLC-R and MLC_k problems. The MLC_k problem is a generalization of the MLC-R problem where the rectangles are rectilinear c-gons, for $c \leq k$ and k is a constant. We also present the first polynomial-time constant ratio approximation algorithm for the Group Traveling Salesperson Problem (GTSP) for a rectangular boundary partitioned into rectilinear c-gons as in the MLC_k problem. Our algorithms are based on the restriction and relaxation approximation techniques.

Categories and Subject Descriptors: F.2.2 [**Analysis of Algorithms and Problem Complexity**]: Nonnumerical Algorithms and Problems—*Gometric problems and computations*; G.4 [**Mathematics of Computing**]: Mathematical Software—*Algorithm design and analysis*

General Terms: Algorithms, Design, Theory

Additional Key Words and Phrases: Corridors, approximation algorithms, restriction and relaxation techniques, complexity, computational geometry

ACM Reference Format:

Gonzalez-Gutierrez, A. and Gonzalez, T. F. 2010. Approximating corridors and tours via restriction and relaxation techniques. ACM Trans. Algor. 6, 3, Article 56 (June 2010), 36 pages.
DOI = 10.1145/1798596.1798609 http://doi.acm.org/10.1145/1798596.1798609

A. Gonzalez-Gutierrez is on leave under the PROMEP program from the Universidad Autónoma de Querétaro, and is supported in part by a Grant from UC MEXUS-CONACyT.

Authors' addresses: A. Gonzalez-Gutierrez, T. F. Gonzalez, Department of Computer Science, University of California, Santa Barbara, CA 93106; email: {aglez, teo}@cs.ucsb.edu.

DOI 10.1145/1798596.1798609 http://doi.acm.org/10.1145/1798596.1798609

1. *Introduction*

An instance I of the Minimum-Length Corridor (MLC-R) problem consists of a pair (F, R), where F is a rectangle partitioned into a set R of rectangles[1] (or rooms) $R_1, R_2, \ldots, R_r$. A *corridor* $T(I)$ for instance I is a set of connected line segments, each of which lies along the line segments that form F and/or the boundary of the rooms, and must include at least one point from every room and from rectangle F. The objective of the MLC-R problem is to construct a corridor of least total length. It is simple to see that the line segments in an optimal corridor do not form any loops, that is, no two points have more than one path between them along an optimal corridor. A generalization of the MLC-R problem where the rooms are rectilinear polygons is called the MLC problem. The MLC problem becomes the MLC_k problem when every room is a rectilinear c-gon, for $c \leq k$ and k is a constant. The MLC problem was initially defined by Naoki Katoh [Demaine and O'Rourke 2001] and subsequently Eppstein [2001] discussed the MLC-R problem. Experimental evaluations of several heuristics for the MLC problem are discussed in Jin and Chong [2003]. The question as to whether or not the decision version of each of these problems is NP-complete is raised in the preceding three references. Mitchell [2000] raised the question as to whether or not the Group Steiner Tree problem (a problem related to the MLC problem) for a set of points in 2D space has a polynomial-time constant ratio approximation algorithm.

Recently Gonzalez-Gutierrez and Gonzalez [2007b], and independently Bodlaender et al. [2006], proved that the decision version of the MLC problem is strongly NP-complete. Gonzalez-Gutierrez and Gonzalez [2007b] also showed that the decision version of the MLC-R problem is strongly NP-complete as well as some of its variants. In virtue of these results, attention has shifted to the corresponding approximation problems.

Bodlaender et al. [2006] consider several restricted versions of the MLC problem. One of these restricted versions of the MLC problem is called the *geographic clustering* problem. In this case, there is a square with side length q which can be enclosed by each room (rectilinear polygon), and the perimeter of each room is bounded above by $c \cdot q$, where $c \geq 4$ is a constant. Clearly, not all instances of the MLC-R problem are geographic clustering problem instances. It is not known whether the decision version of the geometric clustering problem is NP-complete. A Polynomial-Time Approximation Scheme (PTAS) for the geographic clustering version of the MLC and related problems is presented in Bodlaender et al. [2006].

Another restricted version of the MLC problem considered in Bodlaender et al. [2006] is when each room has a size ρ_i defined as the side length of the smallest enclosing square of the room, and each room R_i has perimeter at most $4\rho_i$. A room with this property is said to be a room with *square perimeter*. A room R_i is called α-fat if for every square Q whose boundary intersects R_i and its center is inside R_i, the intersection area of Q and R_i is at least $\frac{\alpha}{4}$ times the area Q. In general $\alpha \in [0, 1]$. For square rooms α is equal to one, but for rectangular rooms α tends to zero. It is not known whether the decision version of the MLC problem, when all the rooms have square perimeter and are α-fat, is an NP-complete problem;

[1] Throughout this article we assume that all the rectangles (and rectilinear polygons) consist only of horizontal and vertical line edges.

however, there is a polynomial-time approximation algorithm with approximation ratio $\frac{16}{\alpha} - 1$ [Bodlaender et al. 2006]. In the case when all the rooms are squares the approximation ratio is 15. Clearly, $\frac{16}{\alpha} - 1$ is not bounded above by a constant for the MLC-R.

The *Group Steiner Tree* (GST) problem may be viewed as a generalization of the MLC problem. Reich and Widmayer [1990] introduced the GST problem, motivated by applications in VLSI design. The GST problem, as it is defined by Reich and Widmayer, consists of a given connected undirected edge-weighted graph $G = (V, E, w)$, where $w : E \rightarrow \mathbb{R}^+$ is an edge-weight function; a nonempty set C, $C \subseteq V$, of *terminals*; and a partition $P = \{C_1, C_2, \ldots, C_k\}$ of C. The objective of the GST problem is to find a tree $T(P) = (V', E')$, where $E' \subseteq E$ and $V' \subseteq V$, such that at least one terminal from each set C_i is in the tree $T(P)$ and the total edge-length $\sum_{e \in E'} w(e)$ is minimized. The graph Steiner Tree (ST) problem is a special case of the GST problem where each set C_i is a single vertex. Karp [1972] proved that the decision version of the ST problem is NP-complete. Since the GST problem includes the ST problem, the decision version of the GST problem is also NP-complete. There is a simple and straightforward reduction from the MLC problem to the GST problem which can be used to show that any constant ratio approximation algorithm for the GST problem is a constant ratio approximation algorithm for the MLC problem.

For an instance of the GST problem in which C is partitioned into k subsets, Ihler [1991] gives a heuristic which has a performance bound of $(k-1) \cdot opt$, where opt is the value of an optimal solution. Bateman et al. [1997] give the first known heuristic with a *sublinear* performance bound of $(1 + \ln(\frac{k}{2})) \cdot \sqrt{k} \cdot opt$. Helvig et al. [2001] give a polynomial-time $O(k^{\epsilon})$-approximation algorithm for any fixed $\epsilon > 0$. However, neither these approximation algorithms nor other known ones for the GST problem are constant ratio approximation algorithms for the MLC-R problem.

Slavik [1997, 1998] and Safra and Schwartz [2006] defined a more general version of the GST problem where $Q = \{C_1, C_2, \ldots, C_k\}$ is not necessarily a partition of C, but each errand i can be performed at any vertex in $C_i \subseteq C$ and $\cup_i C_i = C$, that is, a vertex in C may be in more than one set C_i. This version of the GST problem is called by Slavik [1997, 1998] the *Tree (Errand) Cover* (TEC) problem. The TEC problem is formally defined as follows.

INPUT: A connected undirected edge-weighted graph $G = (V, E, w)$, where $w : E \rightarrow \mathbb{R}^+$ is an edge-weight function; a nonempty set C, $C \subseteq V$, of *terminals*; and a set $Q = \{C_1, C_2, \ldots, C_k\}$, where $C_i \subseteq C$ and $\cup_i C_i = C$.

OUTPUT: A tree $T(Q) = (V', E')$, where $E' \subseteq E$ and $V' \subseteq V$, such that at least one terminal from each set C_i is in the tree $T(Q)$ and the total edge-length $\sum_{e \in E'} w(e)$ is minimized.

Slavik [1997, 1998] developed an approximation algorithm for the TEC problem with approximation ratio 2ρ, when each errand can be performed in at most ρ locations. Safra and Schwartz [2006] established inapproximability results for the 2D version of the TEC problem when each set is connected, but again, these results do not seem to carry over to the MLC-R problem. The TEC and GST (as defined earlier) are computationally equivalent problems. The TEC problem is often referred to as the GST problem. We differentiate between these problems.

As we have seen, our problems are restricted versions of more general problems reported in the literature. However, previous results for those problems do not establish NP-completeness results, inapproximability results, nor constant ratio approximation algorithms for our problems.

In this article we present a polynomial-time approximation algorithm for the MLC-R problem with approximation ratio 30. This is the first constant ratio approximation algorithm for the MLC-R problem. We also present a polynomial-time constant ratio approximation algorithm for the MLC_k problem, when k is a constant.

An application for the MLC problem is when laying optical fiber in metropolitan areas and every block (or set of blocks) is connected through its own gateway which may be placed anywhere on the boundary of the set of blocks. The objective is to find a minimum-length corridor interconnecting all the gateways (one for each set of blocks) in the area. Our problems also have applications in VLSI and floorplanning when laying wires for clock signals or power, and when laying wires for an electrical network or optical fibers for data communications. There has been recent research activity for related problems arising in intelligent transportation as well as in modern spatial database systems for trip planning queries [Li et al. 2005]. Section 6 discusses our problems under other objective functions.

In Section 2 we discuss preliminary results and define the p-MLC-R problem, a restricted version of the MLC-R problem, used to approximate the solution of the MLC-R problem. Then in Section 3 we present our parameterized algorithm for the p-MLC-R problem. The parameterized algorithm takes in a parameter S that identifies a subset of boundary points from each rectangle and calls them *critical points*. The p-MLC-R_S is exactly like the p-MLC-R problem except that every feasible corridor must include a critical point from each rectangle. Then Slavik's approximation algorithm for the TEC problem [Slavik 1997, 1998] is used to generate a corridor for the p-MLC-R_S problem instance. This is the corridor that our algorithm generates for the p-MLC-R problem instance. When the maximum number of critical points identified at each rectangle is k_S, our corridor has length at most $2k_S$ times the length of an optimal corridor for the p-MLC-R_S problem instance. The approximation ratio for this algorithm depends on the ratio (r_S) between optimal solutions for p-MLC-R_S and p-MLC-R problem instances. Therefore, the approximation ratio for the parameterized algorithm is $2k_S \cdot r_S$. There are many selector functions S for which k_S is a constant, for example, the selector function that identifies from each rectangle its corners as critical points has $k_S = 4$. However, for some functions the ratio r_S cannot be bounded above by a constant. In Section 4 we discuss several simple selector functions for which the parameterized algorithm is not a constant ratio approximation algorithm for the p-MLC-R problem because one cannot bound the r_S by a constant. These negative results shed some light for possible design principles that result in constant ratio approximation algorithms. Two constant ratio approximation algorithms arising from the design principles are discussed in Section 5. One selector function identifies from each rectangle two opposite corners and a special point. Special points are defined in Section 4. The main thrust of the article is Section 5 where we prove that r_S is bounded above by a constant for the preceding selector function. In Section 6 we discuss results for related problems, conclusions, and open problems. In particular we discuss a polynomial-time constant ratio approximation algorithm for the MLC_k problem

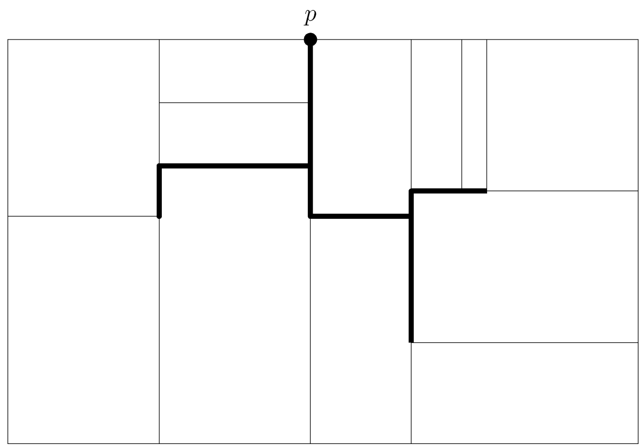


FIG. 1. Optimal corridor for a p-MLC-R problem instance.

for any constant k. We also discuss our approximation algorithm for the Group Traveling Salesperson Problem (GTSP) for a rectangle partitioned into rectilinear c-gons as in the MLC_k problem.

2. *Preliminaries*

Consider the restricted versions of the MLC and MLC-R problems where the input has additionally an access point p, located on the edges of rectangle F, and the solution must include this access point p. We call these problems the *p-access point* version of the problems or simply the p-MLC and p-MLC-R problems. The decision version of each of these problems is shown to be NP-complete in Gonzalez-Gutierrez and Gonzalez [2007b]. An optimal corridor for the p-MLC-R problem instance given in Figure 1 is represented by the thick line segments. The solution to any instance of the MLC (respectively, MLC-R) problem can be obtained by finding a corridor for the p-MLC (respectively, p-MLC-R) problem at each intersection point p located along the edges of rectangle F, and then selecting the best of these corridors. Based on this observation we state the following theorem.

THEOREM 2.1. *Any polynomial-time constant ratio approximation algorithm for the p-MLC-R problem is also a polynomial-time constant ratio approximation algorithm for the MLC-R problem. The approximation ratio is identical for both algorithms.*

PROOF. By the preceding discussion. □

A theorem similar to this one can be established for the MLC and related problems. As a result of Theorem 2.1, we have reduced the MLC-R approximation problem to the p-MLC-R approximation problem. Hereafter we concentrate our efforts on the p-MLC-R approximation problem.

It is convenient to transform the geometric representation of the p-MLC-R problem into the following graph representation. There is a vertex for every distinct point located at the intersection of two orthogonal line segments representing the edges of rectangles and the rectangular boundary F. A vertical (respectively,

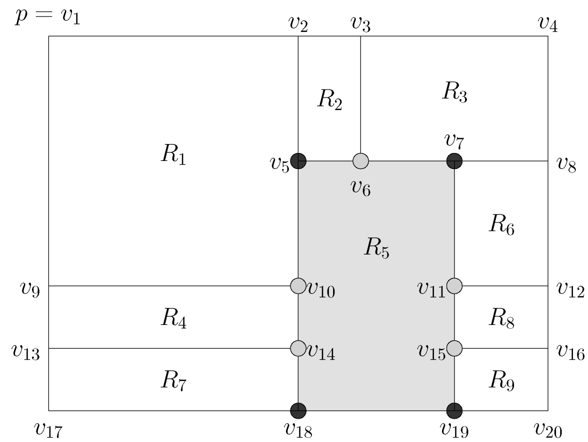


FIG. 2. Instance of the p-MLC-R problem.

horizontal) line segment in the instance I of the p-MLC-R problem is called an *edge* if it includes exactly two points represented by vertices, and the two points are the segment's endpoints. We assume without loss of generality that p is located at a vertex representing a point located on the rectangular boundary F. Every instance I of the p-MLC-R problem is represented by the graph $G(I) = (V, E, w)$, where the set V of vertices and the set of E of edges is defined earlier, and the weight of an edge ($w(e)$) corresponds to the length of the line segment represented by the edge. In this article we use the geometric and graph representation of the p-MLC-R problem interchangeably, and mix the two notations. We use $V(R_i)$ to denote all the vertices located along the boundary of rectangle R_i. Note that $|V(R_i)| \geq 4$ (see Figure 2 where $V(R_5)$ has 9 vertices). We use $C(R_i)$ to denote the set of vertices that corresponds to the corners of R_i. Every vertex is a noncorner point of at most one rectangle (the number of noncorner points of rectangle R_5 is five ($|V(R_5)| - 4$); see Figure 2).

The instance of the TEC problem corresponding to the instance (F, R, p) of the p-MLC-R problem is defined for the (metric) graph $G(V, E, w)$ constructed from (F, R, p) with an errand E_i for each rectangle R_i located at all the vertices $V(R_i)$, plus the errand E_0 located at vertex p. Clearly every feasible solution to the p-MLC-R problem instance is also a feasible solution for the corresponding TEC problem instance and vice versa. Furthermore, the objective function value of every feasible solution to both problem instances is identical.

Let I represent any instance of the p-MLC-R problem. Let $T(I)$ be any corridor for instance I and $t(I)$ be its edge-length. Let *OPT*(I) be an optimal corridor for instance I and let *opt*(I) be its edge-length. An approach to generate suboptimal solutions for the p-MLC-R problem instance (F, R, p) is to construct an instance of the TEC problem and then invoke an approximation algorithm for the TEC problem instance. The solution generated by the algorithm for the TEC problem instance is the solution to the p-MLC-R problem instance. Currently one uses Slavik's [1997, 1998] approximation algorithm for the TEC problem, which is based on *relaxation* techniques. A direct application of this approach to the p-MLC-R problem generates a corridor whose total edge-length is at most $2\rho \cdot opt(I)$, where $\rho = max_i\{|V(R_i)|\}$. Unfortunately, this simple approach does not result in a constant ratio approximation for the p-MLC-R problem since, as we pointed out before, $|V(R_i)|$ is not bounded above by any constant. In the next section we discuss our parameterized approximation algorithm which is a refined version of this approach.

3. *Parameterized Algorithm*

To establish a constant ratio approximation for the p-MLC-R problem, we need to refine our previous strategy that uses an approximate solution to the TEC problem. The idea is to *restrict* the solution space by limiting in each rectangle R_i the possible vertices, from which at least one must be part of the corridor. Consider the p-MLC-R$_S$ problem where the selector function S identifies from each rectangle R_i a set of at most k_S of its boundary points from which at least one must be included by every corridor. The points selected for each rectangle R_i are called the *critical points of* R_i. Usually, the k_S critical points for each rectangle R_i defined by S include a subset of its corner points as well as some points with a special connectivity property. This connectivity property will be defined later on. The objective function of the p-MLC-R$_S$ problem is to find a minimum edge-length corridor that includes for each rectangle R_i at least one of its critical points.

Given S and an instance I of the p-MLC-R problem, we use I_S to denote the instance of the corresponding p-MLC-R$_S$ problem. The instance of the TEC problem, denoted by J_S, is constructed from the instance I_S of the p-MLC-R$_S$ problem using the same approach as the one used for the p-MLC-R problem, but limiting the errands from each rectangle to the critical points of the rectangle. Clearly every feasible solution to the p-MLC-R$_S$ problem instance I_S is also a feasible solution to the instance J_S of the TEC problem, and vice versa. Furthermore, the objective function value of every feasible solution to both problems is identical. Slavik's algorithm applied to the instance J_S of the TEC problem generates a solution $T(J_S)$ from which we construct a corridor $T(I)$ with edge-length $t(I)$ for the p-MLC-R problem. We call our approach the parameterized algorithm $Alg(S)$, where S is the parameter. Since Slavik's approximation algorithm is based on relaxation techniques and we apply it to a restricted version of the p-MLC-R problem, we say that our approximation algorithm is based on *restriction* and *relaxation* approximation techniques. Let $OPT(I_S)$ be an optimal corridor for I_S and let $opt(I_S)$ be its edge-length. Theorem 3.1 establishes the approximation ratio for our parameterized algorithm $Alg(S)$. It is simple to see that the total edge-length of an optimal solution of the instance I_S, $opt(I_S)$, corresponding to the p-MLC-R$_S$ problem, is at least as large as the total edge-length of an optimal solution of the instance I, $opt(I)$, of the p-MLC-R problem. We define the ratio between $opt(I_S)$ and $opt(I)$ as r_S (with $r_S \geq 1$). In other words, one needs to prove that $opt(I_S) \leq r_S \cdot opt(I)$ for every instance I of the p-MLC-R problem, in order to use the following theorem.

THEOREM 3.1. *Parameterized algorithm $Alg(S)$ generates for every instance I of the p-MLC-R problem a corridor $T(I)$ of length $t(I)$ at most $2k_S \cdot r_S$ times $opt(I)$, provided that $opt(I_S) \leq r_S \cdot opt(I)$.*

PROOF. Applying Slavik's approximation algorithm [Slavik 1997, 1998] we generate a corridor $T(I_S)$ of length $t(I_S) \leq 2 \cdot k_S \cdot opt(I_S)$. Clearly $T(I_S)$ is also a corridor for I, so the solution generated, $T(I)$, is simply $T(I_S)$. By our assumption,

$opt(I_S) \leq r_S \cdot opt(I)$. It then follows that the length of the corridor generated by our parameterized algorithm $Alg(S)$ is $t(I) \leq 2k_S \cdot r_S \cdot opt(I)$. $\square$

For the previous approach to yield a constant ratio approximation algorithm we need both k_S and r_S to be bounded above by constants. For example, when S selects from each rectangle R_i its four corner points, k_S is four. However, in order for our parameterized algorithm $Alg(S(4C))$, when $S(4C)$ selects the four corners from each rectangle R_i, to be a constant ratio approximation algorithm for the p-MLC-R problem, we need to show that $opt(I_S) \leq r_S \cdot opt(I)$, for some r_S bounded above by a constant.

For most selector functions S, proving that $opt(I_S) \leq r_S \cdot opt(I)$, for every instance I of the p-MLC-R problem, is difficult because we do not have the optimal solutions at hand. Instead we establish a bound for all corridors. That is, we prove that for every corridor $T(I)$ with edge-length $t(I)$ there is a corridor for I_S denoted by $T(I_S)$ with edge-length $t(I_S) \leq r_S \cdot t(I)$. Applying this to $T(I) = OPT(I)$ we know that there is a corridor $T(I_S)$ such that $t(I_S) \leq r_S \cdot opt(I)$. Since $opt(I_S) \leq t(I_S)$, we know that $opt(I_S) \leq r_S \cdot opt(I)$.

We discuss several selector functions in the next section. For all of these selector functions, r_S cannot be bounded above by a constant. However, we will identify some important characteristics of selector functions that will enable us to come up with one for which r_S can be bounded above by a constant (Section 5).

4. *Selector Functions*

In this section we consider several different basic selector functions S. We show that incorporating them into our parameterized algorithm $Alg(S)$ does not result in constant ratio approximations. These facts point us in the direction of a method for selecting a set of points called *special points*. Special points will turn out to be very important when combined with other critical points to generate constant ratio approximation algorithms (Section 5).

4.1. SELECTING THE FOUR CORNERS. Consider the selector function $S(4C)$ that identifies from each rectangle R_i its four corners $C(R_i)$, that is, the four critical points for R_i are its corner points. Clearly, $k_{S(4C)} = 4$. We now show that $r_{S(4C)}$ cannot be bounded above by a constant. Thus the resulting parameterized algorithm $Alg(S(4C))$ is not a constant ratio approximation algorithm. To prove this we give a family of problem instances with parameter j such that $r_{S(4C)}$ is proportional to j, and j can be made arbitrarily large.

Consider the j-layer family of instances $I(j)$ of the p-MLC-R problem represented in Figure 3. The rectangle F has width 4 and height $\epsilon << 4$. The point p is the top-right corner of F. The rectangle F is partitioned into j layers of rectangles. Layer $i = 1$ is formed by two rectangles of size $(\frac{\epsilon}{2} - \delta) \times 2$, above three rectangles with (very tiny) height δ and width 1, 2, and 1, respectively (see the bottom part of Figure 3). The rectangle in layer 1 with height δ and width 2 is colored gray. Layer $i + 1$ consists of two copies of layer i scaled by 50% laid side by side and placed on top of layer i (see Figure 3). In other words, layer 2 has four rectangles of size $(\frac{\epsilon}{4} - \frac{\delta}{2}) \times 1$, above two sets of three rectangles each with (very tiny) height $\frac{\delta}{2}$ and width $\frac{1}{2}$, 1, and $\frac{1}{2}$, respectively. The two rectangles in layer 2 with height

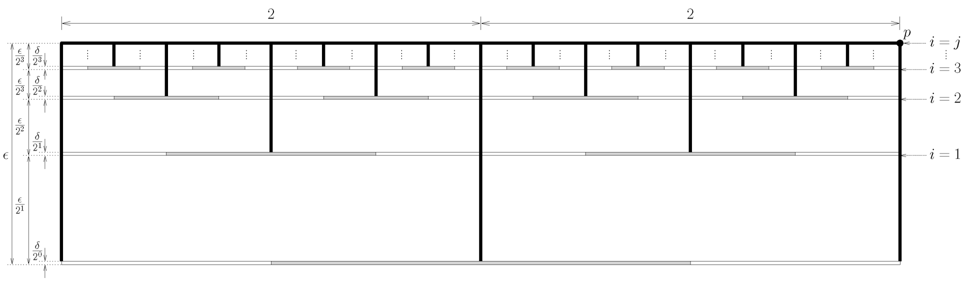


FIG. 3. Optimal solution $OPT(I(j))$ for the family of instances $I(j)$ of the p-MLC-R problem.

$\frac{\delta}{2}$ and width 1 are colored gray. In general, layer $i \geq 1$ has 2^i rectangles of size $(\frac{\epsilon}{2^i} - \frac{\delta}{2^{i-1}}) \times \frac{4}{2^i}$ above 2^{i-1} sets of three rectangles each of (very tiny) height $\frac{\delta}{2^{i-1}}$ and width $\frac{1}{2^{i-1}}$, $\frac{1}{2^{i-2}}$, and $\frac{1}{2^{i-1}}$, respectively. The 2^{i-1} rectangles in layer i with height $\frac{\delta}{2^{i-1}}$ and width $\frac{1}{2^{i-2}}$ are colored gray.

An optimal solution $OPT(I(j))$ of the j-layer family of problem instances $I(j)$ is given by the thick black lines in Figure 3. The total edge-length of the optimal corridor $OPT(I(j))$ is given by

$$opt(I(j)) = 4+2(\epsilon-\delta)+2^0\left(\frac{\epsilon}{2^0}-\frac{\delta}{2^0}\right)+2^1\left(\frac{\epsilon}{2^1}-\frac{\delta}{2^1}\right)+\ldots+2^{j-1}\left(\frac{\epsilon}{2^{j-1}}-\frac{\delta}{2^{j-1}}\right).$$

Thus, $opt(I(j)) = 4 + (j+2)(\epsilon - \delta)$.

It is simple to show that in an optimal solution for instance $I_{S(4C)}(j)$ of the p-MLC-$R_{S(4C)}$ problem, denoted by $OPT(I_{S(4C)}(j))$, there must be a segment of length 1 to connect a corner point of the gray rectangle in layer 1, there must be two segments of length $\frac{1}{2}$ to connect a corner point of the two gray rectangles in layer 2, and so on. Furthermore all of these segments must be distinct. Therefore, the total length of $OPT(I_{S(4C)}(j))$ is $opt(I_{S(4C)}(j)) > j$ and the ratio $r_{S(4C)} = \frac{opt(I_{S(4C)}(j))}{opt(I(j))}$ is greater than $\frac{j}{4+(j+2)(\epsilon-\delta)}$. Making δ and ϵ approach zero, the ratio is about $\frac{j}{4}$, and j can be made arbitrarily large.

Let $S(F4C)$ be a function that identifies fewer corners than $S(4C)$ for each rectangle. It is simple to see that by using the same example, our parameterized algorithm $Alg(S(F4C))$ has an approximation ratio that cannot be bounded above by any constant.

4.2. SELECTING k POINTS AT RANDOM. Randomization is a powerful technique to generate near-optimal solutions to some problems. Lets apply it to restrict the sets $V(R_i)$. Consider the selector function $S(Rk)$ that identifies randomly at most $k \geq 1$ critical points among the vertices of each rectangle. If $k = 7$ the example given in Figure 3 will have $opt(I_{S(Rk)}) = opt(I(j))$ because every rectangle has at most 7 vertices. Actually, one can show that $opt(I_{S(Rk)}) = opt(I(j))$ holds even when $k = 5$. However, there is a large set of problem instances for which the parameterized algorithm $Alg(S(Rk))$ does not generate solutions with an expected approximation ratio bounded above by a constant. These instances include the problem instances given in Figure 3 after stacking k rectangles at every rectangle on the left and right sides of all the gray rectangles [Gonzalez-Gutierrez and Gonzalez 2007a].

4.3. SELECTING A SET OF SPECIAL POINTS. Considering the family of instances given in Figure 3, the only "good" selector functions S for our parametrized algorithm are those that include for every gray rectangle its middle point as a critical point. We call these points *special points*, and we formally define them shortly. For the definition of special point, assume that "rectangle" R_0, which is just point p, is included in R, and p is said to be a corner of R_0. Depending on the selector function S, a subset of the corner points $C(R_i)$ are called the *fixed points* $F(R_i)$ of rectangle R_i. Now, the set of critical points consists of the union of the disjoint sets of fixed points and special points.

The middle point of each gray rectangle in Figure 3 has the *minimum connectivity distance property*. By this we mean, in very general terms, that if given all partial corridors that do not include a point from rectangle R_i, but include points from all other rectangles, then a *special point* of R_i is a vertex in $V(R_i)$ that is not in $F(R_i)$, and the maximum edge-length needed to connect it to each one of the partial corridors is least possible. Finding special points in this way is in general time consuming. Also, this definition is not valid for all problem instances as the set of partial corridors, where every corridor includes vertices from all the rectangles except from R_i, may be empty. In what follows we define special points precisely for all problem instances in a way that is computationally easy to identify a set of special points for each rectangle. Special points are identified using an upper bound on the connectivity distance.

Given that we have selected a set $F(R_i)$ of fixed points for each room (rectangle) R_i, we define a special point as follows. Let $u \in V(R_i)$ and let T_u be a tree of shortest paths rooted at u to all other vertices $(\bigcup_{j \neq i} V(R_j))$ along the edges of rectangle F and the edges of the rooms. Let $SP(u, v)$ be the length of the (shortest) path from vertex u to vertex v along T_u. Let $FP(u, R_j)$ be the length of the (shortest) path from point $u \in V(R_i)$ to the "farthest" vertex of rectangle R_j along T_u, for $i \neq j$, that is,

$$FP(u, R_j) = max_{v \in V(R_j)}\{SP(u, v) | u \in V(R_i),\, j \neq i\}.$$

In other words, the edge-length needed to connect vertex u of room R_i to any corridor through the connection of room R_j is at most $FP(u, R_j)$.

We define the *connectivity distance* $CD(u, R)$ of vertex u in room R_i as

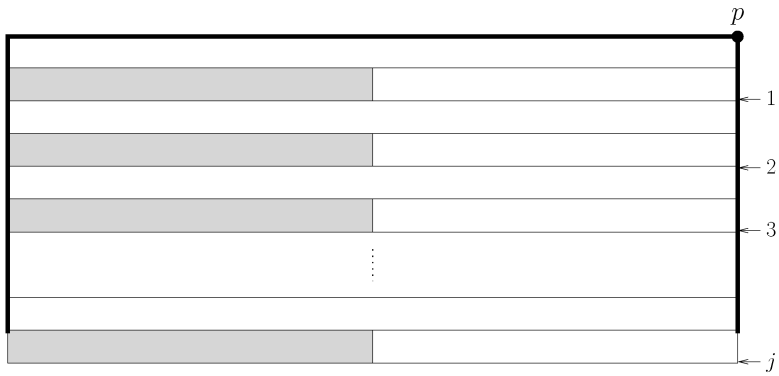
$$min_{j \neq i}\{FP(u, R_j) | R_j \in R\}.$$

If $F(R_i) \subset V(R_i)$ we define the connectivity distance $CD(R_i, R)$ of room R_i as

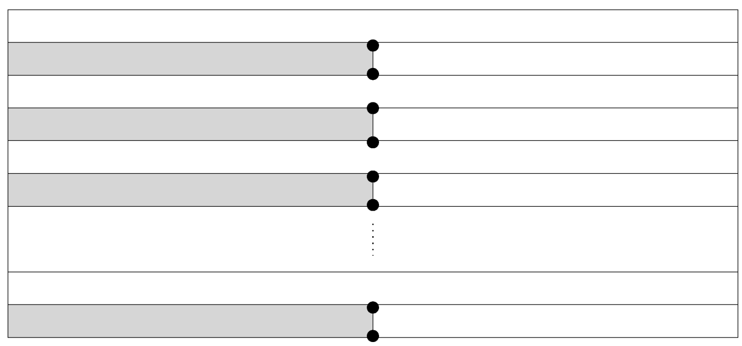
$$min_{u \in V(R_i) \setminus F(R_i)}\{CD(u, R)\}.$$

In other words, $CD(R_i, R)$ is the edge-length needed to connect some specific vertex in $V(R_i) \setminus F(R_i)$ to any corridor through the connection of another room. The special point of R_i is a vertex $u \in V(R_i) \setminus F(R_i)$ such that $CD(u, R) = CD(R_i, R)$. Notice that there may be more than one point satisfying this condition, in which case we select any of these points as the special point. When $F(R_i) = V(R_i)$ then there is no special point. It is important to remember that for the definition of special point, R_0 which is simply p is included in R.

Consider now the selector function $S(+)$ that identifies one special point from each room R_i. The special point for R_i is referred to as SpP_i. In this case, $F(R_i)$ is the empty set. For the problem instance given in Figure 3, $opt(I(j)) = opt(I_{S(+)}(j))$. However, as we shall see shortly, this property does not hold in general. Consider the



(a) optimal corridor of length $opt(I(j))$



(b) two candidates of special points for each gray rectangle

FIG. 4. Family of instances $I(j)$ of the p-MLC-R problem.

j-layer family of instances $I(j)$ of the p-MLC-R problem given in the Figure 4(a), where the height h of F is very small compared to its width w. The rectangle F is partitioned into j layers of rectangles. Each layer is formed by one rectangle of size $w \times \delta$ on top of two rectangles, each of size $\frac{w}{2} \times \delta$, where $2j\delta = h$. An optimal solution is represented by the thick black lines in Figure 4(a), and its total edge-length $opt(I(j)) = w + 2(h - \delta)$. The special point of each gray rectangle is either its top-right or bottom-right corner. An optimal solution for the instance $I_{S(+)}$ must include a segment from a rightmost corner of each gray rectangle to the left- or right-hand side of rectangle F. The total edge-length of the optimal solution for the instance $I_{S(+)}$ is $opt(I_{S(+)}) = j\frac{w}{2} + h - \delta$. Making δ and h approach zero, the ratio $\frac{opt(I_{S(+)})}{opt(I(j))}$ is about $\frac{j}{2}$, and j can be made arbitrarily large. Therefore, $r_{S(+)}$ is not bounded above by any constant. Thus, the restriction $S(+)$ does not result in constant ratio approximations.

The selector function $S(K+)$ consisting of $k > 1$ special points from each rectangle does not result in constant ratio approximations. There are many instances that show that the ratio $r_{S(K+)}$ is not bounded by any constant. These instances include the problem instances given in Figure 4 after stacking k rectangles at every rectangle on the right side of each of the gray rectangles [Gonzalez-Gutierrez and Gonzalez 2007a].

4.4. SELECTING TWO ADJACENT CORNERS AND ONE SPECIAL POINT. The case for $S(2AC+)$ having as critical points two adjacent corners and a special point does not result in a constant ratio approximation, that is, $k_{S(2AC+)}$ is 3, but $r_{S(2AC+)}$ is not bounded above by a constant. The same family of instances of the p-MLC-R problem given in the previous subsection can be used to establish this claim [Gonzalez-Gutierrez and Gonzalez 2007a]. Obviously, the case for S consisting of one corner and one special point does not result in constant ratio approximations.

5. *Constant Ratio Approximation Algorithms*

In this section we present two selector functions that result in constant ratio approximations for the parameterized algorithm. These approximations turn out to be the first two polynomial-time constant ratio approximation algorithms for the p-MLC-R problem and, by Theorem 2.1, for the MLC-R problem.

Our approach is motivated by two facts. First, the family of instances of the p-MLC-R problem given in Figure 3 suggests the selection of one special point, if any, from each rectangle R_i. Second, the family of instances of the p-MLC-R problem discussed in Sections 4.3 and 4.4 suggest the need of at least two opposite corners of each rectangle R_i.

We have analyzed the two most promising approaches. The first approach consists of the selection of two opposite corners and one special point from each rectangle R_i. The second one consists of the selection of the four corners and one special point from each rectangle R_i. Both approaches result in algorithms with identical approximation ratios. As it turns out, the analysis for the first approach is considerably easier than the second one. In Section 5.2 we discuss and prove that the first approach results in a constant ratio approximation. In Section 5.3 we briefly sketch the analysis for the second approach. In Section 5.1 we define new terms that are used extensively throughout this section.

5.1. PRELIMINARIES AND DEFINITIONS. Let $T(I)$ be a corridor for instance I of the p-MLC-R problem and let $S(2OC+) = \{TR, BL, SpP\}$ be the selector function that identifies from each rectangle the top-right and bottom-left corners, and a special point. Now, overlap the corridor $T(I)$ with the rectangles in R. All the rectangles that do not have a critical point along the corridor are called *no-critical-point-exposed* rectangles (*ncpe* rectangles). The remaining rectangles are called *critical-point-exposed* rectangles (*cpe* rectangles). Notice that the definition of the ncpe rectangles depends on the points selected by the selector function $S(2OC+)$. Figure 5 shows an instance I of the p-MLC-R problem with a corridor $T(I)$ and all the ncpe rectangles defined by using the selector function $S(2OC+)$ (the cpe rectangles are omitted).

We say that the *edges* of the ncpe rectangle n_i are the line segments consisting of $\{BR_i\} \cup P_i^r$, $\{BR_i\} \cup P_i^b$, $\{TL_i\} \cup P_i^l$, or $\{TL_i\} \cup P_i^t$, where BR_i and TL_i are the bottom-right and top-left corners of R_i, respectively; and P_i^r, P_i^b, P_i^l, and P_i^t are the right, bottom, left, and top side of R_i, respectively. Note that P_i^r, P_i^b, P_i^l, and P_i^t do not include the corners of R_i.

Let τ be the counterclockwise tour of corridor $T(I)$. This tour starts at the access point p moving in the counterclockwise direction with respect to p on the exterior of F. Then τ traverses the edges of the corridor without "crossing" it and ends again at the access point p. The tour τ associated to the corridor $T(I)$ of the instance I given in Figure 5 is shown in Figure 6. The counterclockwise direction of τ is

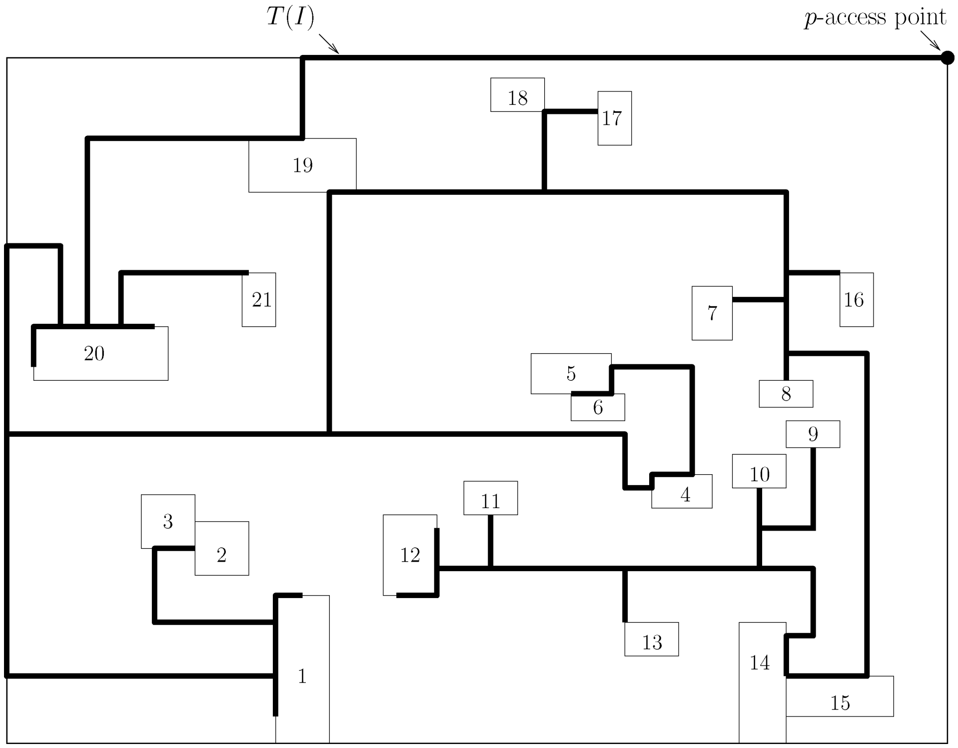


FIG. 5. Corridor $T(I)$ and ncpe rectangles defined by $S(2OC+) = \{TR, BL, SpP\}$ for an instance I of the p-MLC-R problem.

pointed out by the circular arrow at the p-access point. Note that there exists a clockwise tour that traverses the corridor in the opposite direction of τ.

Consider now the instance I of the p-MLC-R problem given in Figure 7(a), where τ visits first all the rectangles located below and to the right of $T(I)$ and then the ones above and to the left of $T(I)$. For complex corridors we do not have a consistent pattern of "below" before "above" or vice versa. For example, rectangle R_i is visited before rectangle R_j by the tour τ of Figure 7(b), which begins on the top part at the access point p in the counterclockwise direction (as pointed out by the arrow). The first point where R_i is visited by τ is w_i and the first point where R_j is visited by τ is w_j. For our instance of Figure 5, the numbers inside the ncpe rectangles denote the order in which they are visited by the counterclockwise tour τ given in Figure 6.

Lets mark the tour τ at the first point where each ncpe rectangle is visited for the first time. The order in which the ncpe rectangles are first visited is called the *canonical order* of the ncpe rectangles. Rename the ncpe rectangles according to this canonical order as $VR = \{n_1, n_2, \ldots, n_q\}$. For example, in Figure 7(b), n_1 is R_i and n_2 is R_j. For $1 \leq i \leq q$, we define the *reaching point* RP_i of the ncpe rectangle n_i as the first point in the ncpe rectangle n_i visited by the tour τ; and the *leaving point* LP_i of the ncpe rectangle n_i as the last point in the ncpe rectangle n_i visited by τ before it visits the reaching point RP_{i+1} of rectangle n_{i+1} (see Figure 8). For convenience we add the points LP_0 and RP_{q+1} which correspond to the access point $p = R_0$.

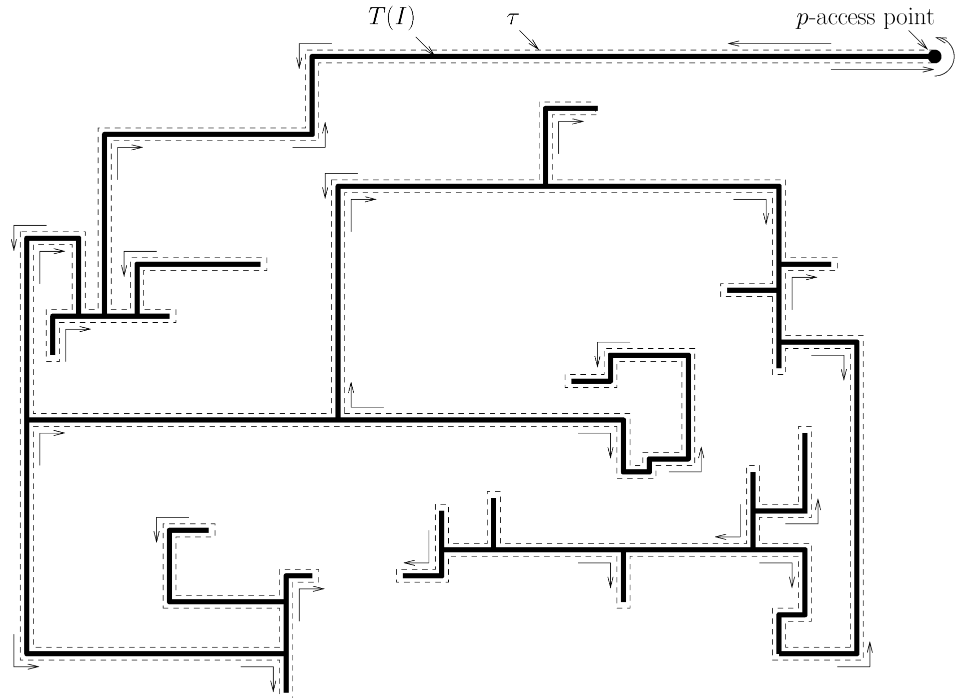


FIG. 6. Counterclockwise tour τ for corridor $T(I)$.

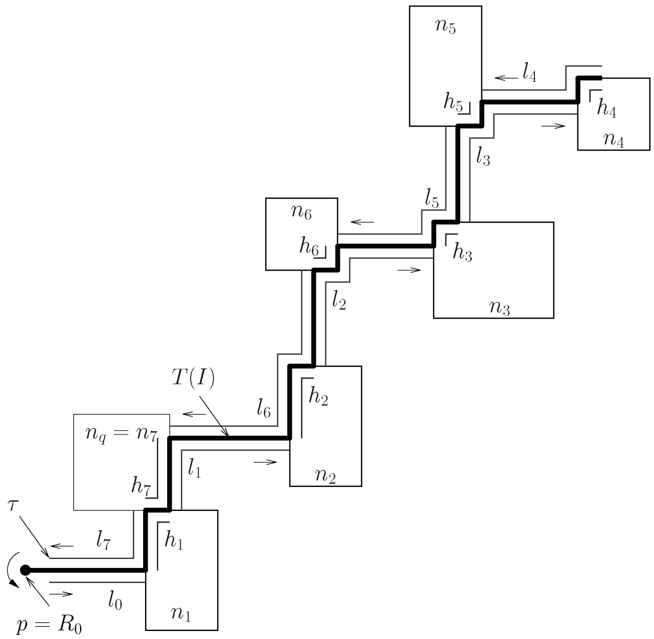
Let $\tau(Z_1, Z_2)$ be the path along tour τ from the point Z_1 to the point Z_2, where $Z_1, Z_2 \in \{LP_0, LP_1, \ldots, LP_q, RP_1, \ldots, RP_{q+1}\}$ and Z_1 appears before Z_2 on the tour τ. For $0 \leq i \leq q$, let l_i be the length of the path $\tau(LP_i, RP_{i+1})$. Let h_j be the length of the path $\tau(RP_j, LP_j)$, for $1 \leq j \leq q$. Figures 7(a) and 8(a) show line segments of length l_i and h_j for a set of ncpe rectangles.

By using the shortest path along the portion of the corridor $T(I)$ from LP_i to RP_{i+1}, we define the *exit point* X_i of ncpe n_i and the *entry point* Y_{i+1} of n_{i+1} as the intersection of the shortest path with the edges of n_i and n_{i+1}, respectively. For convenience we also add the points X_0 and Y_{q+1} which correspond to the access point p. For $i < j$ we define $T(X_i, Y_j)$ as the shortest path along the corridor $T(I)$ from the exit point X_i to the entry point Y_j, and $t(X_i, Y_j)$ denotes its total edge-length. It is simple to prove that $t(X_i, Y_{i+1}) \leq l_i$ for $0 \leq i \leq q$, and $t(Y_i, X_i) \leq h_i$ for $1 \leq i \leq q$ (see Figure 8).

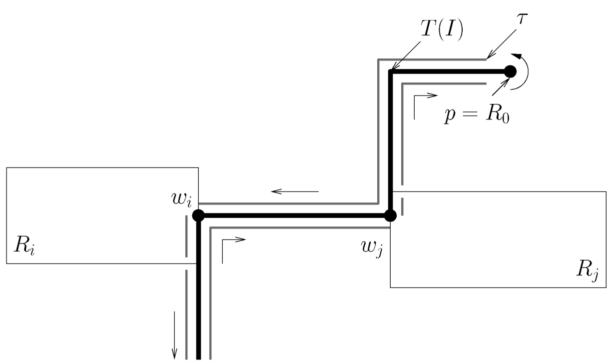
5.2. Selecting Two Opposite Corners and One Special Point.

In this subsection we analyze the parameterized algorithm $Alg(S)$ when each set $V(R_i)$ is restricted to two opposite corners and one special point by the selector function $S(2OC+)$. Without loss of generality, we select the top-right and bottom-left corners as the two opposite corners. Clearly $k_{S(2OC+)} = 3$, and we will prove later on that $r_{S(2OC+)} = 5$. By Theorem 3.1 this results in an approximation ratio 30 for the parameterized algorithm.

THEOREM 5.1. *For every instance I of the p-MLC-R problem, the parameterized algorithm $Alg(S(2OC+))$ generates, in polynomial time with respect to the*



(a) line segments of τ of length l_i and h_j



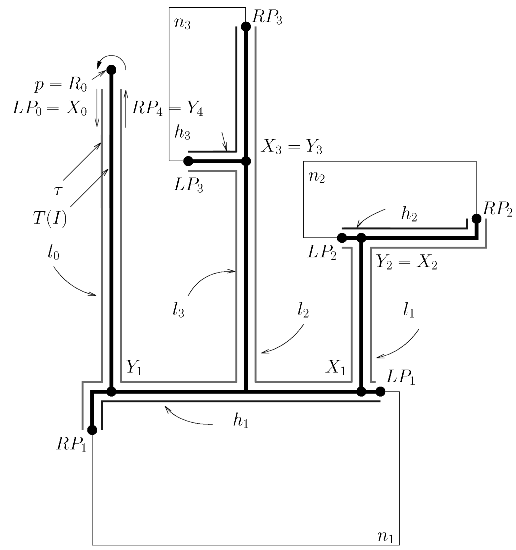
(b) points w_i, w_j where R_i, R_j are visited the first time by τ, respectively

FIG. 7. Two sets of ncpe rectangles and tour τ for corridor $T(I)$.

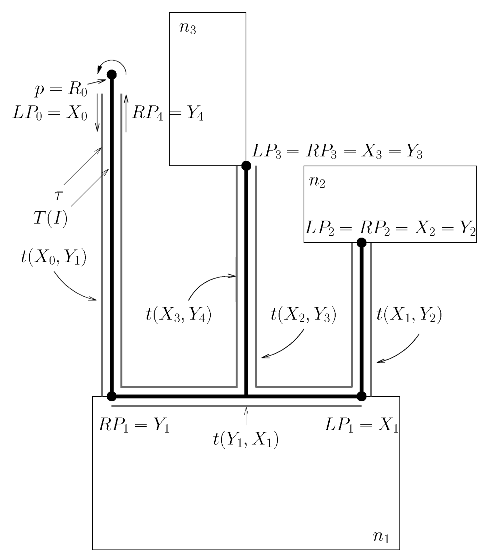
total number r of rectangles in the partition of F, a corridor of length at most $30 \cdot opt(I)$.

PROOF. By the preceding discussion. The time complexity of the algorithm is the one for Slavik's algorithm [Slavik 1997, 1998], which is bounded above by a polynomial in terms of the number of rectangles. □

By applying Theorem 2.1 we know that selecting the best corridor generated by the parameterized algorithm $Alg(S(2OC+))$, when executing it with every point p being a vertex located along the boundary of F, results in a constant ratio approximation algorithm for the MLC-R problem. We call this algorithm the general parameterized algorithm $ALG(S(2OC+))$.



(a) X_i and LP_i, and Y_i and RP_i are not identical pair of points, except when $i \in \{0, 4\}$



(b) X_i matches LP_i, and Y_i matches RP_i, for $0 \leq i \leq 4$

FIG. 8. Corridor $T(I)$ and tour τ visiting three ncpe rectangles.

COROLLARY 5.1. *The general parameterized algorithm $ALG(S(2OC+))$ generates a corridor of length at most $30 \cdot ov(I)$ for every instance I of the MLC-R problem, where $ov(I)$ is the edge-length of a minimum-length corridor for problem instance I.*

PROOF. By the preceding discussion. □

In order to establish our result that the parameterized algorithm with $S(2OC+)$ is a constant ratio approximation algorithm we need to prove Theorem 5.2. This theorem uses Lemma 5.4 where we show that for each ncpe rectangle n_i one can connect one of its critical points to the corridor by adding line segments of length at most $l_{i-1} + h_i + l_i$.

THEOREM 5.2. *Given any corridor $T(I)$ of length $t(I)$ for any instance I of the p-MLC-R problem, there is a corridor for instance $I_{S(2OC+)}$ of the p-MLC-$R_{S(2OC+)}$ problem with edge-length $t(I_{S(2OC+)}) < 5 \cdot t(I)$.*

PROOF. By Lemma 5.4, which we establish shortly, the corridor $T(I)$ can be extended to reach a critical point of each ncpe rectangle n_i by adding a set of line segments of length at most $l_{i-1} + h_{i-1} + h_i$. Adding all of these terms we know that the total length of the line segments that need to be introduced is at most $l_0 + \sum_{j=1}^{q} h_j + 2\sum_{j=1}^{q-1} l_j + l_q < 4t(I)$, since $\sum_{j=0}^{q} l_j + \sum_{j=1}^{q} h_j = |\tau| = 2t(I)$, and $|\tau|$ represents the total edge-length of τ (see Figure 7(a)). Therefore, the total length of the additional line segments plus the length of the corridor T is less than $5t(I)$. So the resulting corridor $T(I_{S(2OC+)})$ has edge-length $t(I_{S(2OC+)}) < 5t(I)$. $\square$

To complete the proof of our claims we need to prove that a critical point of ncpe rectangle n_i can be connected to the corridor by adding line segments of length at most $l_{i-1} + h_i + l_i$ (Lemma 5.4). The proof of Lemma 5.4 is based on the following general approach. First we define a special type of turns in paths, which is called an *inversion of direction* (or simply *iod-subpaths*) of a path. Iod-subpaths are vertical or horizontal. We then define the *beam* of an iod-subpath as the set of points that are "visible" from an iod-subpath. A path is said to be *type-1* if it has vertical and horizontal iod-subpaths, otherwise it is said to be *type-2*. In Lemma 5.1 we establish that if a path has both vertical and horizontal iod-subpaths, then it has two *adjacent* ones whose beams intersect. When this condition holds, we can show that there is at least one rectangle in R that is completely inside the "region" of the adjacent iod-subpaths (Lemma 5.2). This property is used to establish that from either end of the path that includes the adjacent iod-subpaths, or from a point which we call the bifurcation point to one of the iod-subpaths, there exist rectangles completely inside the region of the adjacent iod-subpaths. At least one of these rectangles has its bottom-side (right-side) completely overlapping with the path between the two adjacent iod-subpaths (Lemma 5.3).

Lemmas 5.1, 5.2, and 5.3 are used to prove Lemmas 5.4, 5.5, 5.6, and 5.8. The idea behind the proof of Lemma 5.4 is to examine the paths $T(X_{i-1}, Y_i)$ and $T(X_{i-1}, Y_{i+1})$ neighboring ncpe rectangle n_i. These paths will be labeled either type-1 or type-2 depending on some of their properties. If $T(X_{i-1}, Y_i)$ (respectively, $T(X_{i-1}, Y_{i+1})$) is type-1 then in Lemma 5.5 (respectively, Lemma 5.6) we show, by using Lemma 5.3, that the special point in n_i can be connected to the corridor by adding line segments of length at most l_{i-1} (respectively, $l_{i-1} + h_i + l_i$). For the remaining case, we characterize the form of the type-2 paths $T(X_{i-1}, Y_i)$ and $T(X_{i-1}, Y_{i+1})$ (Lemma 5.7). Then in Lemma 5.8 we use this characterization to show that the path $T(Y_i, Y_{i+1})$ is type-1, and we show that the special point of n_i can be connected to the corridor by adding line segments of length at most $l_{i-1} + h_i + l_i$.

We need to introduce additional notation and establish preliminaries results (Lemmas 5.5–5.8) before we formulate Lemmas 5.1–5.4. As we traverse the path $T(X_i, Y_j)$ from X_i to Y_j we identify a sequence of alternating horizontal and

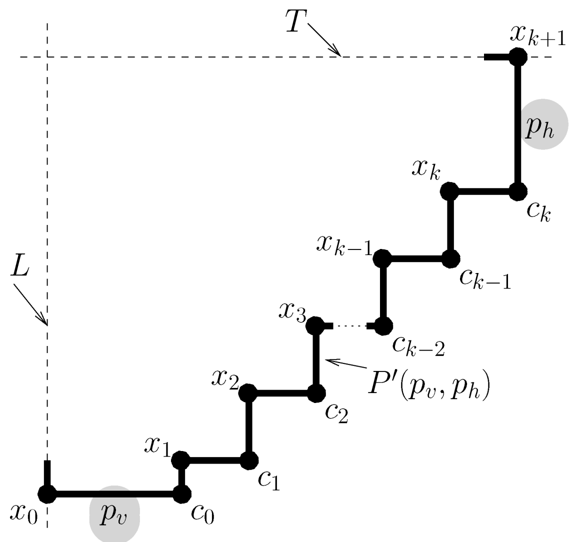


FIG. 9. Pair of adjacent iod-subpaths.

vertical line segments, $s_1, s_2, \ldots, s_l$, such that the first endpoint of s_1 visited is X_i; the last endpoint of s_{k-1} visited coincides with the first endpoint of s_k, for $1 < k \leq l$; and the last endpoint of s_l visited is Y_j. The length of each of these segments is greater than zero, except when $X_i = Y_j$.

The line segments s_{k-1} and s_k, for $1 < k < l$, in path $T(X_i, Y_j)$ are said to be *adjacent*. For the path $T(X_i, Y_j)$ we say that the line segment s_1 is *adjacent* to the edge of ncpe n_i where X_i is located if s_1 intersects the edge, and s_1 and the edge are perpendicular. We define adjacency for s_l and an edge of n_j similarly.

An *inversion of direction subpath* (or *iod-subpath*) of a path $T(X_i, Y_j)$ and ncpe rectangles n_i and n_j, consists of a line segment s_k of the path and on each of its two ends there is either an adjacent edge of ncpe n_i or n_j, or an adjacent line segment s_{k-1} or s_{k+1}; and adjacent line segments and/or portions of the adjacent edges of ncpe rectangles must be on the same side of s_k (i.e., the segments and/or portions of the edges are on the same side of the line that completely includes s_k). The line segment s_k of an iod-subpath is called the *central segment* and the two adjacent line segments and/or edges are called the *end segments*. If the end segments are vertical then it is a *vertical iod-subpath*, otherwise it is a *horizontal iod-subpath*. Figure 9 shows a vertical (horizontal) iod-subpath labeled p_v (p_h).

The subpath $P'(p_v, p_h)$ between the two iod-subpaths p_v and p_h of a path $T(X_i, Y_j)$ consists of all the line segments in the path between the farthest (with respect to the distance along the path) endpoints of their central line segments (path from x_0 to x_{k+1} in Figure 9). Two iod-subpaths of a path are said to be *adjacent* if the subpath $P'(p_v, p_h)$ between them does not contain another iod-subpath.

A vertical iod-subpath p_v is said to be to the left of a horizontal iod-subpath p_h if its central segment is completely located to the left of the vertical line that completely includes the central line segment of p_h. A vertical iod-subpath with portions of its two vertical end segments above its central line segment is called an *up-vertical* iod-subpath. If portions of its two vertical end segments are below its central line segment, the vertical iod-subpath is called *down-vertical* iod-subpath.

We define *right-horizontal* or *left-horizontal* iod-subpath similarly. The *beam* of an iod-subpath consists of all the points visited by the central line segment as we move it perpendicularly and in the direction of the end segments of the iod-subpath until it reaches the rectangular boundary F.

We say that point x precedes the point y along a subpath $P'(p_v, p_h)$, denoted by $x \prec y$, if x is visited before y when traversing the subpath $P'(p_v, p_h)$ from p_v to p_h. We now establish a fundamental lemma which will be used extensively in the proofs of subsequent lemmas.

LEMMA 5.1. *If a path (a, b) contains both vertical and horizontal iod-subpaths, then it contains a vertical iod-subpath adjacent to a horizontal iod-subpath, and their beams intersect.*

PROOF. Suppose we traverse the path from a to b. Since the path contains vertical and horizontal iod-subpaths, let p_v^{first} be the first vertical iod-subpath, and p_h^{first} be the first horizontal iod-subpath. Without loss of generality, assume that p_v^{first} precedes p_h^{first}. Let p_v^{last} be the last vertical iod-subpath that precedes p_h^{first}. Assume without loss of generality that p_v^{last} is an up-vertical iod-subpath, and that as we traverse the path from a to b we will exit p_v^{last} through its right end. Consider the subpath $P'(p_v^{last}, p_h^{first})$. Since there are no iod-subpaths between p_v^{last} and p_h^{first} it must be that as we traverse the path from p_v^{last} to p_h^{first} the horizontal segments are traversed from left to right and the vertical ones are traversed in the upwards direction, as otherwise there would be a vertical or horizontal iod-subpath before p_h^{first}, which contradicts our earlier assumptions. This means that p_h^{first} is a left-horizontal iod-subpath and the beams of p_v^{last} and p_h^{first} intersect. $\square$

A pair of adjacent vertical p_v and horizontal p_h iod-subpaths whose beams intersect are said to be in *canonical form* if p_v is an up-vertical iod-subpath located to the left of p_h. A proof similar to the one of Lemma 5.1 can be used to show that as we traverse the subpath $P'(p_v, p_h)$ from p_v to p_h one only moves up and to the right (see Figure 9). Let L be the leftmost vertical line traced by the beam of p_v, and let T be the topmost horizontal line traced by the beam of p_h. The area delineated by L, $P'(p_v, p_h)$ and T is called the region of the canonical iod-subpaths (p_v-p_h) and it is specified by $r(p_v, p_h)$ (see Figure 9).

We say that a rectangle $R_i \in R$ is *contained* by $r(p_v, p_h)$ if it is completely inside $r(p_v, p_h)$. A rectangle $Z_b \in R$ ($Z_r \in R$) is said to be *bottom (right) contained* if Z_b (Z_r) is contained by $r(p_v, p_h)$ and its bottom (right) edge is completely contained by a horizontal (vertical) line segment of the subpath $P'(p_v, p_h)$.

Traverse the subpath $P'(p_v, p_h)$ from p_v to p_h. Label the first point x_0; label each corner point visited as $c_0, x_1, c_1, x_2, \ldots, c_{k-1}, x_k, c_k$; and label the last point x_{k+1} (see Figure 9). Let $CP = \{c_0, c_1, \ldots, c_k\}$ and let $XP = \{x_1, x_2, \ldots, x_k\}$. For $0 \leq i \leq k$, let $Z_i \in R$ be the rectangle that has its bottom-right corner on the corner point $c_i \in CP$.

LEMMA 5.2. *Given any pair of adjacent vertical (p_v) and horizontal (p_h) iod-subpaths in canonical form, there are both a bottom-side contained rectangle $Z_b \in R$ and a right-side contained rectangle $Z_r \in R$ in $r(p_v, p_h)$.*

PROOF. Consider first the case when Z_0 has its top-edge above line T. It follows that Z_k has its left edge inside the region $r(p_v, p_h)$, and Z_k is a right-side contained

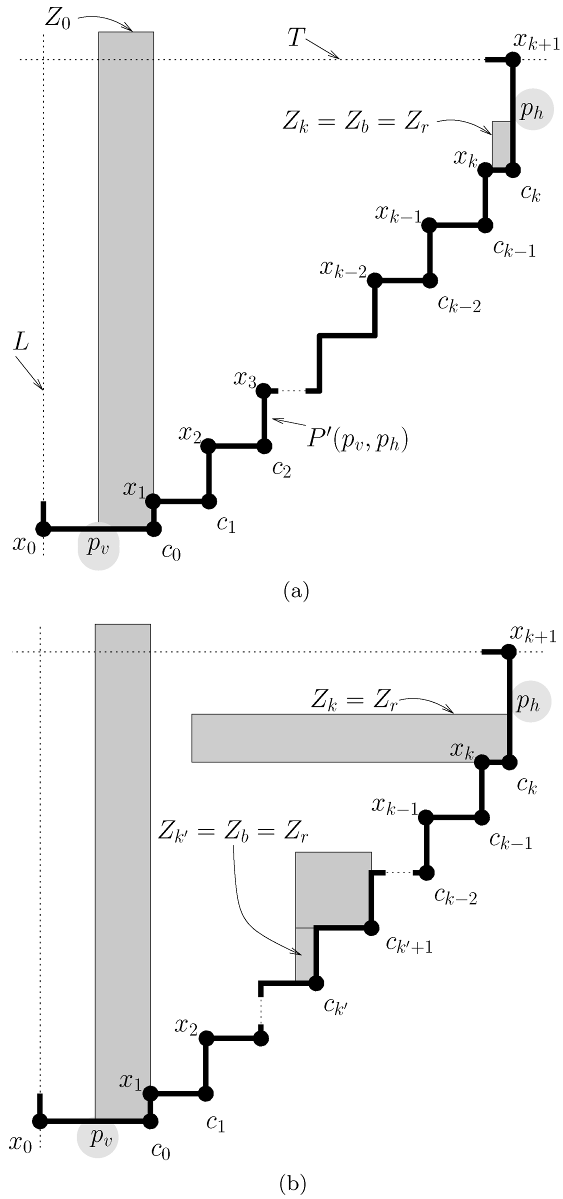


FIG. 10. Bottom-side and right-side contained rectangles Z_b and Z_r.

rectangle (i.e., Z_r). If Z_k is bottom-side contained then Z_k is both Z_b and Z_r, and the lemma follows (see Figure 10(a)).

Otherwise, Z_k is not a bottom-side contained rectangle, and its bottom side extends to the left of the corner point x_k (see Figure 10(b)). Let k' be the largest positive integer such that the bottom side of Z_l for every $l \leq k'$ is completely contained by a horizontal segment of $P'(p_v, p_h)$. Clearly such an l exists as the property holds for Z_0 but does not for Z_k. Since $Z_{k'+1}$ extends to the left of the vertical line with the x-coordinate value of $c_{k'}$, it follows that $Z_{k'}$ is in the region

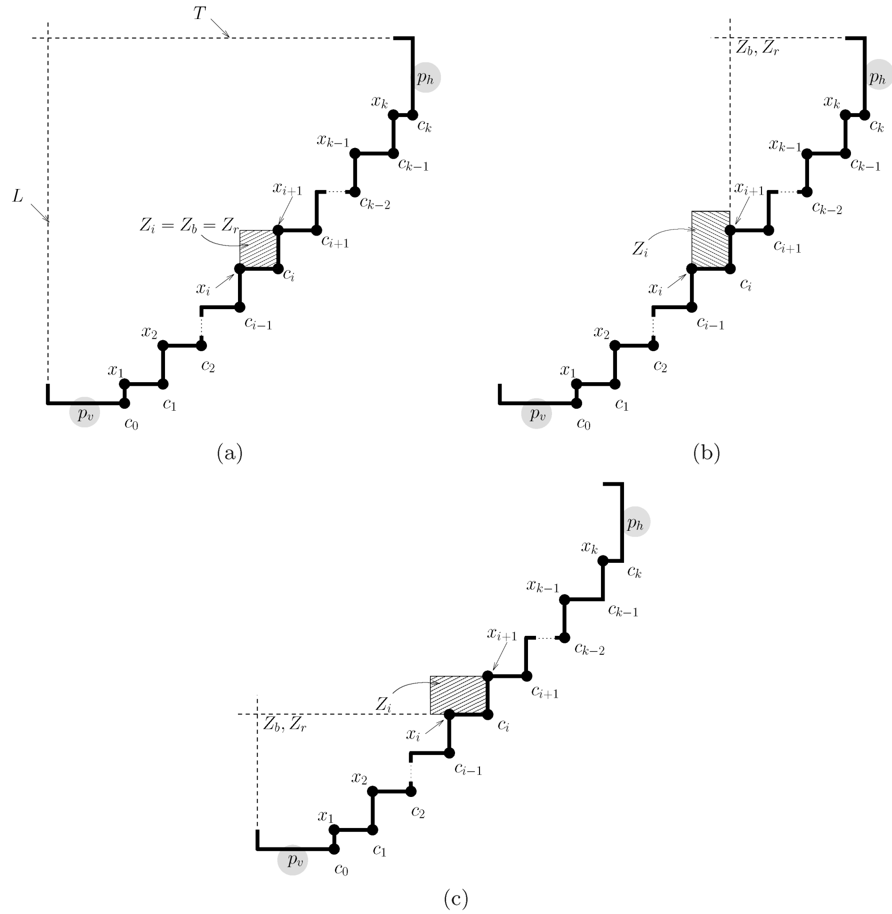


FIG. 11. Bottom-side and right-side contained rectangles with respect to the points $x_i \prec x_{i+1}$, where $x_i, x_{i+1} \in XP$.

$r(p_v, p_h)$. Thus, $Z_{k'}$ is both a right-side and a bottom-side contained rectangle (see Figure 10(b)). Therefore both types of contained rectangles exist in $r(p_v, p_h)$.

Similar arguments can be used for the case when the top-edge of Z_0 is not above line T. □

LEMMA 5.3. *For a pair of adjacent vertical (p_v) and horizontal (p_h) iod-subpaths in canonical form, and two points $x_i \prec x_{i+1}$ along $P'(p_v, p_h)$, where $x_i, x_{i+1} \in XP$, there exist both a bottom-side contained rectangle Z_b and a right-side contained rectangle Z_r either: (a) along the path from p_v to x_{i+1}, or (b) along the path from x_i to p_h. Furthermore, these rectangles have their bottom-right corner point coincide with a point in CP.*

PROOF. Consider the corner point $c_i \in CP$ such that $x_i \prec c_i \prec x_{i+1}$. It is easy to see that if Z_i has its top and left edges bounded by horizontal and vertical lines overlapping the points x_{i+1} and x_i, respectively, then Z_i is vertically and horizontally contained, and it is along both paths from p_v to x_{i+1} and from x_i to p_h (see Figure 11(a)). Assume now that the top edge of rectangle Z_i is above the horizontal line overlapping the point x_{i+1}. It then follows that the right edge of rectangle Z_i and the path from x_{i+1} to p_h produce an up iod-subpath[2] (see Figure 11(b)). Therefore, both horizontal and vertical iod-subpaths exist along the path from x_{i+1} to p_h. By Lemma 5.2 we know that there exist both a bottom-side and a right-side contained rectangles. Furthermore, these rectangles are on the path from x_{i+1} to p_h and consequently on the path from x_i to p_h. Therefore the lemma follows. The remaining case is when the left edge of rectangle Z_i is on the left side of the vertical line overlapping the point x_i (see Figure 11(c)). Similar arguments to the previous case can be used to show that there exist both horizontally and vertically contained rectangles on the path from p_v to x_i and consequently on the path from p_v to x_{i+1}, and the lemma follows. $\square$

Consider the ncpe rectangles n_{i-1}, n_i, and n_{i+1}. By reflexion (with respect to the axis x, and y intersecting the central point of n_i) and rotations (of 90, 180, and 270 degrees with respect to the axis that is perpendicular to the plane passing through the central point of n_i), we know that we only need to consider the case when the exit point X_{i-1} either belongs to P^r_{i-1} or is the BR_{i-1} corner. Therefore, in what follows when we consider a path starting at X_{i-1}, we assume that X_{i-1} is located at $P^r_{i-1} \cup BR_{i-1}$. A path is said to be *type-1* if it contains vertical and horizontal iod-subpaths, and *type-2* otherwise.

To establish our main result in Theorem 5.2 we need to prove Lemma 5.4.

LEMMA 5.4. *For $1 \leq i \leq q$, it is possible to connect at least one of the critical points of every ncpe rectangle $n_i \in VR$ to the corridor, by adding line segments of length at most $l_{i-1} + h_i + l_i$.*

PROOF. By definition of special points, $CD(SpP_1, R) = CD(n_1, R) \leq l_0$, and $CD(SpP_q, R) = CD(n_q, R) \leq l_q$. Therefore, joining SpP_1 (SpP_q) to its nearest point in the corridor requires line segments of length at most l_0 (respectively, l_q).

Now consider each ncpe rectangle n_i for $2 \leq i \leq q-1$. The paths $T(X_{i-1}, Y_i)$ and $T(X_{i-1}, Y_{i+1})$ by our previous definitions are either type-1 or type-2. When $T(X_{i-1}, Y_i)$ ($T(X_{i-1}, Y_{i+1})$) is type-1, we establish in Lemma 5.5 (Lemma 5.6, respectively) shortly that ncpe rectangle n_i can be joined to the corridor by connecting one of its critical points to the nearest point in the corridor by adding line segments of length at most $l_{i-1} + h_i + l_i$. The remaining case is when both $T(X_{i-1}, Y_i)$ and $T(X_{i-1}, Y_{i+1})$ are type-2. In this case we characterize in Lemma 5.7 the path $T(X_{i-1}, Y_j)$ for $j \in \{i, i+1\}$ and then we use this characterization in Lemma 5.8 to prove that rectangle n_i can be joined to the corridor by connecting its critical points to the nearest point in the corridor by adding line segments of length at most $l_{i-1} + h_i + l_i$. The proof of the lemma follows from Lemmas 5.5, 5.6, and 5.8 given next. $\square$

Let us now proceed with Lemmas 5.5, 5.6, and 5.8 needed for the proof of Lemma 5.4.

[2]Note that since Z_i is not an ncpe, the right edge of Z_i and the path from x_{i+1} to p_h is not technically an iod-subpath. However, iod-subpaths may also be defined with cpe rectangles and have the same properties which are needed to prove the lemma.

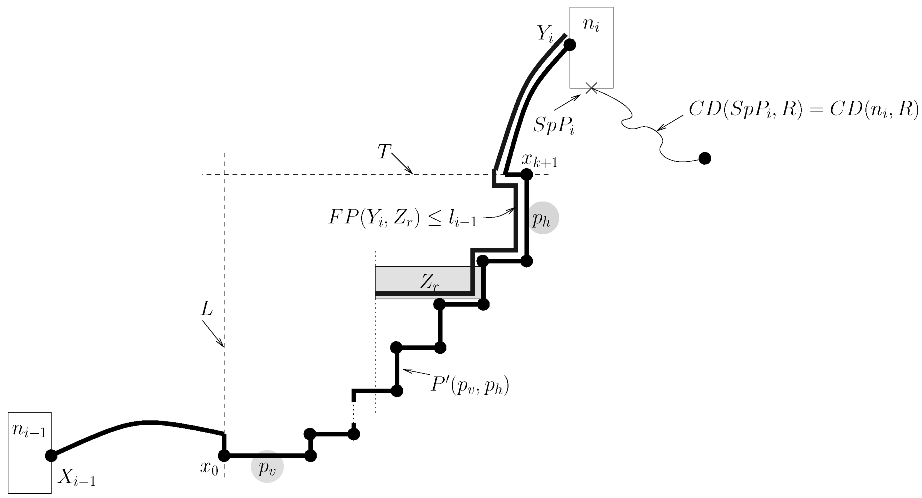


FIG. 12. Type-1 path $T(X_{i-1}, Y_i)$.

LEMMA 5.5. *If path $T(X_{i-1}, Y_i)$ is type-1, then a critical point of n_i can be connected to the corridor by adding line segments of length at most l_{i-1}.*

PROOF. By assumption X_{i-1} is located at $P^r_{i-1} \cup \{BR_{i-1}\}$. Since path $T(X_{i-1}, Y_i)$ is type-1 then by Lemma 5.1 we know it has a pair of adjacent vertical and horizontal iod-subpaths (p_v and p_h) whose beams intersect and are in canonical form. By Lemma 5.2, the region $r(p_v, p_h)$ has a bottom-side contained rectangle $Z_b \in R$ and a right-side contained rectangle $Z_r \in R$ (see Figure 12). We define $FP_c(x_i, Z_j)$ as the length of the path along a restricted set c of edges from x_i to the farthest point in rectangle Z_j. Let c be the restricted set of edges $T(X_{i-1}, Y_i)$ plus the bottom edge of Z_r. By definition $FP(Y_i, Z_r) \leq FP_c(Y_i, Z_r)$. By projecting the bottom edge of Z_r to the corridor $T(X_{i-1}, Y_i)$, we know that $FP_c(Y_i, Z_r) \leq t(X_{i-1}, Y_i)$ and $t(X_{i-1}, Y_i) \leq l_{i-1}$, as we established earlier in previous section. Therefore, $FP(Y_i, Z_r) \leq l_{i-1}$. By the definition of a special point, the special point of n_i can be connected to its closest point of the corridor by segments of length at most $FP(Y_i, Z_r) \leq l_{i-1}$ (see Figure 12). □

LEMMA 5.6. *If path $T(X_{i-1}, Y_{i+1})$ is type-1, then a critical point in n_i can be connected to the corridor by adding line segments of length at most $l_{i-1} + h_i + l_i$.*

PROOF. Let X be a point along the path $T(X_{i-1}, Y_i)$ that is the bifurcation point of the paths $T(X_{i-1}, Y_{i+1})$ and $T(X_{i-1}, Y_i)$. Since $T(X_{i-1}, Y_{i+1})$ is type-1, then applying similar arguments as in the proof of Lemma 5.5, we know that the path $T(X_{i-1}, Y_{i+1})$ has a pair of adjacent vertical and horizontal iod-subpaths (p_v and p_h) whose beams intersect. By Lemma 5.2, the region $r(p_v, p_h)$ has both a bottom-side contained rectangle $Z_b \in R$ and a right-side contained rectangle $Z_r \in R$ (see Figures 13 and 14). If the bifurcation point X is located in the portion of the path $T(X_{i-1}, Y_{i+1})$ that is not included in the path from x_0 to x_{k+1}, then a proof similar to the one for Lemma 5.5 can be used to show that the special point of n_i can be connected to its closest point in the corridor by segments of length at most $FP(X_i, Z_b) \leq l_i$ (Figure 13), or $FP(X_i, Z_r) \leq l_{i-1}$ (Figure 14). Otherwise, X is in

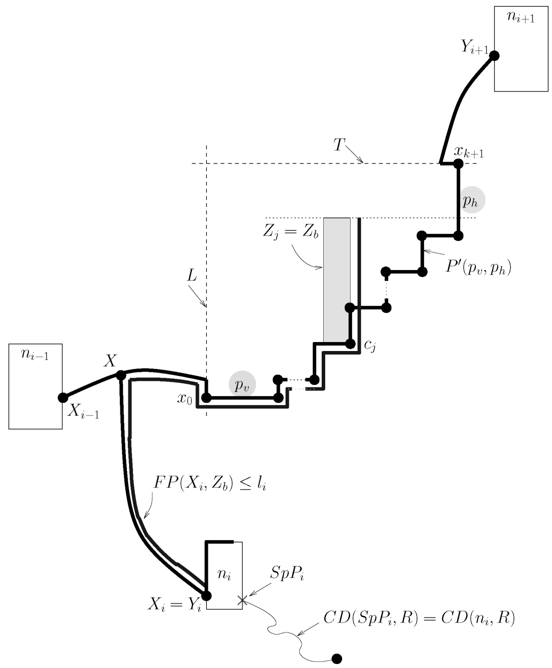


FIG. 13. Type-1 path $T(X_{i-1}, Y_{i+1})$: X is on the left side of subpath $P'(p_v, p_h)$.

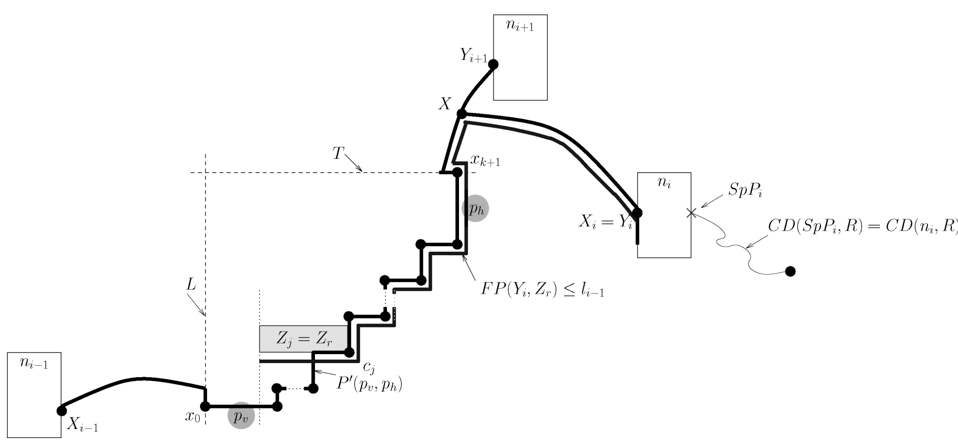


FIG. 14. Type-1 path $T(X_{i-1}, Y_{i+1})$: X is on the right side of subpath $P'(p_v, p_h)$.

the path from x_0 to x_{k+1}. Suppose that the bifurcation point X is in between x_j and x_{j+1} in XP. By Lemma 5.2 we know that there are both a bottom-side contained rectangle $Z_b \in R$ and a right-side contained rectangle $Z_r \in R$ in the path from x_0 to x_{j+1}, or in the path from x_j to x_{k+1}. If there is a bottom-side contained rectangle Z_b (Z_r) whose bottom-side (right-side) is between a pair of corners $c_{j+1}, \ldots, c_k$ ($c_0, \ldots, c_{j-1}$), then arguments similar to the ones for Lemma 5.5 can be used to show that $FP(X_i, Z_b) \leq l_i$ ($FP(X_i, Z_r) \leq l_{i-1}$), and therefore the special point of n_i can be connected to its closest point in the corridor by a set of line segments of

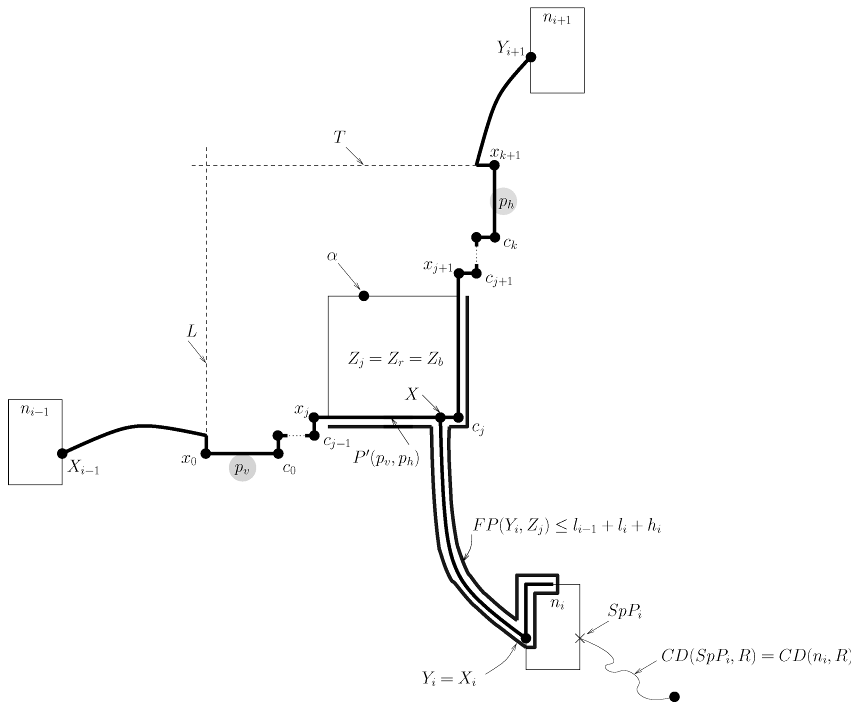


FIG. 15. Type-1 path $T(X_{i-1}, Y_{i+1})$: X is along the subpath $P'(p_v, p_h)$.

length at most $FP(X_i, Z_b) \leq l_i$ ($FP(X_i, Z_r) \leq l_{i-1}$). One can show that the only remaining case is when the rectangle Z_j (the one with bottom-right corner at c_j) is both a bottom-side and right-side contained rectangle. In this case we can easily show that $FP(X_i, \alpha)$ is at most $l_i + h_i + l_{i+1}$, where α is any point (as indicated in Figure 15) along the edges of Z_j where the corridor reaches it. Therefore the special point of n_i can be connected to its closest point in the corridor by a set of line segments of length at most $FP(X_i, Z_j) \leq l_{i-1} + h_i + l_i$ (see Figure 15). This concludes the proof of the lemma. □

Consider the case when both $T(X_{i-1}, Y_i)$ and $T(X_{i-1}, Y_{i+1})$ are type-2 paths. Before we prove Lemma 5.8, we need to characterize type-2 paths and establish some important properties. Since both $T(X_{i-1}, Y_i)$ and $T(X_{i-1}, Y_{i+1})$ are type-2 paths, then path $T(X_{i-1}, Y_j)$, for $j \in \{i, i+1\}$, does not have both a vertical and a horizontal iod-subpath. A path is said to be type-V, type-H, or type-N when it has only vertical iod-subpaths, only horizontal iod-subpaths, or no iod-subpaths, respectively. A type-V path is said to be type-VMR if traversing each segment of the path from X_{i-1} to Y_j takes us up, down, or right. A type-H path is said to be type-HMD if traversing each segment of the path from X_{i-1} to Y_j takes us left, right, or down. A type-N path is said to be type-ND if traversing each segment of the path from X_{i-1} to Y_j takes us right or down, and $X_{i-1} \neq Y_j$; and type-ND$_1$, if $X_{i-1} = Y_j$.

LEMMA 5.7. *A type-2 path $T(X_{i-1}, Y_j)$, for $j \in \{i, i+1\}$, is of one of the following forms:*

(1) *Path $T(X_{i-1}, Y_j)$ is type-VMR, $X_{i-1} \in P^r_{i-1} \cup \{BR_{i-1}\}$ and the entry point $Y_j \in P^l_j \cup \{TL_j\}$.*
(2) *Path $T(X_{i-1}, Y_j)$ is type-HMD, $X_{i-1} = BR_{i-1}$, and the entry point $Y_j \in P^t_j \cup \{TL_j\}$.*
(3) *Path $T(X_{i-1}, Y_j)$ is type-ND, $X_{i-1} = BR_{i-1}$ and the entry point $Y_j = TL_j$.*
(4) *Path $T(X_{i-1}, Y_j)$ is type-ND$_1$, consists of only one point, $X_{i-1} = Y_j$ and $X_{i-1} \in P^r_{i-1} \cup \{BR_{i-1}\}$, and $Y_j \in P^l_j \cup \{TL_j\}$.*

PROOF. By assumption X_{i-1} is located at $P^r_{i-1} \cup \{BR_{i-1}\}$. There are four cases depending on the type of the path $T(X_{i-1}, Y_j)$.

Case 1. Path $T(X_{i-1}, Y_j)$ is type-V.
By definition the path has vertical iod-subpaths. While traversing the path from X_{i-1} to Y_j, if a horizontal segment is traversed in the left direction, then there will be a horizontal iod-subpath contradicting that $T(X_{i-1}, Y_j)$ is a type-2 path. So as we traverse $T(X_{i-1}, Y_j)$ from X_{i-1}, horizontal segments are traversed from left to right and the vertical segments are traversed in either direction. So the alternatives for Y_j are from $P^l_j \cup P^t_j \cup P^b_j \cup \{TL_j, BR_j\}$. However, if $Y_j \in P^t_j \cup P^b_j \cup \{BR_j\}$, then there is a horizontal iod-subpath, contradicting that the path $T(X_{i-1}, Y_j)$ is type-2. It must then be that $Y_j \in P^l_j \cup \{TL_j\}$. Therefore, the path $T(X_{i-1}, Y_j)$ is type-VMR (see Figure 16(a)[3]).

Case 2. Path $T(X_{i-1}, Y_j)$ is type-H.
Since the path has at least one horizontal iod-subpath and the path is type-2, it must be that $X_{i-1} = BR_{i-1}$ and as we traverse the path $T(X_{i-1}, Y_j)$ from X_{i-1} to Y_j, all its vertical segments are traversed in the downward direction. The horizontal segments of $T(X_{i-1}, Y_j)$ are traversed in either direction. So the alternatives are that Y_j is located in $P^t_j \cup P^l_j \cup P^r_j \cup \{TL_j, BR_j\}$. However, if $Y_j \in P^l_j \cup P^r_j \cup \{BR_j\}$, then there is also a vertical iod-subpath, contradicting that $T(X_{i-1}, Y_j)$ is type-2. It must then be that $Y_j \in P^t_j \cup \{TL_j\}$, and the path $T(X_{i-1}, Y_j)$ is type-HMD (see Figure 16(b)).

Case 3. Path $T(X_{i-1}, Y_j)$ is type-N and $X_{i-1} \neq Y_j$.
Since the path has no iod-subpaths, it must be that $X_{i-1} = BR_{i-1}$, and as we traverse the path $T(X_{i-1}, Y_j)$ from X_{i-1} to Y_j, all its vertical segments are traversed in the downward direction. It cannot be that a horizontal line segment of $T(X_{i-1}, Y_j)$ is traversed from right to left because that would mean there is a horizontal iod-subpath, contradicting that $T(X_{i-1}, Y_j)$ is a type-N path. Thus, as we traverse the path $T(X_{i-1}, Y_j)$ from X_{i-1} to Y_j, the segments are traversed in the downward and rightward direction. If $Y_j \in P^t_j \cup P^l_j$, a horizontal or vertical iod-subpath is formed, contradicting that $T(X_{i-1}, Y_j)$ is type-N. It must then be that $Y_j = TL_j$. Therefore, the path $T(X_{i-1}, Y_j)$ is type-ND (see Figure 16(c)).

[3] Thick edges and filled circles in the corners of ncpes indicate valid connection points between the corridor $T(X_{i-1}, Y_j)$ and the ncpe rectangles. Unfilled circles in the corners of ncpes indicate invalid connection points between the corridor $T(X_{i-1}, Y_j)$ and ncpes (because of the definition of ncpe rectangles).

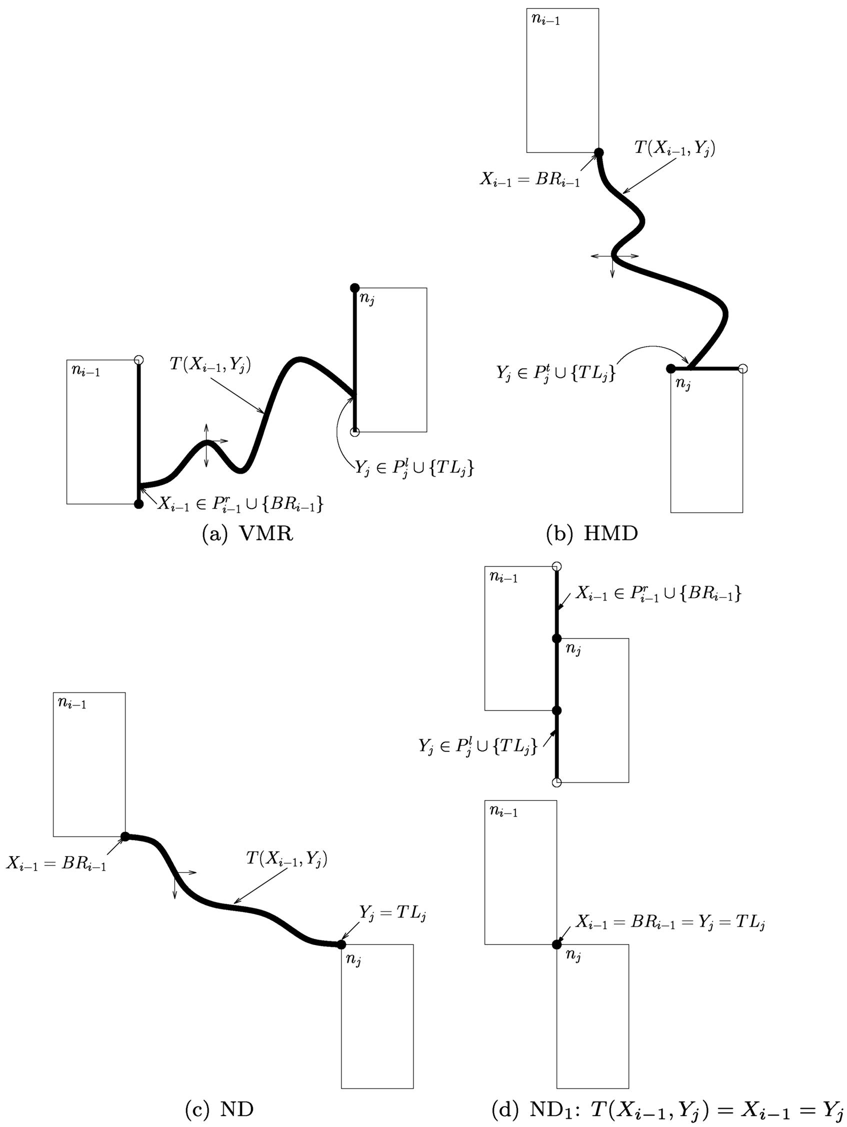


FIG. 16. Possible path types for $T(X_{i-1}, Y_j)$.

Case 4. Path $T(X_{i-1}, Y_j)$ is type-N and $X_{i-1} = Y_j$.

If $X_{i-1} \in P^r_{i-1} \cup \{BR_{i-1}\}$ then $Y_j \in P^l_j \cup \{TL_j\}$, and if $X_{i-1} = BR_{i-1}$ then $Y_j = TL_j$. So the path is type-ND_1 (see Figure 16(d)). □

LEMMA 5.8. *If paths $T(X_{i-1}, Y_i)$ and $T(X_{i-1}, Y_{i+1})$ are type-2, then a critical point of n_i can be connected to the corridor by adding line segments of length at most $l_{i-1} + h_i + l_i$.*

PROOF. By assumption X_{i-1} is located at $P^r_{i-1} \cup \{BR_{i-1}\}$. Since paths $T(X_{i-1}, Y_i)$ and $T(X_{i-1}, Y_{i+1})$ are type-2 then by Lemma 5.7 they can only be type VMR, HMD, ND, or ND_1. On the path $T(X_{i-1}, Y_{i+1})$, let X be the

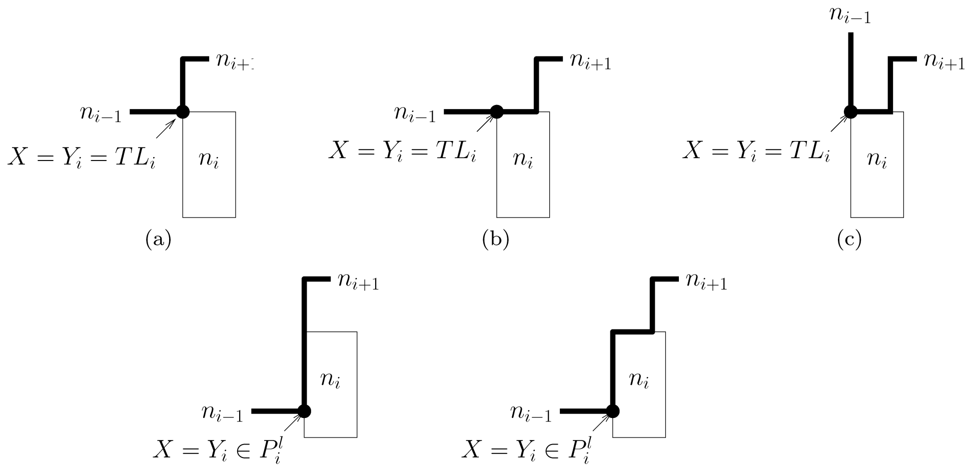


FIG. 17. Connection of ncpe rectangle n_i along the path $T(X_{i-1}, Y_{i+1})$ of type-VMR: (a)-(c) $X = Y_i = TL_i$, (d)-(e) $X = Y_i \in P_i^l$.

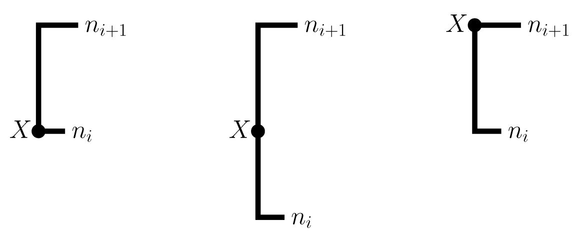


FIG. 18. Bifurcation point X of paths $T(X_{i-1}, Y_{i+1})$ and $T(X_{i-1}, Y_i)$, both of type-VMR.

bifurcation point where the paths $T(X_{i-1}, Y_i)$ and $T(X_{i-1}, Y_{i+1})$ break into two (see Figures 13, 14, and 15), or where $T(X_{i-1}, Y_i)$ ends (i.e., $X = Y_i$) or where $T(X_{i-1}, Y_{i+1})$ ends (i.e., $X = Y_{i+1}$). We will show that the path $T(Y_i, Y_{i+1})$ is type-1. There are four cases depending on the type of the path $T(X_{i-1}, X)$.

Case 1. Path $T(X_{i-1}, X)$ is type-VMR. Clearly, both $T(X_{i-1}, Y_i)$ and $T(X_{i-1}, Y_{i+1})$ are also type-VMR paths. Therefore, the paths $T(X, Y_i)$ and $T(X, Y_{i+1})$ are type-VMR, type-ND, or type-ND$_1$. There are three cases depending on X: (a) $X = Y_i$, (b) $X = Y_{i+1}$, and (c) $X \neq Y_i$ and $X \neq Y_{i+1}$. For case (a) ($X = Y_i$) when $Y_i = TL_i$ the first two[4] (or three) segments of $T(X, Y_{i+1})$ form a horizontal iod-subpath with P_i^t (see Figure 17(a)–(c)). When $Y_i \in P_i^l$ the first two (or four) segments of $T(X, Y_{i+1})$ form a horizontal iod-subpath with P_i^t (see Figure 17(d)–(e)). Applying similar arguments for case (b) ($X = Y_{i+1}$), we know that $T(X, Y_i)$ has a horizontal iod-subpath. For case (c) (when $X \neq Y_i$ and $X \neq Y_{i+1}$), the first two segments of $T(X, Y_i)$ and the first two segments of $T(X, Y_{i+1})$ form a horizontal iod-subpath in $T(Y_i, Y_{i+1})$ (see the three cases in Figure 18). Thus, if $T(X, Y_i)$ or the $T(X, Y_{i+1})$ is a type-VMR path and there exists

[4] In the proof of this lemma we state that there are several segments in a path. We do not give a proof of this because it is straightforward when we use the edge of the ncpe rectangle at the end of the path.

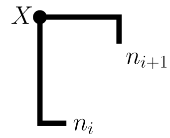


FIG. 19. Bifurcation point X of paths $T(X_{i-1}, Y_{i+1})$ and $T(X_{i-1}, Y_i)$, both of type-ND.

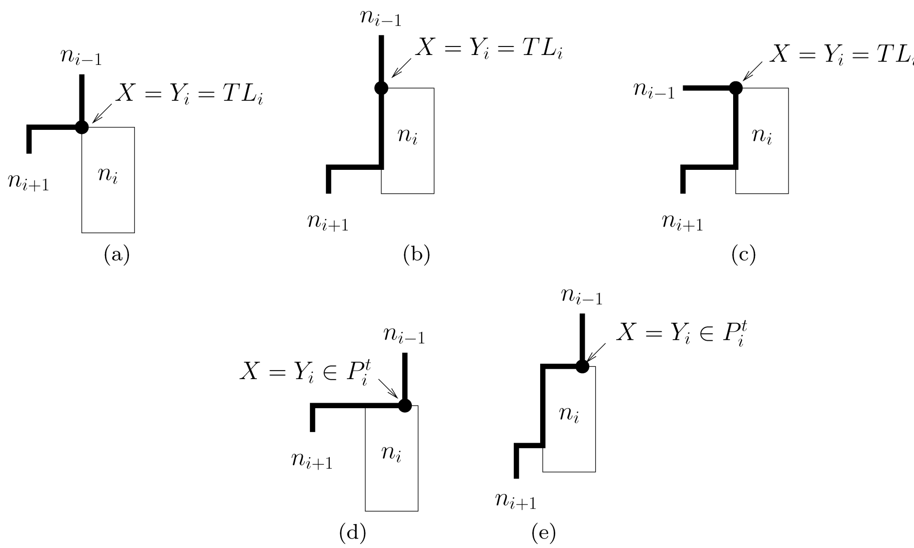


FIG. 20. Connection of ncpe rectangle n_i along the path $T(X_{i-1}, Y_{i+1})$ of type-HMD: (a)-(c) $X = Y_i = TL_i$, (d)-(e) $X = Y_i \in P_i^t$.

a horizontal iod-subpath in $T(Y_i, Y_{i+1})$, then the path $T(Y_i, Y_{i+1})$ is type-1. On the other hand, since $X \neq Y_i$ and $X \neq Y_{i+1}$, it must be that both $T(X, Y_i)$ and $T(X, Y_{i+1})$ are type-ND. The two segments in the path $T(X, Y_i)$ and the first two segments in the path $T(X, Y_{i+1})$ form a horizontal and vertical iod-subpaths (see Figure 19). Therefore path $T(Y_i, Y_{i+1})$ is type-1.

Case 2. Path $T(X_{i-1}, X)$ is type-HMD. Clearly, both $T(X_{i-1}, Y_i)$ and $T(X_{i-1}, Y_{i+1})$ are type-HMD. Therefore, the paths $T(X, Y_i)$ and $T(X, Y_{i+1})$ are type-HMD, type-ND, or type-ND$_1$. There are three cases depending on X: (a) $X = Y_i$, (b) $X = Y_{i+1}$, and (c) $X \neq Y_i$ and $X \neq Y_{i+1}$. For case (a) ($X = Y_i$) when $Y_i = TL_i$ the first two or three segments of $T(X, Y_{i+1})$ form a vertical iod-subpath with P_i^l (see Figure 20(a)–(c)). When $Y_i \in P_i^t$, the first three or four segments of $T(X, Y_{i+1})$ form also a vertical iod-subpath with P_i^l (see Figure 20(d)–(e)). Applying similar arguments for case (b) ($X = Y_{i+1}$), we know that $T(X, Y_i)$ has a vertical iod-subpath. For case (c) (when $X \neq Y_i$ and $X \neq Y_{i+1}$), the first two segments of $T(X, Y_i)$ and the first two segments of $T(X, Y_{i+1})$ form a vertical iod-subpath in $T(Y_i, Y_{i+1})$ (see the three cases in Figure 21). Thus, if $T(X, Y_i)$ or $T(X, Y_{i+1})$ is a path of type-HMD and there exists a vertical iod-subpath in $T(Y_i, Y_{i+1})$, then the path $T(Y_i, Y_{i+1})$ is type-1. On the other hand, since $X \neq Y_i$ and $X \neq Y_{i+1}$ it must be that both $T(X, Y_i)$ and $T(X, Y_{i+1})$ are type-ND. The first two segments in path $T(X, Y_i)$ and the first two segments in path $T(X, Y_{i+1})$ form a vertical and horizontal iod-subpaths (see

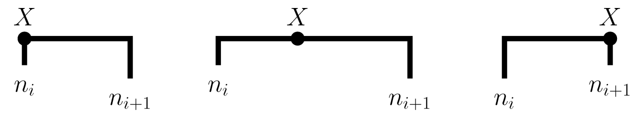


FIG. 21. Bifurcation point X of paths $T(X_{i-1}, Y_{i+1})$ and $T(X_{i-1}, Y_i)$, both of type-HMD.

Figure 19). Thus $T(Y_i, Y_{i+1})$ has a vertical and horizontal iod-subpaths and therefore it is type-1.

Case 3. Path $T(X_{i-1}, X)$ is type-ND. Then the paths $T(X_{i-1}, Y_i)$ and $T(X_{i-1}, Y_{i+1})$ are types VMR, HMD, or ND. Therefore, paths $T(X, Y_i)$ and $T(X, Y_{i+1})$ are types VMR, HMD, ND, or ND_1. If one of these paths is type-VMR and the other is type-HMD, then clearly, $T(Y_i, Y_{i+1})$ has a vertical and horizontal iod-subpaths and it is type-1. So both paths $T(X, Y_i)$ and $T(X, Y_{i+1})$ are VMR, ND, or one is ND_1; or they both are HMD, ND, or one is ND_1. In either case, arguments similar to the ones for Cases 1 and 2 can be used to establish that the path $T(Y_i, Y_{i+1})$ is type-1.

Case 4. Path $T(X_{i-1}, X)$ is type-ND_1. Since n_i is visited after n_{i-1} but before n_{i+1}, one can show that $X_{i-1} = X = Y_i$. Since these two ncpe rectangles share a point, it must be that at least one of the rectangles has a corner point at point X. But this corner point cannot be the *TR* or *BL* corner as otherwise n_{i-1} or n_i would not be an ncpe rectangle. So it has to be that one corner point is a *TL* corner or it is a *BR* corner. Since $X_{i-1} \in P^r_{i-1} \cup \{BR_{i-1}\}$, then $X_i \in P^t_i$ or $X_i = TL_i$. In either case, it must be that the path $T(X, Y_{i+1})$ is type-VMR, containing a vertical iod-subpath. But a horizontal iod-subpath is formed by the top-edge of n_i. Therefore path $T(Y_i, Y_{i+1})$ is type-1.

In all the cases, there exist both horizontal and vertical iod-subpaths along the path $T(Y_i, Y_{i+1})$. Now, by Lemmas 5.1, 5.2, and 5.3 we know that there exist both bottom-side and right-side contained rectangles along $T(Y_i, Y_{i+1})$. Using arguments similar to those of Lemma 5.5 and the fact that $T(Y_i, Y_{i+1})$ is type-1, one can prove that a critical point of n_i can be connected to the corridor by adding line segments of length at most $l_{i-1} + h_i + l_i$. This completes the proof of the lemma. □

5.3. SELECTING THE FOUR CORNERS AND ONE SPECIAL POINT. The critical points in $S(4C+)$ for each rectangle $R_i \in R$ are its four corners and a special point. For this case $k_{S(4C+)} = 5$ and we can show that $r_{S(4C+)} = 3$. Therefore, the approximation ratio of the parametrized algorithm is 30 as in the case of $S(2OC+)$.

We now briefly discuss our proof strategy to show that given any corridor $T(I)$ there is a corridor $T(I_{S(4C+)})$ such that $t(I_{S(4C+)}) \leq 3 \cdot t(I)$. As in the case of $S(2OC+)$, given any corridor $T(I)$ we identify all the ncpe rectangles and establish an ordering $(n_1, n_2, \ldots, n_q)$ between them. Assume there exists at least one ncpe rectangle ($q \geq 1$), otherwise $t(I_{S(4C+)}) = t(I)$ and the result follows. For each ncpe we find a shortest path from one of its critical points to the corridor $T(I)$. We then select this path to connect a critical point to the corridor $T(I)$. Clearly after deleting some edges to remove any cycle that may have been created, corridor $T(I)$ plus a subset of these connections give a corridor $T(I_{S(4C+)})$. Next we need to show that the sum of the length of the segments introduced is at most $2 \cdot t(I)$. This is the part that is more complex than the one for the selector function $S(2OC+)$.

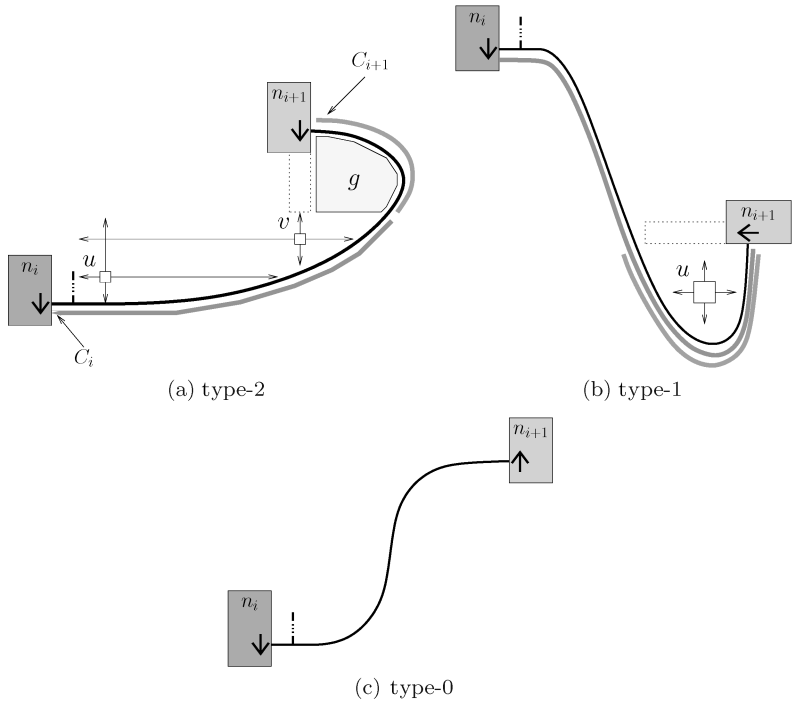


FIG. 22. Region types.

We characterize the region between every pair of adjacent ncpe rectangles. The region between two adjacent ncpe rectangles is said to be of type 0, 1 or 2 (see Figure 22). The region between ncpe rectangles n_i and n_{i+1} is type-2 (see Figure 22(a)), if the distance along the corridor between n_i and n_{i+1}, which we call in Section 5.1 l_i, is larger than the edge-length needed to connect both a critical point from n_i and one from n_{i+1} to the corridor. This is the most desirable case. If this were to be the case for every pair of adjacent ncpe rectangles the proof of the approximation ratio would be simple to establish (in fact we would even be able to establish a better ratio). However, this is not always the case. The region between ncpe n_i and n_{i+1} is type-1 (see Figure 22(b)), if l_i is larger than the edge-length needed to connect either a critical point from n_i to the corridor, or one from n_{i+1} to the corridor, but not both. If this were the case for every pair of adjacent ncpe rectangles the proof would also be simple. The main problem is when the region between ncpe n_i and n_{i+1} is type-0 (see Figure 22(c)). In this case, the edge-length needed to connect a critical point of either n_i or n_{i+1} cannot be bounded above by l_i. This is where the proof is complex because we need to consider a sequence of ncpe rectangles, not just three as in the previous subsection.

Now suppose that there is a sequence of type-0 adjacent ncpe rectangles as shown in Figure 23. The connection of the first ncpe rectangle n_1 to the corridor has already been accounted for in l_0. Now we need to charge the connection of a critical point of each ncpe rectangles $n_2, \ldots, n_9$ to the corridor. The connection for the ncpe n_2 is charged to the horizontal distance from ncpe n_1 to ncpe n_2, and the vertical distance from ncpe n_2 to ncpe n_3 because one can show that the area

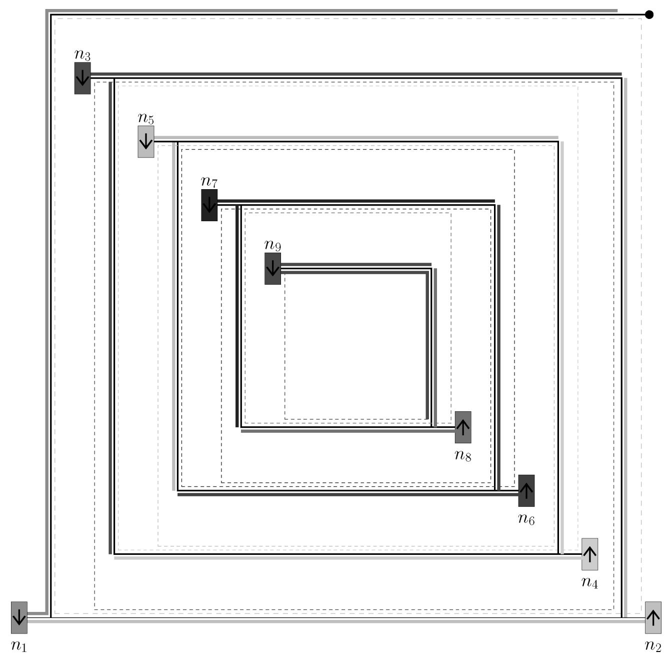


FIG. 23. Sequence of adjacent regions type-0.

includes at least one rectangle already connected to the corridor. Similarly, the cost of connecting ncpe n_3 can be charged to the horizontal distance from ncpe n_2 to ncpe n_3 and the vertical distance from ncpe n_3 to ncpe n_4. And so forth until ncpe n_9, where its connection is charged to the corridor after it because there is a rectangle inside the box in the center of Figure 23.

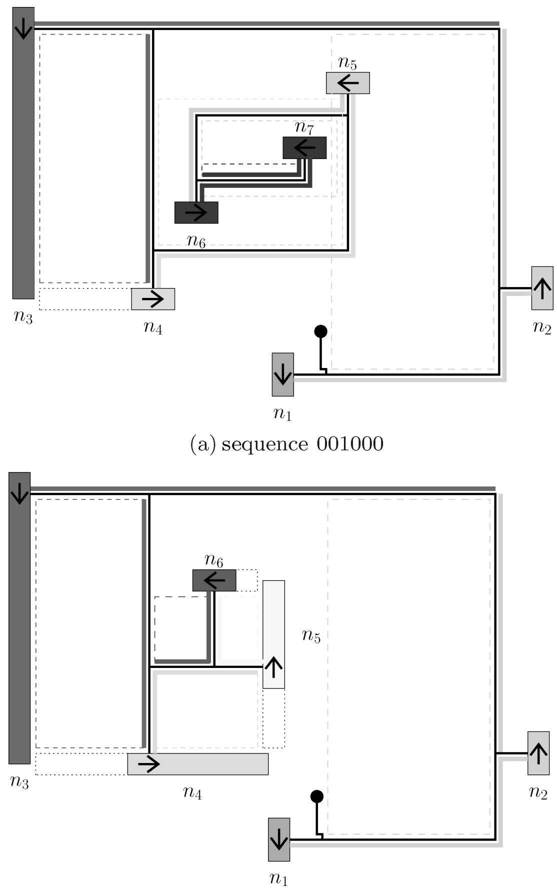
Other complex cases are given in Figure 24(a)–(b) which indicate how to deal with the sequence of adjacent rectangles of types 001000 and sequence 00111. By the sequence $Y_1 Y_2 \dots Y_k$ we mean that the first pair of ncpe rectangles is type Y_1 and the second pair is type Y_2, and so on. There are more critical cases that need to be considered. It is possible to characterize all sequences that need to be solved, but it is quite complex. That is why we only present the analysis for the selector function $S(2OC+)$, which is significantly simpler.

We claim without stating any further details that the approximation ratio of the parameterized algorithm Alg($S(4C+)$) is 30.

6. *Additional Results and Discussion*

Our approximation algorithm is based on restriction and relaxation techniques. The analysis of our approximation algorithm applies (with the same time complexity and approximation ratio) when the boundary of the MLC-R problem is a rectilinear polygon rather than the rectangle F, or when the problem is to find a tree that is not necessary joined to the boundary of F.

Our approximation algorithm can also be adapted to the MLC$_k$ problem, but the approximation ratio depends on k. When k is bounded above by a constant, the algorithm is a constant ratio approximation algorithm. The selector function $S(C+)$, which includes all the corner points plus other points (defined later on),



(a) sequence 001000

(b) sequence 00111

FIG. 24. Assignment of regions between adjacent ncpe rectangles.

and a special point of each rectilinear polygon $R_i \in R$, is more complex. The idea is to introduce for each rectilinear polygon R_i the least number of horizontal line segments (all of which are completely inside R_i) that partition the interior of R_i into rectangles (see Figure 25). All the corners of these rectangles are on the boundary of R_i and are the fixed points for R_i. Then we add a special point for R_i. The total number of critical points for each R_i is at most $\frac{3}{2}k - 1$. Therefore, $k_{S(C+)} = \frac{3}{2}k - 1$.

We now need to determine $r_{S(C+)}$. The process follows the same lines as the one for the MLC-R problem. The only difference is when considering the ncpes. Instead of selecting for each ncpe n_i the rectilinear polygon, we just take the rectangle that the tour intersects first. So the set of ncpes are simply rectangles with four fixed points. The special point is for the whole rectilinear polygon. We claim that our analysis for the MLC-R problem also applies for this case. Nonetheless, it can be simplified since each rectangle includes at least four critical points and we are just interested in establishing a constant approximation ratio (rather than a minimal one). Note that the special point is part of only one rectangle. But since it is associated with the rectilinear polygon, that is enough to carry through

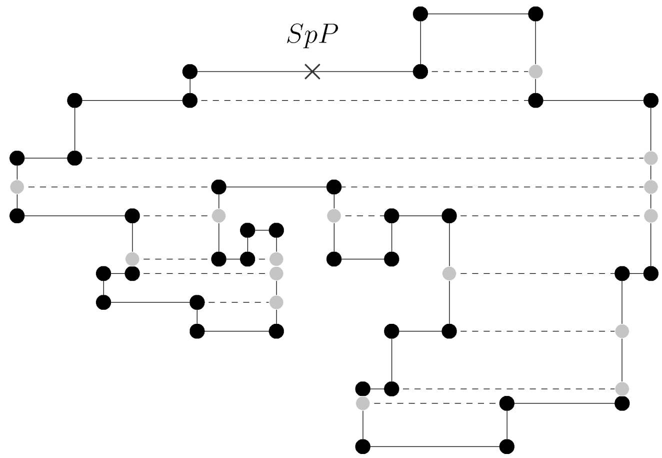


FIG. 25. Partition of polygon R_i into rectangles by introducing horizontal line segments completely inside R_i.

our analysis and show that $r_{S(C+)} = 5$. The approximation bound is therefore $2k_{S(C+)} \cdot r_{S(C+)} = 2 \cdot (\frac{3}{2}k - 1) \cdot 5 = 15k - 10$. For k bounded above by a constant, the approximation ratio is a constant. Note that with a more careful introduction of critical points we can decrease the approximation ratio. For example, we just need the bottom-left and top-right corner of each rectangle. But at this point we are only interested in showing that our approximation algorithm takes polynomial time and it is a constant ratio approximation algorithm for the MLC$_k$ problem.

We have adapted our algorithm (with the same constant ratio) to restricted versions of the MLC problem, but so far we have not been able to adapt it to all cases. The existence of a constant ratio approximation algorithm for the MLC problem remains a challenging open problem. An equally challenging problem is to develop approximation algorithm for the MLC-R problem with a significant smaller approximation ratio, for example two.

An interesting open problem is to develop a constant ratio approximation algorithm for the version of the MLC-R problem when not all the rectangles in R need to access the corridor. This corresponds to the "Steiner" version of the problem, rather than the "spanning tree" version of the problem which we call the MLC-R problem. Our analysis does not seem to apply for this case.

When we restrict the MLC-R problem to $S(2OC+)$ or $S(4C+)$, we use Slavik's algorithm for the TEC problem to generate a suboptimal solution for the MLC-R problem instance. This is the most time-consuming part of our procedures. The first open question is about the development of a faster approximation algorithm for the MLC-R problem restricted by $S(2OC+)$ or $S(4C+)$. The second open question has to do with the development of an algorithm for those problems with a smaller approximation ratio or even one that generates an optimal solution. But since the MLC-R problem is NP-hard when restricted by $S(2OC+)$ or $S(4C+)$, it is unlikely one can find an efficient algorithm for its solution. The NP-hardness proof follows the same lines as the one in Gonzalez-Gutierrez and Gonzalez [2007b], but we need to do some modifications to show that the the MLC-R restricted to $S(2OC+)$ ($S(4C+)$) (i.e., every rectangle must intersect the corridor at a critical point defined by $S(2OC+)$ (respectively, $S(4C+)$)) is NP-hard.

The MLC problem (and all its subproblems) can be easily solved in polynomial time when the objective function is to find a tree such that the maximum length of the path from the root to a leaf is least possible. A class of interesting open problems are ones with dual-criteria objective functions, that is, minimize the maximum path length and minimize the total corridor edge-length. Our algorithm and analysis also hold for the case when each edge in the graph $G(V, E, w)$ constructed from the instance (F, R, p) of p-MLC-R problem has an arbitrary weight function rather than the weight being the distance between the two vertices. For example, the weight of an edge might be zero so in our application for fiber optics means that a fiber already exists along the edge. One may also add edges to the graph (constructed from the instance (F, R, p)) with arbitrary weights that represent points connected directly by existing fiber in our application. Our results will also carry to this version of the problem.

A related problem studied by Slavik [1997, 1998] is the *Errand Scheduling* (ES) problem. In this case the problem is to find a *shortest partial tour* visiting a subset of vertices of the given metric graph G such that at least one vertex in $C_i \subset C$ is in the partial tour, where C_i is associated with the *errand* i. When each vertex represents a unique errand, the ES problem is an instance of the well-known *Traveling Salesperson Problem (TSP)*. Therefore the ES problem is NP-hard. The ES problem has also been referred to as the *group TSP (g-TSP)*. Slavik [1997, 1998] shows that the ES problem restricted to metric graphs problem can be approximated to within $\frac{3\rho}{2}$ when each errand can be performed in at most ρ nodes. Another interesting problem is the group-TSP problem when restricted to rectangles as in the case of the of the MLC-R problem, which we call the *rectangular group-TSP*. In this version of the TSP one may visit the same edge or vertex more than once.

We claim that the same approach that we use for the MLC-R problem can also be adapted to the rectangular group-TSP. It is simple to show that the selector functions that do not generate constant ratio approximations for the MLC-R problem do not generate constant ratio approximations for the rectangular group-TSP. However the parameterized algorithms using the selector function $S(2OC+)$ also generates a constant ratio approximation to the group-TSP. In fact we can just use the tour of the corridor (traversing each edge twice) as the solution to the rectangular TSP problem. The approximation ratio in this case will be 60 times the length of an optimal tour. However there is a better algorithm for this case. Instead of using Slavik's approximation algorithm for the TEC problem, we use the one for the ES problem [Slavik 1997, 1998] in the parameterized algorithm. This results in an algorithm with approximation ratio $\frac{3}{2} \cdot k_{S(2OC+)} \cdot r_{S(2OC+)}$, where $k_{S(2OC+)} = 3$ and $r_{S(2OC+)} = 5$, which is 22.5 for the rectangular group-TSP.

The same type of approach used for the MLC_k can be used for the rectilinear c-gon group-TSP ($c \leq k$), resulting in a polynomial-time constant ratio approximation. For brevity we do not discuss additional results. Approximation algorithms for other versions of the group-TSP problem are discussed in Bodlaender et al. [2006].

REFERENCES

Bateman, C. D., Helvig, C. S., Robins, G., and Zelikovsky, A. 1997. Provably good routing tree construction with multiport terminals. In *Proceedings of the International Symposium on Physical Design (ISPD97)*. ACM Press, 96–102.

BODLAENDER, H. L., FEREMANS, C., GRIGORIEV, A., PENNINKX, E., SITTERS, R., AND WOLLE, T. 2006. On the minimum corridor connection and other generalized geometric problems. In *Proceeding of the 4th Workshop on Approximation and Online Algorithms (WAOA)*. T. Erlebach and C. Kaklamanis, Eds. Lecture Notes in Computer Science, vol. 4368. Springer, 69–82.

DEMAINE, E. D. AND O'ROURKE, J. 2001. Open problems from CCCG 2000. In *Proceedings of the 13th Canadian Conference on Computational Geometry (CCCG01)*. 185–187.

EPPSTEIN, D. 2001. Some open problems in graph theory and computational geometry. PDF file (3.89 MB), World Wide Web. http://www.ics.uci.edu/∼eppstein/200-f01.pdf.

GONZALEZ-GUTIERREZ, A. AND GONZALEZ, T. F. 2007a. Approximation algorithms for the minimum-length corridor and related problems. Tech. rep. TR 2007-03, Department of Computer Science, University of California at Santa Barbara. http://www.cs.ucsb.edu/research/tech_reports/.

GONZALEZ-GUTIERREZ, A. AND GONZALEZ, T. F. 2007b. Complexity of the minimum-length corridor problem. *Comput. Geom. Theory Appl. 37,* 2, 72–103.

HELVIG, C. S., ROBINS, G., AND ZELIKOVSKY, A. 2001. An improved approximation scheme for the group Steiner problem. *Netw. 37,* 1, 8–20.

IHLER, E. 1991. Bounds on the quality of approximate solutions to the group Steiner problem. In *Proceedings of the 16th International Workshop on Graph-Theoretic Concepts in Computer Science (WG90)*. Springer, 109–118.

JIN, L. Y. AND CHONG, O. W. 2003. The minimum touching tree problem. PDF file (189 KB), World Wide Web. http://www.yewjin.com/research/MinimumTouchingTrees.pdf.

KARP, R. M. 1972. Reducibility among combinatorial problems. *Complex. Comput. Comput.*, 85–103.

LI, F., CHENG, D., HADJIELEFTHERIOU, M., KOLLIOS, G., AND TENG, S.-H. 2005. On trip planning queries in spatial databases. In *Proceedings of Advances in Spatial and Temporal Databases, 9th International Symposium (SSTD05)*. C. B. Medeiros, M. J. Egenhofer, and E. Bertino, Eds. Lecture Notes in Computer Science, vol. 3633. Springer, 273–290.

MITCHELL, J. S. B. 2000. Geometric shortest paths and network optimization. In *Handbook of Computational Geometry*. Elsevier, North-Holland, Amsterdam, 633–701.

REICH, G. AND WIDMAYER, P. 1990. Beyond Steiner's problem: A VLSI oriented generalization. In *Proceedings of the 15th International Workshop on Graph-Theoretic Concepts in Computer Science (WG89)*. Springer, 196–210.

SAFRA, S. AND SCHWARTZ, O. 2006. On the complexity of approximating TSP with neighborhoods and related problems. *Comput. Complex. 14,* 4, 281–307.

SLAVIK, P. 1997. The errand scheduling problem. Tech. rep. 97-02, Department of Computer Science and Engineering, University of New York at Buffalo. http://www.cse.buffalo.edu/tech-reports/.

SLAVIK, P. 1998. Approximation algorithms for set cover and related problems. Ph. D. thesis 98-06, Department of Computer Science and Engineering, University of New York at Buffalo. http://www.cse.buffalo.edu/tech-reports/.

RECEIVED AUGUST 2007; REVISED AUGUST 2008; ACCEPTED AUGUST 2008