University of Warsaw

Faculty of Mathematics, Informatics and Mechanics

Katarzyna Kowalska

Student no. 371053

Approximation and Parametrized Algorithms for Segment Set Cover

Master's thesis in COMPUTER SCIENCE

Supervisor: dr Michał Pilipczuk Instytut Informatyki

10 Supervisor's statement

- Hereby I confirm that the presented thesis was prepared under my supervision and that it fulfils the requirements for the degree of Master of Computer Science.
- Date Supervisor's signature

$_{14}$ Author's statement

Hereby I declare that the presented thesis was prepared by me and none of its contents was obtained by means that are against the law.

The thesis has never before been a subject of any procedure of obtaining an academic degree.

Moreover, I declare that the present version of the thesis is identical to the attached electronic version.

21 Date Author's signature

22	${f Abstract}$
23 24	The work presents a study of different geometric set cover problems. It mostly focuses on segment set cover and its connection to the polygon set cover.
25	${f Keywords}$
26 27	set cover, geometric set cover, FPT, W[1]-completeness, APX-completeness, PCP theorem, NP-completeness
28	Thesis domain (Socrates-Erasmus subject area codes)
29 30	11.3 Informatyka
31	Subject classification
32 33 34	D. Software D.127. Blabalgorithms D.127.6. Numerical blabalysis
35	Tytuł pracy w języku polskim
36	Algorytmy parametryzowania i trudność aproksymacji problemu pokrywania zbiorów

odcinkami na płaszczyźnie

38 Contents

39	1.	Introduction
40	2.	Definitions
41		2.1. Geometric Set Cover
42		2.2. Approximation
43		2.3. δ -extensions
44	3	APX-completeness Geometric Set Cover
45	Ο.	3.1. APX-completeness for segments parallel to axes
		3.1.1. MAX-(3,3)-SAT and statement of reduction
46		
47		
48		3.1.2.1. VARIABLE-gadget
49		3.1.2.2. OR-gadget
50		3.1.2.3. CLAUSE-gadget
51		3.1.2.4. Summary
52		3.1.2.5. Summary of construction
53		3.1.3. Construction lemmas and proof of Lemma 1
54	4	FPT for Geometric Set Cover for segments with δ -extensions
55	т.	4.1. FPT for segments
		4.1.1. Axis-parallel segments
56		1 0
57		· ·
58		4.2. FPT for weighted segments with δ -extensions
59	5 .	W[1]-completeness for weighted segments in 3 directions
60	6.	Geometric Set Cover with lines
61		6.1. Lines parallel to one of the axis
62		6.2. FPT for arbitrary lines
63		6.3. APX-completeness for arbitrary lines
64		6.4. 2-approximation for arbitrary lines
-		6.5. Connection with general set cover
65		0.9. Connection with general set cover
66	7.	Geometric Set Cover with polygons
67		7.1. State of the art
60	8	Conclusions

69 Chapter 1

Introduction

The Set Cover problem is one of the most common NP-complete problems. [tutaj referencja]
We are given a family of sets and have to choose the smallest subfamily of these sets that cover
all their elements. This problem naturally extends to settings were we put different weights
on the sets and look for the subfamily of the minimal weight. This problem is NP-complete
even without weights and if we put restrictions on what the sets can be. One of such variants
is Vertex Cover problem, where sets have size 2 (they are edges in a graph).

In this work we focus on another such variant where the sets correspond to some geometric shapes and only some points of the plane have to be covered. When these shapes are rectangles with edges parallel to the axis, the problem can be proven to be W[1]-complete (solution of size k cannot be found in $n^o(k)$ time), APX-complete (for suffciently small $\epsilon > 0$, the problem does not admit $1 + \epsilon$ -approximation scheme) [refrencje].

Some of these settings are very easy. Set cover with lines parallel to one of the axis can be solved in polynomial time.

There is a notion of δ -expansions, which loosen the restrictions on geometric set cover. We allow the objects to cover the points after δ -expansion and compare the result to the original setting. This way we can produce both FPT and EPTAS for the rectangle set cover with δ -extensions [referencie].

Our contribution. In this work, we prove that unweighted geometric set cover with segments is fixed parameter tractable (FPT).

Moreover, we show that geometric set cover with segments is APX-complete for unweighted axis-parallel segments, even with 1/2-extensions. So the problem for very thin rectangles also can't admit PTAS. Therefore, in the efficient polynomial-time approximation scheme (EPTAS) for *fat polygons* by [Har-Peled and Lee, 2009], the assumption about polygons being fat is necessary.

Finally, we show that geometric set cover with weighted segments in 3 directions is W[1]-complete. However, geometric set cover with weighted segments is FPT if we allow δ -extension.

This result is especially interesting, since it's counter-intuitive that the unweighed setting is FPT and the weighted setting is W[1]-complete. Most of such problems (like vertex cover or [wiecej przykladow]) are equally hard in both weighted and unweighted settings.

Chapter 2

Definitions

103 2.1. Geometric Set Cover

In the geometric set cover problem we are are given \mathcal{P} – a set of objects, which are connected subsets of the plane, \mathcal{C} – a set of points in the plane. The task is to choose $\mathcal{R} \subseteq \mathcal{P}$ such that every point in \mathcal{C} is inside some element from \mathcal{R} and $|\mathcal{R}|$ is minimized.

In the parametrized setting for a given k, we only look for a solution \mathcal{R} such that $|\mathcal{R}| \leq k$.

In the weighted setting, there is some given weight function $f: \mathcal{P} \to \mathbb{R}^+$, and we would like to find a solution \mathcal{R} that minimizes $\sum_{R \in \mathcal{R}} f(R)$.

110 2.2. Approximation

Let us recall some definitions related to optimization problems that are used in the following sections.

Definition 1. A polynomial-time approximation scheme (PTAS) for a minimization problem Π is a family of algorithms \mathcal{A}_{ϵ} for every $\epsilon > 0$ such that \mathcal{A}_{ϵ} takes an instance I of Π and in polynomial time finds a solution that is within a factor $(1+\epsilon)$ of being optimal. That means the reported solution has weight at most $(1+\epsilon)opt(I)$, where opt(I) is the weight of an optimal solution for I.

Definition 2. A problem Π is **APX-hard** if assuming $P \neq NP$, there exists $\epsilon > 0$ such that there is no polynomial-time $(1 + \epsilon)$ -approximation algorithm for Π .

$_{ ext{0}}$ 2.3. δ -extensions

121 TODO PLACEHOLDER for introductory text

 δ -extensions is one of the modifications to a problem, that makes geometric set cover problem easier, it has been already used in literature (place some refrence here).

Definition 3 (δ -extensions for center-symmetric objects). For any $\delta > 0$ and a center-symmetric object L with centre of symmetry $S = (x_s, y_s)$, the δ -extension of L is the object $L^{+\delta} = \{(1+\delta) \cdot (x-x_s, y-y_s) + (x_s, y_s) : (x,y) \in L\}$, that is, $L^{+\delta}$ is the image of L under homothety centered at S with scale $(1+\delta)$

The geometric set cover problem with δ -extensions is a modified version of geometric set cover where:

- We need to cover all the points in C with objects from $\{P^{+\delta}: P \in P\}$ (which always include no fewer points than the objects before δ -extensions);
- We look for a solution that is no larger than the optimum solution for the original problem. Note that it does not need to be an optimal solution in the modified problem.
- Formally, we have the following.

Definition 4 (Geometric set cover problem with δ-extensions). The geometric set cover problem with δ-extensions is the problem where for an input instance $I = (\mathcal{P}, \mathcal{C})$, the task is to output a solution $\mathcal{R} \subseteq \mathcal{P}$ such that the δ-extended set $\{R^{+\delta} : R \in \mathcal{R}\}$ covers \mathcal{C} and is no larger than the optimal solution for the problem without extensions, i.e. $|\mathcal{R}| \leq |opt(I)|$.

TODO: Some text

130

131

Definition 5 (Geometric set cover PTAS with δ-extensions). We define a PTAS for geometric set cover with δ-extensions as a family of algorithms $\{\mathcal{A}_{\delta,\epsilon}\}_{\delta,\epsilon>0}$ that each takes as an input instance $I = (\mathcal{P}, \mathcal{C})$, and in polynomial-time outputs a solution $\mathcal{R} \subseteq \mathcal{P}$ such that the δ-extended set $\{R^{+\delta}: R \in \mathcal{R}\}$ covers \mathcal{C} and is within a $(1+\epsilon)$ factor of the optimal solution for this problem without extensions, i.e. $(1+\epsilon)|\mathcal{R}| \leq |opt(I)|$.

¹⁴⁵ Chapter 3

APX-completeness Geometric Set Cover

$_{148}$ 3.1. APX-completeness for segments parallel to axes

- 149 In this section we analyze whether there exists PTAS for geometric set cover for rectangles.
- We show that we can restrict this problem to a very simple setting: segments parallel to axes
- and allow (1/2)-extension, and the problem is still APX-hard. Note that segments are just
- degenerated rectangles with one side being very narrow.
- Our results can be summarized in the following theorem and this section aims to prove it.
- Theorem 1. (axis-parallel segment set cover with 1/2-extension is APX-hard).
- Unweighted geometric set cover with axis-parallel segments in 2D (even with 1/2-extension)
- is APX-hard. That is, assuming $P \neq NP$, there does not exist a PTAS for this problem.
- Theorem 1 implies the following.
- Corollary 1. (rectangle set cover is APX-hard). Unweighted geometric set cover with rectangles (even with 1/2-extension) is APX-hard.
- We prove Theorem 1 by taking a problem that is APX-hard and showing a reduction. For this problem we choose MAX-(3,3)-SAT which we define below.

3.1.1. MAX-(3,3)-SAT and statement of reduction

- Definition 6. MAX-3SAT is the following maximization problem. We are given a 3-CNF formula, and need to find an assignment of variables that satisfies the most clauses.
- Definition 7. MAX-(3,3)-SAT is a variant of MAX-3SAT with an additional restriction
- that every variable appears in exactly 3 clauses. Note that thus, the number of clauses is
- equal to the number of variables.
- In our proof of Theorem 1 we use hardness of approximation of MAX-(3,3)-SAT proved in [Håstad, 2001] and described in Theorem 2 below.
- Definition 8 (α-satisfiable MAX-3SAT formula). MAX-3SAT formula of size n is at most α-satisfiable, if every assignment of variables satisfies no more than αn clauses.
- 172 Theorem 2. [Håstad, 2001]
- For any $\epsilon > 0$, it is NP-hard to distinguish satisfiable (3,3)-SAT formulas from at most (7/8 + ϵ)-satisfiable (3,3)-SAT formulas.

Given an instance I of MAX-(3,3)-SAT, we construct an instance J of axis-parallel segment set cover problem, such that for a sufficiently small $\epsilon > 0$, a polynomial time $(1 + \epsilon)$ approximation algorithm for J would be able to distinguish whether an instance I of MAX(3,3)-SAT is fully satisfiable or is at most $(7/8+\epsilon)$ -satisfiable. However, according to (Theorem
2) the latter problem is NP-hard. This would imply P = NP, contradicting the assumption.
The following lemma encapsulates the properties of the reduction described in this section,

and it allows us to prove Theorem 1.

Lemma 1. Given an instance S of MAX-(3,3)-SAT with n variables and optimum value opt(S), we can construct an instance I of geometric set cover with axis-parallel segments in 2D, such that:

- 185 (1) For every solution X of instance I, there exists a solution of S that satisfies at least 15n-|X| clauses.
- 187 (2) For every solution of instance S that satisfies w clauses, there exists a solution of I of size 15n w.
- 189 (3) Every solution with 1/2-extensions of I is also a solution to the original instance I. 190 Therefore, the optimum size of a solution of I is opt(I) = 15n - opt(S).

We prove Lemma 1 in subsequent sections, but meanwhile let us prove Theorem 1 using Lemma 1 and Theorem 2.

193 Proof of Theorem 1.

194

196

197 198

199

200

201

203

204

Consider any $0 < \epsilon < 1/(15 \cdot 8)$.

Let us assume that there exists a polynomial-time $(1 + \epsilon)$ -approximation algorithm for unweighted geometric set cover with axis-parallel segments in 2D with (1/2)-extensions. We construct an algorithm that solves the problem stated in Theorem 2, thereby proving that P = NP.

Take an instance S of MAX-(3,3)-SAT to be distinguished and construct an instance of geometric set cover I using Lemma 1. We now use the $(1 + \epsilon)$ -approximation algorithm for geometric set cover on I. Denote the size of the solution returned by this algorithm as approx(I). We prove that if in S one can satisfy at most $(\frac{7}{8} + \epsilon)n$ clauses, then $approx(I) \ge 15n - (\frac{7}{8} + \epsilon)n$ and if S is satisfiable, then $approx(I) < 15n - (\frac{7}{8} + \epsilon)n$.

Assume S satisfiable. From the definition of S being satisfiable, we have:

$$opt(S) = n.$$

From Lemma 1 we have:

$$opt(I) = 14n.$$

Therefore,

$$approx(I) \le (1+\epsilon)opt(I) = 14n(1+\epsilon) = 14n + 14\epsilon \cdot n =$$

$$= 14n + (15\epsilon - \epsilon)n < 14n + \left(\frac{1}{8} - \epsilon\right)n = 15n - \left(\frac{7}{8} + \epsilon\right)n$$

Assume S is at most $(\frac{7}{8} + \epsilon)$ satisfiable. From the defintion of S being at most $(\frac{7}{8} + \epsilon)$ n satisfiable, we have:

$$opt(S) \le \left(\frac{7}{8} + \epsilon\right)n$$

From Lemma 1 we have:

$$opt(I) \geq 15n - \left(\frac{7}{8} + \epsilon\right)n$$

Since a solution to I with $\frac{1}{2}$ -extensions is also a solution without extentions, by Lemma 1 (3.), we have:

$$approx(I) \ge opt(I) = 15n - \left(\frac{7}{8} + \epsilon\right)n$$

Therefore, by using the assumed $(1 + \epsilon)$ -approximation algorithm, it is possible to distinguish the case when S is satisfiable from the case when it is at most $(\frac{7}{8} + \epsilon)n$ satisfiable, it suffices to compute approx(I) with $15n - (\frac{7}{8} + \epsilon)n$. Hence, the assumed approximation algorithm cannot exist, unless P = NP.

3.1.2. Reduction

211

We proceed to the proof of Lemma 1. That is, we show a reduction from MAX-(3,3)-SAT problem to geometric set cover with segments parallel to axis. Moreover, the obtained instance of geometric set cover will be robust to 1/2-extensions (have the same optimal solution after 1/2-extension).

The construction will be composed of 2 types of gadgets: VARIABLE-gadgets and

The construction will be composed of 2 types of gadgets: **VARIABLE-gadgets** and **CLAUSE-gadgets**. CLAUSE-gadgets would be constructed using two **OR-gadgets** connected together.

19 3.1.2.1. VARIABLE-gadget

VARIABLE-gadget is responsible for choosing the value of a variable in a CNF formula. It allows two minimum solutions of size 3 each. These two choices correspond to the two Boolean values of the variable corresponding to this gadget.

Points. Define points a, b, c, d, e, f, g, h as follows, where L = 12n:

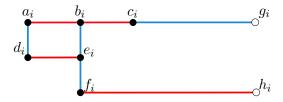


Figure 3.1: **VARIABLE-gadget.** We denote the set of points marked with black circles as $pointsVariable_i$, and they need to be covered (are part of the set \mathcal{C}). Note that some of the points are not marked as black dots and exists only to name segments for further reference. We denote the set of red segments as $chooseVariable_i^{false}$ and the set of blue segments as $chooseVariable_i^{true}$.

$$a = (-L, 0)$$
 $b = (-\frac{2}{3}L, 0)$ $c = (-\frac{1}{3}L, 0)$ $d = (-L, 1)$
 $e = (-\frac{2}{3}L, 1)$ $f = (-\frac{2}{3}L, 2)$ $g = (L, 0)$ $h = (L, 2)$

Let us define:

pointsVariable =
$$\{a, b, c, d, e, f\}$$

and

224

$$pointsVariable_i = pointsVariable + (0, 4i)$$

We denote $a_i = a + (0, 4i)$ etc.

226 **Segments.** Let us define:

chooseVariable
$$_i^{true} = \{(a_i, d_i), (b_i, f_i), (c_i, g_i)\}$$

chooseVariable $_i^{false} = \{(a_i, c_i), (d_i, e_i), (f_i, h_i)\}$

 $\mathsf{segmentsVariable}_i^{} = \mathsf{chooseVariable}_i^{true} \cup \mathsf{chooseVariable}_i^{false}$

- Lemma 2. For any $1 \le i \le n$, points in pointsVariable, can be covered using 3 segments from segmentsVariable.
- 229 Proof. We can use either set chooseVariable_i^{true} or chooseVariable_i^{false}.
- Lemma 3. For any $1 \le i \le n$, points in pointsVariable_i can not be covered with fewer than 3 segments from segmentsVariable_i.
- 232 *Proof.* No segment of segmentsVariable_i covers more than one point from $\{d_i, f_i, c_i\}$, therefore 233 pointsVariable_i can not be covered with fewer than 3 segments.
- Lemma 4. For every set $A \subseteq \text{segmentsVariable}_i \ such \ that \ A \ covers \ \mathsf{pointsVariable}_i \ and (c_i, g_i), (f_i, h_i) \in A, \ it \ holds \ that \ |A| \ge 4.$
- 236 *Proof.* No segment from segmentsVariable_i covers more than one point from $\{a_i, e_i\}$, therefore 237 pointsVariable_i $\{c_i, f_i, g_i, h_i\}$ can not be covered with fewer than 2 segments.

238 3.1.2.2. OR-gadget

- OR-segment connects input and output segments that are connected to other parts of constructions.
- Output segment is part of OR-segment, but iunput is not.
- For every solution \mathcal{R} of the whole construction. Define \mathcal{R}' as intersection of \mathcal{R} and the gadget segments. Minimum solution of OR-gadget has size w, i.e. $\mathcal{R}' \leq w$. output segments can be part of \mathcal{R}' only if $input_x$ or $input_y$ are part of the chosen solution \mathcal{R} . If none of them are chosen, then solution containing output segment has weight at least w+1. Therefore the following formula holds:

$$output \in \mathcal{R}' \land |\mathcal{R}'| = w \Rightarrow (x \in R) \lor (y \in R)$$

Only 3 points that belong to this segment: $l_{i,j}, p_{i,j}, v_{i,j}$ can be covered by segment not from the OR-gadget.



Figure 3.2: **OR-gadget.** Figure presenting OR-gadget: segments from $\mathsf{chooseOr}_{i,j}^{false}$ are red, segments from $\mathsf{chooseOr}_{i,j}^{frue}$ are blue, segments from $\mathsf{orMoveVariable}_{i,j}$ are yellow and green. Dark blue segment is an output segment. Grey segments $input_x$ and $input_y$ are input segments that are not part of $\mathsf{segmentsOr}_{i,j}$.

Points.

250

$$l_0 = (0,0)$$
 $m_0 = (0,1)$ $n_0 = (0,2)$ $o_0 = (0,3)$
 $p_0 = (0,4)$ $q_0 = (1,1)$ $r_0 = (1,3)$ $s_0 = (2,1)$
 $t_0 = (2,2)$ $u_0 = (2,3)$ $v_0 = (3,2)$

$$vec_{i,j} = (10i + 3 + 3j, 4n + 2j)$$

Define $\{l_{i,j}, m_{i,j} \dots v_{i,j}\}$ as $\{l_0, m_0 \dots v_0\}$ shifted by $vec_{i,j}$ Note that $v_{i,0} = l_{i,1}$ (see Figure 3.3)

pointsOr_{i, i} =
$$\{l_{i,j}, m_{i,j}, n_{i,j}, o_{i,j}, p_{i,j}, q_{i,j}, r_{i,j}, s_{i,j}, t_{i,j}, u_{i,j}\}$$

Note that $\mathsf{pointsOr}_{i,j}$ does not include point $v_{i,j}$

254 Segments. We define names subsets of segments, to refer to them in lemmas.

$$\begin{split} \text{chooseOr}_{i,j}^{false} &= \{(q_{i,j}, r_{i,j}), (s_{i,j}, u_{i,j})\} \\ \text{chooseOr}_{i,j}^{true} &= \{(m_{i,j}, s_{i,j}), (o_{i,j}, u_{i,j}), (t_{i,j}, v_{i,j})\} \end{split}$$

orMoveVariable_{i,j} = {
$$(l_{i,j}, n_{i,j}), (n_{i,j}, p_{i,j})$$
}

Segments in OR-gadget:

$$\mathsf{segmentsOr}_{i,j} = \mathsf{chooseOr}_{i,j}^{false} \cup \mathsf{chooseOr}_{i,j}^{true} \cup \mathsf{orMoveVariable}_{i,j}$$

Lemma 5. For any $1 \le i \le n, j \in \{0,1\}$ and $x \in \{l_{i,j}, p_{i,j}\}$, points in points $\operatorname{Or}_{i,j} - \{x\} \cup \{v_{i,j}\}$ can be covered with 4 segments from segments $\operatorname{Or}_{i,j}$.

Proof. We can do that using one segment from orMoveVariable_{i,j}, the one that does not cover x, and all segments from chooseOr $_{i,j}^{true}$.

Lemma 6. For any $1 \le i \le n, j \in \{0, 1\}$, points in points $Or_{i,j}$ can be covered with 4 segments from segments $Or_{i,j}$.

262 Proof. We can do that using segments from $\mathsf{orMoveVariable}_{i,j}$ and $\mathsf{chooseOr}_{i,j}^{false}$.

3.1.2.3. CLAUSE-gadget

CLAUSE-gadget is responsible for calculating if variables values assigned in variable gadgets satisfy the respective clause in CNF. It has minimum solution of weight w if and only if the clause is satisfied, i.e. at least one of the respective variables is assigned a correct value. Otherwise it has minimum solution of weight w+1. This way, by analyzing the minimum solution for the whole problem, we can tell how many clauses were possible to satisfy in the optimum solution of CNF.

The CLAUSE-gadgets consist of two OR-gadgets. It would be inconvenient to position the CLAUSE-gadents in between the very long variable segments. Instead, we use a simple auxiliary gadget to transfer whether the segment is in a solution, i.e. segments $(x_{i,0}, x_{i,1}), (y_{i,0}, y_{i,1}), (z_{i,0}, z_{i,1})$. Each gadget consists of two segments $(x_{i,0}, x_{i,1}), (x_{i,1}, a)$. These are the only segments that can cover $x_{i,1}$. If $x_{i,0}$ is already covered by some other gadget, we can cover $x_{i,1}$ by the other segment covering another point from the gadget, say a. If $x_{i,0}$ is not covered, then the only way to cover $x_{i,0}$ is to use segment $(x_{i,0}, x_{i,1})$. Intuitively, the two segments transfer the state of $x_{i,0}$ onto a, but there are less restrictions on where a can be placed, simplifying the construction.



Figure 3.3: **CLAUSE-gadget.** This figure presents CLAUSE-gadget. Every green rectangle is an OR-gadget. y-coordinates of $x_{i,0}$, $y_{i,0}$ and $z_{i,0}$ depend on the variables in the i-th clause. Grey segments corresponds to the values of variables satisfying the i-th clause.

Points. TODO: Rephrase it

Assuming clause $C_i = a \lor b \lor c$, function idx(w) returns index of the variable w, function neg(w) returns whether variable w is negated in a clause.

$$x_{i,0} = (10i+1, 4 \cdot idx(a) + 2 \cdot neg(c)) \quad x_{i,1} = (10i+1, 4n)$$

$$y_{i,0} = (10i+2, 4 \cdot idx(b) + 2 \cdot neg(b)) \quad y_{i,1} = (10i+2, 4n+4)$$

$$z_{i,0} = (10i+3, 4 \cdot idx(c) + 2 \cdot neg(c)) \quad z_{i,1} = (10i+3, 4n+6)$$

$$\text{moveVariable}_i = \{x_{i,j} : j \in \{0,1\}\} \cup \{y_{i,j} : j \in \{0,1\}\} \cup \{z_{i,j} : j \in \{0,1\}\}$$

$$\text{pointsClause}_i = \text{moveVariable}_i \cup \text{pointsOr}_{i,0} \cup \text{pointsOr}_{i,1} \cup \{v_{i,1}\}$$

$$\text{Segments.}$$

$$\text{segmentsClause}_i = \{(x_{i,0}, x_{i,1}), (y_{i,0}, y_{i,1}), (z_{i,0}, z_{i,1}), (x_{i,1}, l_{i,0}), (y_{i,1}, p_{i,0}), (z_{i,1}, p_{i,1}), \} \cup \{x_{i,1}, x_{i,1}, x$$

 $\cup \; \mathsf{segmentsOr}_{i,0} \cup \mathsf{segmentsOr}_{i,1}$

Lemma 7. For any $1 \le i \le n$ and $a \in \{x_{i,0}, y_{i,0}, z_{i,0}\}$, there is a solClause $_i^{true,a} \subset \text{segmentsClause}_i^{table}$ with $|\text{solClause}_i^{true,a}| = 11$ that covers points in pointsClause $_i - \{a\}$.

Proof. For $a=x_{i,0}$ (analogous proof for $y_{i,0}$): First we use Lemma 5 twice with excluded $x=l_{i,0}$ and $x=l_{i,1}=v_{i,0}$, resulting with 8 segments ${\sf chooseOr}_{i,0}^{true} \cup {\sf chooseOr}_{i,1}^{true}$ which cover all required points apart from $x_{i,1},y_{i,0},y_{i,1},z_{i,0},z_{i,1},l_{i,0}$. We cover those using additional 3 segments: $\{(x_{i,1},l_{i,0}),(y_{i,0},y_{i,1}),(z_{i,0},z_{i,1})\}$

For $a=z_{0,i}$: Using Lemma 6 and Lemma 5 with $x=p_{i,1}$, resulting with 8 segments chooseOr_{i,0}^{false} \cup chooseOr_{i,1}^{true} which cover all required points apart from $x_{i,0}, x_{i,1}, y_{i,0}, y_{i,1}, z_{i,1}, p_{i,1}$. We cover those using additional 3 segments: $\{(x_{i,0}, x_{i,1}), (y_{i,0}, y_{i,1}), (z_{i,1}, p_{i,1})\}$.

Lemma 8. For any $1 \le i \le n$ there is solClause $_i^{false} \subset \text{segmentsClause}_i$ with $|\text{solClause}_i^{false}| = 12$ that covers points in pointsClause $_i$.

Proof. Using Lemma 6 twice we can cover pointsOr_{i,0} and pointsOr_{i,1} with 8 segments.

To cover the remaining points we additionally use: $\{(x_{i,0}, x_{i,1}), (y_{i,0}, y_{i,1}), (z_{i,0}, z_{i,1}), (t_{i,1}, v_{i,1})\}$

Lemma 9. For any $1 \le i \le n$:

304

- 298 (1) points in points Clause, can not be covered using any subset of segments from segments Clause, of size smaller than 12;
- 300 (2) points in pointsClause $_i \{x_{i,0}, y_{i,0}, z_{i,0}\}$ can not be covered using any subset of segments 301 from segmentsClause $_i$ of size smaller than 11.

Proof of (1). No segment in segmentsClause_i covers more than 2 points from $\{x_{i,0}, y_{i,0}, z_{i,0}, l_{i,0}, p_{i,0}, q_{i,0}, u_{i,0}, v_{i,0}, u_{i,0}, v_{i,0}, u_{i,1}, u_{i$

Therefore we need to use at least 12 segments.

Proof of (2). We can choose disjoint sets X, Y, Z such that $X \cup Y \cup Z \subseteq \mathsf{pointsClause}_i$ $\{x_{i,0}, y_{i,0}, z_{i,0}\}$ and there are no segments covering points from different sets. And we prove lower bounds for each of these sets.

$$X = \{x_{i,1}, y_{i,1}, z_{i,1}\}$$

No two points in X are covered with one segment of segmentsClause_i, so it must be covered with 3 different segments.

$$Y = \mathsf{pointsOr}_{i,0} - \{l_{i,0}, p_{i,0}\}$$

$$Z = \mathsf{pointsOr}_{i,1} - \{l_{i,1}, p_{i,1}\}$$

For both Y and Z we can check all of the subsets of 3 segments of segmentsClause_i with brutforce that none of them cover the set of points, so both Y and Z have to be covered with disjoined sets of 4 segments.

TODO: Funny fact, neither Y nor Z doesn't have independent set of size 4.

Therefore pointsClause_i must be covered with at least 3+4+4=11 segments.

3.1.2.4. Summary

Add some smart lemmas that sets will be exclusive to each other.

Lemma 10. Robustness to 1/2-extensions. For every segment $s \in \mathcal{P}$, s and $s^{+1/2}$ cover the same points from \mathcal{C} .

Proof. We can just check every segment. Most of the segments s are collinear only with points that lay on s, so trivially $s^{+\frac{1}{2}}$ cannot cover more points than s does.

TODO: list problematic segments here

In the same gadget: $(n_{i,j}, p_{i,j})$ does not cover $m_{i,j}$ and symmetrically. $(t_{i,j}, v_{i,j})$ does not cover $n_{i,j}$. $(o_{i,0}, u_{i,0})$ does not cover $m_{i,1}$ and symmetrically. $(y_{i,1}, p_{i,0})$ does not cover $n_{i,j}$.

From different gadgets: (b_i, f_i) after $\frac{1}{2}$ -extensions does not cover b_{i+1} point.

VARIABLE-gadget's (a_i, c_i) after $\frac{1}{2}$ -extensions does not cover any points $x_{i,0}, y_{i,0}$ or $z_{i,0}$ from CLAUSE-gadget.

3.1.2.5. Summary of construction

We define:

324

327

329

330

331

332

$$\mathcal{C} := \bigcup_{1 \leq i \leq n} \mathsf{pointsVariable}_i \cup \mathsf{pointsClause}_i$$

$$\mathcal{P} := \bigcup_{1 \leq i \leq n} \mathsf{segmentsVariable}_i \cup \mathsf{segmentsClause}_i$$

The subsequent sections define these sets.

We prove some properties of different gadgets. Every segment for a gadget will only cover points in this gadget (won't interact with any different gadget), so we can prove lemmas *locally*.

TODO: y axis is increasing values downward on figures (not upwards like in normal).

333 3.1.3. Construction lemmas and proof of Lemma 1

In order to prove Lemma 1 we introduce several auxiliary lemmas proving properties of the construction described in the previous section.

Consider an instance S of MAX-(3,3)-SAT of size n with optimum solution satisfying k clauses. Let us construct an instance (C, P) of geometric set cover as described in Section 3.1.2 for instance S of MAX-(3,3)-SAT.

Lemma 11. Instance $(\mathcal{C}, \mathcal{P})$ of geometric set cover admits a solution of size 15n - k.



Figure 3.4: General schema.

General layout of VARIABLE-gadget and CLAUSE-gadget and how they interact with each other.

TODO: Rename Choose X to VARIABLE-gadget and Clause C to CLAUSE-gadget.

Proof. Let the clauses in S be $c_1, c_2 \ldots c_n$ and the variables be $x_1, x_2 \ldots x_n$. Let the assignment of the variables in the optimum solution to S be $\phi: \{x_1, x_2 \dots x_n\} \to \{\texttt{true}, \texttt{false}\}.$

340

341

342

344

348

350

351

352

353

355

356

357

358

361

362

363

365

366

367

368

369

We cover every VARIABLE-gadget with solution described in Lemma 2, in the i-th gadget choosing the set of segments corresponding to the value of $\phi(x_i)$.

For every clause that is satisfied, say c_i , let us name the variable that is **true** in it as x_i and point corresponding to x_i in pointsClause, as a. Points in pointsClause, are covered with set solClause $_i^{true,a}$ described in Lemma 7. For every clause that is not satisfied, say c_j , points in pointsClause_i are covered with set solClause_i described in Lemma 8.

Formally we define sets responsible for choosing variable and satisfing the variable, R_i and C_i respectively, as following:

$$R_i = \begin{cases} \mathsf{chooseVariable}_i^{true} & \text{if } \phi(x_i) = \mathsf{true} \\ \mathsf{chooseVariable}_i^{false} & \text{if } \phi(x_i) = \mathsf{false} \end{cases}$$

$$C_i = \begin{cases} \mathsf{solClause}_i^{true,a} & \text{if } c_i \text{ satisfied} \\ \mathsf{solClause}_i^{false} & \text{if } c_i \text{ not satisfied} \end{cases}$$

$$\mathcal{R} = \bigcup_{i=1}^n \{R_i \cup C_i : 1 \leq i \leq n\}.$$

This set covers all the points from C, because the sets R_i , C_i individually cover their corresponding gadgets, as proved in the respective lemmas.

All of these sets are disjoint, so the size of the obtained solution is:

$$|\mathcal{R}| = \sum_{i=1}^{n} R_i + \sum_{i=1}^{n} C_i = 3n + 11k + 12(n-k) = 15n - k.$$

354

Lemma 12. Suppose we have a solution \mathcal{R} of the instance $(\mathcal{C}, \mathcal{P})$ of geometric set cover. Then there exists a solution \mathcal{R}' , such that $|\mathcal{R}'| \leq |\mathcal{R}|$, and for each VARIABLE-gadget \mathcal{R}' contains at most one of the segments (c_i, g_i) and (f_i, h_i) .

Proof. Assume that we have $\{(c_i, g_i), (f_i, h_i)\}\subseteq \mathcal{R}$ for some i. We will show how to modify \mathcal{R} into \mathcal{R}' , such that the number of such i decreases, while \mathcal{R}' is still a valid solution of $(\mathcal{C}, \mathcal{P})$, and $|\mathcal{R}'| \leq |\mathcal{R}|$. Then, by repeating this procedure, we can eventually construct a solution 359 satisfying the property from the Lemma. 360

To construct \mathcal{R}' , we remove either (c_i, g_i) or (f_i, h_i) from \mathcal{R} , and then add one extra segment to make \mathcal{R}' valid. Recall that the *i*-th VARIABLE-gadget corresponds to variable x_i in S. As every variable in S is used in exactly 3 clauses, one of the ways of setting x_i (to either true or false) must satisfy at least 2 clauses. If that setting is $x_i = \text{true}$, then we remove (f_i, h_i) , otherwise we remove (c_i, g_i) . Now, there exists at most one CLAUSE-gadget which needs adjustment to make \mathcal{R}' valid; we do that by adding $(t_{j,1}, v_{j,1})$ to \mathcal{R}' .

TODO: Can we really just remove one segment and add another one? I'd think we need to "restructure" R around points Variable; (saving one segment due to Lemma 3 and Lemma 4) and then again restructure \mathcal{R} around the clause that we need to fix?

Lemma 13. Suppose we have a solution \mathcal{R} of the instance $(\mathcal{C}, \mathcal{P})$ of geometric set cover that is of size w. Then there exists a solution of S that satisfies at least 15n - w clauses.

Proof. Let the clauses in S be $c_1, c_2 \ldots c_n$ and the variables be $x_1, x_2 \ldots x_n$. Given a solution \mathcal{R} of the instance $(\mathcal{C}, \mathcal{P})$ of geometric set cover, we use Lemma 12 to modify \mathcal{R} such that for any i it contains at most one of (c_i, g_i) and (f_i, h_i) ; this may decrease the cost of \mathcal{R} , but that does not matter in the subsequent construction. To simplify notation, in the remainder of this proof we use \mathcal{R} to refer to the modified solution.

Given \mathcal{R} , we construct a solution of S by constructing an assignment of variables ϕ : $\{x_1, x_2 \dots x_n\} \to \{\text{true}, \text{false}\}\$ that satisfies at least 15n-w clauses in S.

Variables Recall that due to Lemma 12, \mathcal{R} contains at most one of (c_i, g_i) and (f_i, h_i) . We define the value $\phi(x_i)$ for the variable x_i as follows:

$$\begin{cases} \phi(x_i) = \text{true} & \text{if } (c_i, g_i) \in \mathcal{R} \\ \phi(x_i) = \text{false} & \text{if } (f_i, h_i) \in \mathcal{R} \\ \phi(x_i) = \text{false} & \text{otherwise} \end{cases}$$
(3.1)

Moreover, from Lemma 3 we get $|pointsVariable_i \cap \mathcal{R}| \geq 3$ for every i.

Clauses For a clause $C_i = x \vee y \vee z$, \mathcal{R} needs to use at least 11 segments to cover pointsClause_i $-\{x,y,z\}$ in CLAUSE-gadget (Lemma 9).

TODO: maybe put something with cases and names of sets as above

Moreover, if all of the points $\{x_{i,0}, y_{i,0}, z_{i,0}\}$ are not covered by the segments from $\mathcal{R} \cap \mathsf{pointsVariable}_i$, then \mathcal{R} needs to cover $\mathsf{pointsClause}_i$ with at least 12 segments by Lemma 9.

TODO: Maybe remove section below, because we do this calculation at the end anyway We covered CLAUSE-gadget with at least 11 or at least 12 segments:

$$|\bigcup_{i=1}^n \mathsf{segmentsClause}_i \cap \mathcal{R}| \geq 11n + a$$

where a is the number of clauses where none of the points $x_{i,0}, y_{i,0}, z_{i,0}$ were covered by $\mathcal{R} \cap \mathsf{segmentsVariable}_i$ for their respective variable x_j .

Satisfied clauses with chosen variable assignment. Consider a clause, say c_i . If none of the points $x_{i,0}, y_{i,0}, z_{i,0}$ in pointsClause_i were covered by segments from $\mathcal{R} \cap \text{segmentsVariable}_j$, this clause is not satisfied by assignment ϕ .

If one of these points is covered by segments from VARIBALE-gadget (TODO better this or $\mathcal{R} \cap \mathsf{segmentsVariable}_j$), then denote this point as t and say it corresponds to variable x_j . Consider the cases of choosing value of $\phi(x_j)$ in equation (3.1).

If \mathcal{R} contains exactly one of the segments (c_j, g_j) and (f_j, h_j) , then the value $\phi(x_j)$ satisfies c_i .

If \mathcal{R} contains neither (c_j, g_j) nor (f_j, h_j) , then it is impossible that t is covered by segments in $\mathcal{R} \cap \text{segmentsVariable}_i$.

This means that ϕ satisfies all but at most a clauses in S.

To conclude, we proved that given a solution of $(\mathcal{C}, \mathcal{P})$ of size w, we have constructed a variables assignment ϕ that satisfies at least n-a clauses of S. Finally, note that

$$w \ge 3n + 11(n - a) + 12a = 3n + 11n + a = 14n + a,$$

hence

379

380

383

388

389

390

391

392

394

395

396

397

398

399

400

401

$$15n - w \le 15n - 14n - a = n - a.$$

So ϕ satisfies at least 15n - w clauses of S.

We are ready to conclude the proof of Lemma 1.

Proof of Lemma 1. By Lemma 11, we know that there exists a solution to $(\mathcal{C}, \mathcal{P})$ of size 15n - k, so:

$$opt((\mathcal{C}, \mathcal{P})) \le 15n - k.$$

Since the optimum solution of S satisfies k clauses, then according to Lemma 13:

$$opt((\mathcal{C}, \mathcal{P})) \ge 15n - k.$$

Therefore, the solution given by Lemma 11 of size 15n - k is an optimum solution to the instance $(\mathcal{C}, \mathcal{P})$.

Chapter 4

FPT for Geometric Set Cover for segments with δ -extenstions

4.1. FPT for segments

In this section we consider the fixed-parameter tractable algorithms for unweighted geometric set cover with segments. Setting where segments are limited to be axis-parallel (or limited to constant number of directions) has an FPT algorithm already present in literature. We present an FPT algorithm for unweighted geometric set cover with segments, where segments are in arbitrary directions.

4.1.1. Axis-parallel segments

416

417

418

420

421

422

423

424

425

426

428

429

430

431

432

You can find this in Platypus book. (TODO add referece)

We show an $\mathcal{O}(2^k)$ -time branching algorithm. In each step, the algorithm selects a point a which is not yet covered, branches to choose one of the two directions, and greedily chooses a segment in that direction to cover a. This proceeds until either all points are covered or k segments are chosen.

Let us take the point $a = (x_a, y_a)$ which is the smallest among points that are not yet covered in the lexicographic ordering of points in \mathbb{R}^2 . We need to cover a with some of the remaining segments.

Branch over the choice of one of the coordinates (x or y); without loss of generality, let us assume we chose x. Among the segments lying on line $x = x_a$, we greedily add to the solution the one that covers the most points. As a was the smallest in the lexicographical order, then all points on line $x = x_a$ have the y-coordinate larger than y_a . Therefore, if we denote the greedily chosen segment as s, then any other segment on $x = x_a$ that covers a can only cover a (possibly improper) subset of points covered by s. Thus, greedily choosing s is optimal.

In each step of the algorithm we add one segment to the solution, thus each branch can stop at depth k. If no branch finds a solution, then that means a solution of size at most k does not exist.

TODO: Maybe split it into theorem + algorithm + explanation like in section 4.1.2

Remark 1. The same algorithm can be used for segments in d directions, where we branch over d directions and it runs in complexity $\mathcal{O}(d^k)$.

4.1.2. Segments in arbitrary directions 435

- In this section we consider setting where segments are not constrained to only d directions. 436 We present a fixed-parameter tractable algorithm, where parameter is the size of the solution.
- **Theorem 3.** (FPT for segment cover). There exists an algorithm that given a family 438 \mathcal{P} of n segments (in any direction), a set of m points \mathcal{C} and a parameter k, runs in time 439 $k^{O(k)} \cdot (nm)^2$, and outputs a subfamily $\mathcal{R} \subseteq \mathcal{P}$ such that $|\mathcal{R}| \leq k$ and \mathcal{R} covers all points in C, or determines that such a set R does not exist.
- We will need the following lemmas.

442

456

457

458

459

466

- **Lemma 14.** Given an instance $(\mathcal{P},\mathcal{C})$ of the segment cover problem, without a loss of generality we can assume that no segment covers a superset of what another segment covers. That 444 is, for any distinct $A, B \in \mathcal{P}$, we have $A \cap C \not\subseteq B \cap C$ and $A \cap C \not\supseteq B \cap C$. 445
- Proof. Trivial. 446
- **Lemma 15.** Given an instance $(\mathcal{P},\mathcal{C})$ of the segment cover problem, if there exists a line L 447 with at least k+1 points on it, then there exists a subset $A \subseteq \mathcal{P}$, $|A| \leq k$, such that every solution \mathcal{R} with $|\mathcal{R}| \leq k$ satisfies $|\mathcal{A} \cap \mathcal{R}| \geq 1$. Moreover, such a subset can be found in polunomial time. 450
- *Proof.* First we use Lemma 14. 451
- Let us enumerate the points from \mathcal{C} that lie on L as $x_1, x_2, \ldots x_t$ in the order in which 452 they appear on L. Every segment that is not collinear with L can cover at most one of these 453 points. Therefore, in any solution of size not larger than k, among any k of these points at least one must be covered with segment collinear with L. 455
 - Therefore, every solution needs to take one of the segments collinear with L that covers any of the points $x_1, x_2, \dots x_k$. After using reduction from Lemma 14, there are at most k such segments that are distinct.
 - We are ready to prove Theorem 3.
- Proof of Theorem 3. 460
- We will prove this theorem by presenting a branching algorithm that works in desired 461 complexity. It branches over the choice of segments to cover lines with a lot of points, then 462 finally solving the small instance, where every line has at most k points by checking all possible 463 solutions. 464
- **Algorithm.** First we use Lemma 14. 465
 - Next, we present a recursive algorithm. Given an instance of the problem:
- (1) If there exist a line with at least k+1 points from \mathcal{C} , we branch over adding to the 467 solution one of the at most k possible segments provided by Lemma 15; name this 468 segment S. Then we find a solution \mathcal{R} for the problem for points $\mathcal{C}-S$, segments 469 $\mathcal{P} - \{S\}$, and parameter k - 1. We return $\mathcal{R} \cup \{S\}$. 470
- (2) If every line has at most k points on it and $|\mathcal{C}| > k^2$, then answer NO. 471
- (3) If $|\mathcal{C}| \leq k^2$, solve the problem by brute force: check all subsets of \mathcal{P} of size at most k. 472

Correctness. Lemma 15 proves that at least one segment that we branch over in (1) must be present in every solution \mathcal{R} with $|\mathcal{R}| \leq k$. Therefore, the recursive call can find a solution, provided there exists one.

In (2) the answer is no, because every line covers no more than k points from \mathcal{C} , which implies the same about every segment from \mathcal{P} . Under this assumption we can cover only k^2 points with a solution of size k, which is less than $|\mathcal{C}|$.

Checking all possible solutions in (3) is trivially correct.

Complexity. In the leaves of recursion we have $|\mathcal{C}| \leq k^2$, so $|\mathcal{P}| \leq k^4$, because every segments can be uniquely identified by the two extreme points it covers (by Lemma 14). Therefore, there are $\binom{k^4}{k}$ possible solutions to check, each can be checked in time $O(k|\mathcal{C}|)$. Therefore, (3) takes time $k^{O(k)}$.

In this branching algorithm our parameter k is decreased with every recursive call, so we have at most k levels of recursion with branching over k possibilites. Candidates to branch over can be found on each level in time $O((nm)^2)$.

Reduction from Lemma 14 can be implemented in time $O(n^2m)$.

It follows that the overall complexity is $O((nm)^2 \cdot k^O(k))$

489 4.2. FPT for weighted segments with δ -extensions

490 TODO: Some intro

476

488

508

Theorem 4 (FPT for weighted segment cover with δ -extensions). There exists an algorithm that given a family \mathcal{P} of n weighted segments (in any direction), a set of m points \mathcal{C} , and parameters k and $\delta > 0$, runs in time $f(k,\delta) \cdot (nm)^c$ for some computable function f and a constant c, and outputs a set $\mathcal{R} \subseteq \mathcal{P}$ such that $|\mathcal{R}| \leq k$ and $\mathcal{R}^{+\delta}$ covers all points in \mathcal{C} , or determines that such a set \mathcal{R} does not exist.

To solve this problem we will introduce a lemma about choosing a *good* subset of points.

TODO: Some intuition

Definition 9. For a set of collinear points C, a subset $A \subseteq C$ is (k, δ) -good if for any set of segments R that covers A and such that $|R| \leq k$, it holds that $R^{+\delta}$ covers C.

Lemma 16. There exists an algorithm that for any set of collinear points C, $\delta > 0$ and $k \ge 1$, outputs a (k, δ) -good set $A \subseteq C$ of size at most $f(k, \delta)$ for some computable function f. This algorithm runs in time $O(|C| \cdot f(k, \delta))$.

Froof. We prove this for a fixed δ by induction over k.

Inductive hypothesis. For any set of collinear points C, there exists an algorithm that runs in time $O(|C|k(1+\frac{1}{\delta}))$ and finds a set A such that:

- A is (ℓ, δ) -good for every $1 \le \ell \le k$,
- A has size $|A| < f(\delta, k)$ for some computable function f,
 - extreme points from C are in A.

Base case for k=1. It is sufficient that A consists of 2 points: extreme points from C 509 or a single point if |C|=1. 510

If they are covered with one segment, it must be a segment that includes the extreme points from C, so it covers the whole set C.

Inductive step. Assuming inductive hypothesis for any set of collinear points C and for parameter k, we will prove hypothesis for k+1.

Let be s the minimal segment that includes all points from C. That is, the extreme points of C are endpoints of s.

We define $M = \left[1 + \frac{2}{\delta}\right]$ subsegments of s in the following way. We split s into M parts v_i of equal length, that is $|v_i| = \frac{|s|}{M}$ for each $1 \le i \le M$.

Let C_i be the subset of C consisting of points laying on v_i .

Let t_i be the segment with endpoints being the extreme points of C_i (it might be degenerated segment if $|C_i| = 1$ or it might be empty if C_i is empty).

TODO: Add a picture with v_i and t_i here

511

513

515

516

517

518

519

521

522

523

524

529

530

531

532

536

538

541

543

We use the inductive hypothesis to choose (k, δ) -good sets A_i for sets C_i . Note that if $|C_i| \leq 1$, then $A_i = C_i$ and it's still a (k, δ) -good set for C_i .

Then we define $A = \bigcup_{i=1}^{M} A_i$. Thus A includes the extreme points of C, because they are included in the sets A_1 and A_M .

Proof that A is (k, δ) -good for C. Let us take any cover of A with k+1 segments and call it \mathcal{R} . 528

For every segment t_i , if there exists a segment x in \mathcal{R} that is disjoint with t_i , then we have a cover of A_i with at most k segments using $\mathcal{R} - \{x\}$. Since A_i is (k, δ) -good for t_i and C_i , then $(\mathcal{R} - \{x\})^{+\delta}$ covers C_i . So $\mathcal{R}^{+\delta}$ covers C_i as well.

If there exists a segment t_i for which a segment x as defined above does not exist, then all k+1 segments that cover A_i intersect with t_i . (Note: There may exist only one such segment t_i). From the inductive hypothesis end points of s are in A_1 and A_M respectively, so \mathcal{R} must cover them. Hence there must exist segments starting in the ends of s and ending somewhere in t_i . Let us call these two segments y and z. It follows that: $|y| + |z| + |t_i| \ge |s|$. Since $|t_i| \le |v_i| = \frac{|s|}{M} \le \frac{|s|}{1+\frac{2}{\delta}} = \frac{|s|\delta}{\delta+2}$, we have $\max(|y|,|z|) \ge |s|(1-\frac{\delta}{\delta+2})/2 = \frac{|s|}{\delta+2}$.

TODO: Add a picture with such segments here

After δ -extension, the longer of these segments will lengthen both ways by at least:

$$\frac{|s|\delta}{\delta+2} = \frac{|s|}{1+\frac{2}{\delta}} > \frac{|s|}{M} = v_i > t_i.$$

Therefore the longer of segments y and z will cover the segment t_i after δ -extension, 539 therefore $\mathcal{R}^{+\delta}$ covers C_i . Since $C = \bigcup_{i=1}^M C_i$, then $\mathcal{R}^{+\delta}$ covers C. 540

Complexity We use the recursive algorithm for subsets C_i . Every point from C belongs 542 to at most 2 sets C_i .

Apart from recursive algorithm we perform operations linear in size of |C|+M to calculate 544 the sets C_i .

Therefore it has complexity:

$$O(|C|+M) + \sum_{i=1}^{M} O(|C_{i}|k(1+\frac{1}{\delta})) = O(|C|+(1+\frac{1}{\delta})) + O((\sum_{i=1}^{M} |C_{i}|)k(1+\frac{1}{\delta})) \le O(|C|k(1+\frac{1}{\delta})).$$

Proof of Theorem 4. To construct an algorithm for this problem let us formulate some claims
 about the problem first.

Definition 10. Line is long if there are at least k+1 points from $\mathcal C$ on it.

- Claim 1. If there are more than k long lines, then C can not be covered with k segments.
- Claim 2. If there is more than k^2 points from C that do not lie on any long line, then C can not be covered with k segments.

Applying the above claims, if we have more than k long lines or more than k^2 points form \mathcal{C} that do not lie on any long line, then we answer that there is no solution of size at most k.

Otherwise, we can split C into at most k+1 sets: D, at most k^2 points that do not lie on any long line and C_i – points that lay on i-th long line. Sets C_i do not need to be disjoint.

Then for every set C_i , we can use Lemma 16 to get (k, δ) -good set A_i for C_i .

Then we have set $D \cup \bigcup A_i$ of size at most $f(k, \delta)$ for some computable function f, that if we have a solution \mathcal{R} of size at most k that covers $D \cup \bigcup A_i$, then $\mathcal{R}^{+\delta}$ covers \mathcal{C} . This is because \mathcal{R} already covers points D, they cover C_i , because they cover (k, δ) -good set A_i with at most k segments, so $\mathcal{R}^{+\delta}$ covers C_i .

After that we shrunk down size of \mathcal{C} to size of $f(k,\delta)$ for some computable function f. Then we would like to shrink down size of \mathcal{P} . For every collinear subset of D, we can choose one segment from \mathcal{P} that covers these points and have the lowest weight or decide there is no segment that cover them. There are at most $|D|^2$ different segments, because we can distinguish these collinear sets by their extreme points.

This has complexity $O(|D|^2|\mathcal{P}|)$ and produce shrunk down set \mathcal{P} of size $f(k,\delta)$ for some computable function f.

Then we can iterate over all subsets of shrunk down set \mathcal{P} and choose the set with the lowest sum of weights that cover D. This solution would have weight not larger than optimal solution for the problem without extension, because we iterate over all posibilities of covering the subset of \mathcal{C} .

574 Chapter 5

601

602

603

y coordinate.

 $\{1 \dots n^2\} \to \{1 \dots n\} \times \{1 \dots n\}.$

$_{575}$ W[1]-completeness for weighted $_{576}$ segments in 3 directions

```
TODO: some introduction
     Theorem 5. Weighted geometric set cover with segments in 3 directions is W[1]-
     hard. Consider the problem of covering a set C of points by selecting k axis-parallel or right-
     diagonal weighted segments from a set \mathcal{P} with weights w: \mathcal{P} \to \mathbb{R} with minimal weight.
     Assuming ETH, there is no algorithm for this problem with running time f(k) \cdot (|\mathcal{C}| + |\mathcal{P}|)^{o(\sqrt{k})}
    for any computable function f.
     Corollary 2. Weighted geometric set cover is W[1]-hard. Assuming ETH, there is no
     algorithm for weighted geometric set cover with running time f(k) \cdot (|\mathcal{C}| + |\mathcal{P}|)^{o(\sqrt{k})} for any
584
     computable function f.
     Proof. Trivial from Theorem 5.
                                                                                                         In order to prove theorem 5 we will show reduction from grid tiling problem.
587
    Definition 11. In the grid tiling problem we are given integers n and k, and a function
588
     f:\{1\ldots k\}\times\{1\ldots k\}\to\wp(\{1\ldots n\}\times\{1\ldots n\}) specifying the set of allowed tiles for each cell
    of a k \times k grid. The task is to find functions x, y : \{1 \dots k\} \to \{1 \dots n\} that assign numbers
    from \{1...n\} to respectively columns and rows of the grid, so that (x(i), y(j)) \in f(i, j) for
    all valid i and j, or conclude that such assignment does not exist.
     Theorem 6. Assuming ETH, there is no f(k) \cdot n^{o(\sqrt{k})}-time algorithm for grid tiling problem
    for any computable function f.
        TODO: proof from reference in literature (platypus book page 490)
595
        Let us have an instance of grid tiling problem – size of the grid k, number of colors n and
596
    function of allowed tiles f: \{1...k\} \times \{1...k\} \rightarrow \wp\{1...n\} \times \{1...n\}.
597
        TODO: nice picture of instance of grid tiling with solution
     Construction. We construct an instance of Geometric Set Cover with segments in 3 direc-
     tions with weights (\mathcal{C}, \mathcal{P}, w).
```

First let us choose any ordering of n^2 elements and denote it as bijective function order:

Define $match_v(i,j)$ and $match_h(i,j)$ as functions denoting whether two points share x or

$$match_v(i,j) \iff order(i) = \{x_i, y_i\} \land order(j) = \{x_j, y_j\} \land x_i = x_j$$

 $match_h(i,j) \iff order(i) = \{x_i, y_i\} \land order(j) = \{x_j, y_i\} \land y_i = y_j$

Points. Define points:

$$h_{i,j,t} = (i \cdot (n^2 + 1) + t, j \cdot (n^2 + 1))$$
$$v_{i,j,t} = (i \cdot (n^2 + 1), j \cdot (n^2 + 1) + t)$$

Let's define sets H and V as:

$$H = \{h_{i,j,t} : 1 \le i, j, \le k, 1 \le t \le n^2\}$$
$$V = \{v_{i,j,t} : 1 \le i, j, \le k, 1 \le t \le n^2\}$$

Let us define $\epsilon = \frac{1}{2k^2}$. For a point $p = \{x, y\}$ we define points:

$$p^{L} = \{x - \epsilon, y\},$$

$$p^{R} = \{x + \epsilon, y\},$$

$$p^{U} = \{x, y + \epsilon\},$$

$$p^{D} = \{x, y - \epsilon\}.$$

Then we define:

605

$$\mathcal{C} := H \cup \{p^L : p \in H\} \cup \{p^R : p \in H\} \cup V \cup \{p^U : p \in V\} \cup \{p^D : p \in V\}$$

Segments. Define horizontal segments.

$$\begin{split} & \mathsf{hor}_{i,j,t_1,t_2} = (h^R_{i,j,t_1}, h^L_{i+1,j,t_2}) \\ & \mathsf{ver}_{i,j,t_1,t_2} = (v^U_{i,j,t_1}, v^D_{i,j+1,t_2}) \\ & \mathsf{horBeg}_{i,t} = (h^L_{1,i,1}, h^L_{1,i,t}) \\ & \mathsf{horEnd}_{i,t} = (h^R_{k,i,t}, h^R_{k,i,n^2}) \\ & \mathsf{verBeg}_{i,t} = (v^D_{i,1,1}, v^D_{i,1,t}) \\ & \mathsf{verEnd}_{i,t} = (v^U_{i,k,t}, v^U_{i,k,n^2}) \end{split}$$

$$\begin{split} HOR &= & \{ \mathsf{hor}_{i,j,t_1,t_2} : 1 \leq i < k, 1 \leq j \leq k, 1 \leq t_1, t_2 \leq n^2, match_h(t_1,t_2) \} \\ & \cup & \{ \mathsf{horBeg}_{i,t} : 1 \leq i \leq k, 1 \leq t \leq n^2 \} \\ & \cup & \{ \mathsf{horEnd}_{i,t} : 1 \leq i \leq k, 1 \leq t \leq n^2 \} \end{split}$$

$$\begin{array}{lcl} VER & = & \{ \mathsf{ver}_{i,j,t_1,t_2} : 1 \leq i \leq k, 1 \leq j < k, 1 \leq t_1, t_2 \leq n^2, match_v(t_1,t_2) \} \\ & \cup & \{ \mathsf{verBeg}_{i,t} : 1 \leq i \leq k, 1 \leq t \leq n^2 \} \\ & \cup & \{ \mathsf{verEnd}_{i,t} : 1 \leq i \leq k, 1 \leq t \leq n^2 \} \end{array}$$

$$DIAG := \{(h_{i,j,t}, v_{i,j,t}) : 1 \le i, j \le k, 1 \le t \le n^2, order(t) \in f(i,j)\}$$

TODO: explain that these segments are in fact diagonal

606

609

$$\mathcal{P} := HOR \cup VER \cup DIAG$$

Weight function is equal to length of the segment for HOR and VER and equal to $\delta = \frac{1}{4k^4}$ for DIAG.

TODO: Put a picture of small instance like 3x3 with n=2

$$w(s) = \begin{cases} |s| & \text{if } s \in HOR \cup VER \\ \delta & \text{if } s \in DIAG \end{cases}$$

Lemma 17. If there exists solution for grid tiling, then there exists solution of instance (C, P, w) of geometric set cover with weight $2k^2(n^2 + 1) - 4k^2\epsilon - 4k(1 - \epsilon) + k^2\delta$.

Proof. If there exists a solution to the grid tiling problem x, y, then there exists a solution that covers all points

$$\{h_{i,j,t}: 1 \le i, j \le k, order(t) = (x(i), y(j))\} \cup \{v_{i,j,t}: 1 \le i, j \le k, order(t) = (x(i), y(j))\}$$

with k^2 segments from DIAG and the rest in VER or HOR. This solution has weight $2k^2(n^2+1)-4k^2\epsilon-4k(1-\epsilon)+k^2\delta$.

- Claim 3. Points in $\{p^L: p \in H\} \cup \{p^R: p \in H\}$ have to be covered with sgements from HOR.
- Points in $\{p^U: p \in V\} \cup \{p^D: p \in V\}$ have to be covered with segments from VER.
- Claim 4. For given $1 \le i, j \le n$ and any solution of an instance $(\mathcal{C}, \mathcal{P}, w)$ no two points h_{i,j,t_1}, h_{i,j,t_2} $(v_{i,j,t_1}, v_{i,j,t_2})$ for $1 \le t_1 < t_2 \le n^2$ can be not covered with segments from HOR (VER).
- 620 Proof. Proof for horizontal segments. Proof for vertical is analoguous.
- Assume point h_{i,j,t_1} is not covered with segments from HOR. Point h_{i,j,t_1}^R has to be covered with HOR from Claim 3. And every segment in HOR covering h_{i,j,t_1}^R , but not h_{i,j,t_1} covers also h_{i,j,t_2} .
- Lemma 18. For a constructed instance (C, \mathcal{P}, w) any solution has to have sum of weights from sets HOR and VER at least $W_{hv} = 2k^2(n^2+1) 4k^2\epsilon 4k(1-\epsilon)$.
- Proof. We know that for every $1 \le i, j \le k$ only one there exists only one t_x and t_y such that v_{i,j,t_x} and h_{i,j,t_y} can be not covered by segments from HOR and VER (Claim 4),
- We sum the lower bound for sum of length for horizontal/vertical lines for a single vertical line (the bound is the same for every horizontal line).
- Let us fix $1 \le i \le k$.

(1) Length between $v_{i,1,1}^D$ and v_{i,k,n^2}^U is:

$$(k(n^2+1) + n^2 + \epsilon) - ((n^2+1) + 1 - \epsilon) = k(n^2+1) - 2(1-\epsilon).$$

631 (2) For every $1 \leq j \leq k$ there exists at most one $1 \leq t \leq n^2$ such that $v_{i,j,t}$ is not covered by segments from VER (Claim 4). Its guards $v_{i,j,t}^U$ and $v_{i,j,t}^D$ have to be covered in VER (Claim 3). Therefore at most k spaces of length 2ϵ can be left not covered by segments from VER.

Therefore sum of these lower bounds for vertical and horizontal lines are:

$$2k(k(n^{2}+1)-2k\epsilon-2(1-\epsilon)) = 2k^{2}(n^{2}+1)-4k^{2}\epsilon-4k(1-\epsilon)$$

635

Lemma 19. For a constructed instance (C, P, w) for any solution R with weight equal to $2k^2(n^2+1)-4k^2\epsilon-4k(1-\epsilon)+k^2\delta$, for every $1 \le i,j \le k$ there exists such $1 \le t \le n^2$ that:

- 638 (1) $v_{i,j,t}, h_{i,j,t}$ are not covered by segments from VER or HOR;
- (2) segment $(v_{i,j,t}, h_{i,j,t})$ is in solution \mathcal{R} ;

649

650

651

- 640 (3) $order(t) \in f(i,j)$, ie. it is an allowed tile for (i,j);
- 641 (4) for every $1 \le s \le n^2$, $s \ne t$, $v_{i,j,s}$ is covered in VER;
- 642 (5) for every $1 \le s \le n^2$, $s \ne t$, $h_{i,i,s}$ is covered in HOR.
- Name the function of this t as diagonal: $\{1 \dots k\} \times \{1 \dots k\} \rightarrow \{1 \dots n^2\}$.
- Proof. At most one h_{i,j,t_x} and v_{i,j,t_y} points are covered with DIAG (Claim 4).

 Exactly one h_{i,j,t_x} and v_{i,j,t_y} points are covered with DIAG, because if one of them were not, then we would use too much weight

$$W_{hv} + 2\epsilon > 2k^2(n^2 + 1) - 4k^2\epsilon - 4k(1 - \epsilon) + k^2\delta$$

This points are covered with the same segment from DIAG, because we need to use at least k^2 of them to use exactly one DIAG segment for every pair of $1 \leq i, j \leq k$, if we used 2 segments from DIAG for one pair (i,j), then we would have used too much weight by δ . Since these points h_{i,j,t_x} and v_{i,j,t_y} are covered by segment from DIAG, therefore $t_x = t_y$.

Therefore $diagonal(i, j) = t_x = t_y$ and $order(t_x)$ is an allowed tile for (i, j) because the respective segment is in DIAG.

Lemma 20. For a constructed instance (C, P, w) for any solution of weight $2k^2(n^2 + 1) - 4k^2\epsilon - 4k(1-\epsilon) + k^2\delta$ it holds that diagonal function from Lemma 19:

- 1. for any $1 \le i < k, 1 \le j \le k$, $match_h(diagonal(i, j), diagonal(i + 1, j))$ must be true;
- 655 2. for any $1 \le i \le k, 1 \le j < k$, $match_v(diagonal(i, j), diagonal(i, j + 1))$ must be true.

```
only one segment, because otherwise it would use W_{hv} + \epsilon > 2k^2(n^2+1) - 4k^2\epsilon - 4k(1-\epsilon) + k^2\delta,
657
      therefore such solution would be too costly.
658
          Proof for vertical (2), proof for horizontal is analogous.
     Let us take any 1 \le i < k, 1 \le j \le k and name t_1 = diagonal(i, j) and t_2 = diagonal(i + 1, j). Therefore h_{i,j,t_1} and h_{i+1,j,t_2} are not covered by segments from HOR, h_{i,j,t_1}^R and h_{i+1,j,t_2}^L have to be covered by segments from HOR (Claim 3). Every segment from HOR starts
661
662
     at h_{x,y,z_1}^R segment and finishes at h_{x,y+1,z_2}^L segment for some 1 \leq x \leq k, 1 \leq z < k and
663
     1 \leq z_1, z_2 \leq n^2. Since all of the points between h_{i,j,t_1}^R and h_{i+1,j,t_2}^L are covered by segments in HOR, and there are two different segments covering these points, one of them must begin
     at h_{i,j,t_1}^R and end at h_{i,j+1,z_2}^L and the other one begin at h_{i,j,z_1}^R and end at h_{i+1,j,t_2}^L for some
     1 \le z_1, z_2 \le n^2. Therefore space between h_{i,j,z_1}^R and h_{i,j+1,z_2}^L is covered twice and is longer than \epsilon. By Lemma 18 lower bound for weight of such solution is W_{hv} + \epsilon which contradicts
667
668
     that the given solution has size 2k^2(n^2+1)-4k^2\epsilon-4k(1-\epsilon)+k^2\delta < W_{hv}+\epsilon Therefore
669
     h_{i,j,t_1}^R and h_{i+1,j,t_2}^L must be covered by one segment from HOR and these points are ends of
          h_{i,j,t_1}^R and h_{i+1,j,t_2}^L are ends of a segment from HOR, therefore match_h(t_1,t_2) must be
672
673
      Corollary 3. For a constructed instance (\mathcal{C}, \mathcal{P}, w) for any solution of weight 2k^2(n^2+1) –
674
      4k^2\epsilon - 4k(1-\epsilon) + k^2\delta it holds that diagonal function from Lemma 19:
         1. for any 1 \le i, j \le k, match_h(diagonal(1, j), diagonal(i, j)) must be true;
676
         2. for any 1 \le i, j \le k, match_v(diagonal(i, 1), diagonal(i, j)) must be true.
677
      Proof. Simple inductive proof based on Lemma 20.
                                                                                                                              678
      Lemma 21. If there exists solution of instance (\mathcal{C}, \mathcal{P}, w) with weight at most 2k^2(n^2+1) –
679
      4k^2\epsilon - 4k(1-\epsilon) + k^2\delta, then there exists a solution for grid tiling instance.
680
      Proof. Take diagonal function from Lemma 19.
681
          To define x funtion for every 1 \le i \le k as x(i) = x_i where (x_i, a) = order(v_{i,1}) and y
682
     function for every 1 \le i \le k as y(i) = y_i where (b, y_i) = order(h_{1,i})
683
          To prove that it is a correct solution for grid tiling, we need to prove that for every
      1 \le i, j \le k \ (x(i), y(j)) is in allowed tiles set f(i, j).
685
          Let us take any 1 \le i, j \le k, By Corollary 3 we know that match_h(diagonal(1, j), diagonal(i, j))
686
      and match_v(diagonal(i, 1), diagonal(i, j)) are true. Therefore order(diagonal(i, j)) = (x(i), y(i)).
687
      By Lemma 19 we know that order(diagonal(i,j)) is in f(i,j). Therefore (x(i),y(i)) is in
688
      f(i,j).
689
                                                                                                                              690
     Proof of Theorem 5. Based on Lemmas 17 and 21 this is true.
```

Proof. Every space between points in H and V covered by HOR/VER has to be covered by

656

692 Chapter 6

700

701

702

706

710

Geometric Set Cover with lines

$_{694}$ 6.1. Lines parallel to one of the axis

When \mathcal{R} consists only of lines parallel to one of the axis, the problem can be solved in polynomial time.

We create bipartial graph G with node for every line on the input split into sets: H – horizontal lines and V – vertical lines. If any two lines cover the same point from C, then we add edge between them.

Of course there will be no edges between nodes inside H, because all of them are pararell and if they share one point, they are the same lines. Similar argument for V. So the graph is bipartial.

Now Geometric Set Cover can be solved with Vertex Cover on graph G. Since Vertex Cover (even in weighted setting) on bipartial graphs can be solved in polynomial time.

Short note for myself just to remember how to this in polynomial time:

Non-weighted setting - Konig theorem + max matching

Weighted setting - Min cut in graph of $\neg A$ or $\neg B$ (edges directed from V to H)

6.2. FPT for arbitrary lines

You can find this is Platypus book. We will show FPT kernel of size at most k^2 .

(Maybe we need to reduce lines with one point/points with one line).

For every line if there is more than k points on it, you have to take it. At the end, if there is more than k^2 points, return NO. Otherwise there is no more than k^4 lines.

In weighted settings among the same lines with different weights you leave the cheapest one and use the same algorithm.

6.3. APX-completeness for arbitrary lines

We will show a reduction from Vertex Cover problem. Let's take an instance of the Vertex Cover problem for graph G. We will create a set of |V(G)| pairwise non-pararell lines, such that no three of them share a common point.

Then for every edge in $(v, w) \in E(G)$ we put a point on crossing of lines for vertices v and w. They are not pararell, so there exists exactly one such point and any other line don't cover this point (any three of them don't cross in the same point).

Solution of Geometric Set Cover for this instance would yield a sound solution of Vertex Cover for graph G. For every point (edge) we need to choose at least one of lines (vertices) v or w to cover this point.

Vertex Cover for arbitrary graph is APX-complete, so this problem in also APX-complete.

6.4. 2-approximation for arbitrary lines

Vertex Cover has an easy 2-approximation algorithm, but here very many lines can cross through the same point, so we can do d-approximation, where d is the biggest number of lines crossing through the same point. So for set where any 3 lines don't cross in the same point it yields 2-approximation.

The problematic cases are where through all points cross at least k points and all lines have at least k points on them. It can be created by casting k-grid in k-D space on 2D space.

Greedy algorithm yields $\log |\mathcal{R}|$ -approximation, but I have example for this for bipartial graph and reduction with taking all lines crossing through some point (if there are no more than k) would solve this case. So maybe it works.

Unfortunaly I haven't done this:(

I can link some papers telling it's hard to do.

⁷³⁸ 6.5. Connection with general set cover

Problem with finite set of lines with more dimensions is equivalent to problem in 2D, because we can project lines on the plane which is not perpendicular to any plane created by pairs of (point from C, line from P).

Of course every two lines have at most one common point, so is every family of sets that have at most one point in common equivalent to some geometric set cover with lines?

No, because of Desargues's theorem. Have to write down exactly what configuration is banned.

746 Chapter 7

Geometric Set Cover with polygons

7.1. State of the art

751

752

753

Covering points with weighted discs admits PTAS [Li and Jin, 2015] and with fat polygons with δ -extensions with unit weights admits EPTAS [Har-Peled and Lee, 2009].

Although with thin objects, even if we allow δ -expansion, the Set Cover with rectangles is APX-complete (for $\delta = 1/2$), it follows from APX-completeness for segments with δ -expansion in Section 3.1.

Covering points with squares is W[1]-hard [Marx, 2005]. It can be proven that assuming SETH, there is no $f(k) \cdot (|\mathcal{C}| + |\mathcal{P}|)^{k-\epsilon}$ time algorithm for any computable function f and $\epsilon > 0$ that decides if there are k polygons in \mathcal{P} that together cover \mathcal{C} , Theorem 1.9 in [Marx and Pilipczuk, 2015].

⁷⁵⁸ Chapter 8

Conclusions

We do not know FPT for axis-parallel segments without δ -extensions.

Bibliography

- [Har-Peled and Lee, 2009] Har-Peled, S. and Lee, M. (2009). Weighted geometric set cover
 problems revisited. Journal of Computational Geometry, 3.
- ⁷⁶⁴ [Håstad, 2001] Håstad, J. (2001). Some optimal inapproximability results. J. ACM, ⁷⁶⁵ 48(4):798-859.
- ⁷⁶⁶ [Li and Jin, 2015] Li, J. and Jin, Y. (2015). A PTAS for the weighted unit disk cover problem. CoRR, abs/1502.04918.
- [Marx, 2005] Marx, D. (2005). Efficient approximation schemes for geometric problems? In
 Brodal, G. S. and Leonardi, S., editors, Algorithms ESA 2005, pages 448-459, Berlin,
 Heidelberg. Springer Berlin Heidelberg.
- ⁷⁷¹ [Marx and Pilipczuk, 2015] Marx, D. and Pilipczuk, M. (2015). Optimal parameterized algorithms for planar facility location problems using voronoi diagrams. *CoRR*, abs/1504.05476.