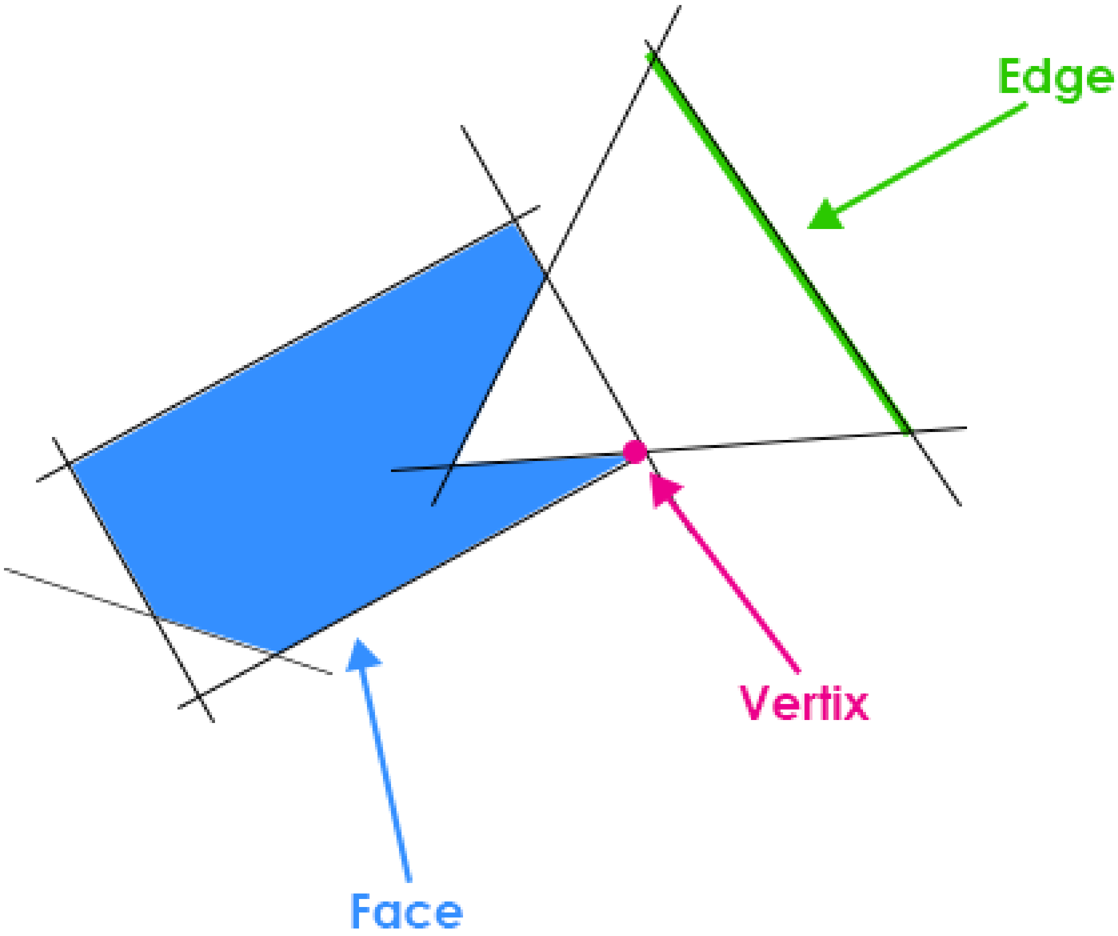
SAMPLING AND THE MOMENT TECHNIQUE

By Sveta Oksen

Overview

- Vertical decomposition
  - Construction
  - Running time analysis
- The bounded moments theorem
  - General settings
  - The sampling model
  - The exponential decay lemma
- Applications
  - Proving the guess (of vertical decomposition)
  - (1/r)-cutting

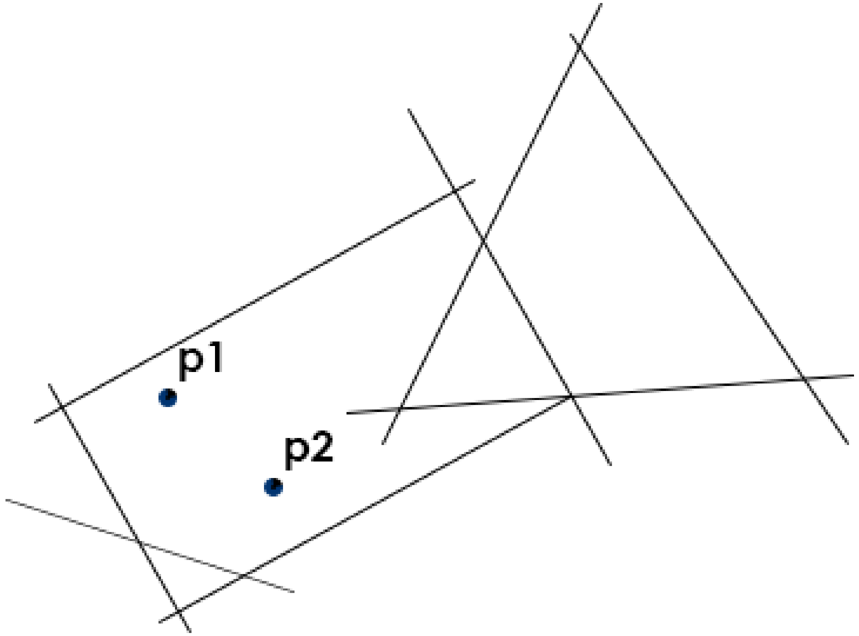
Basic definitions



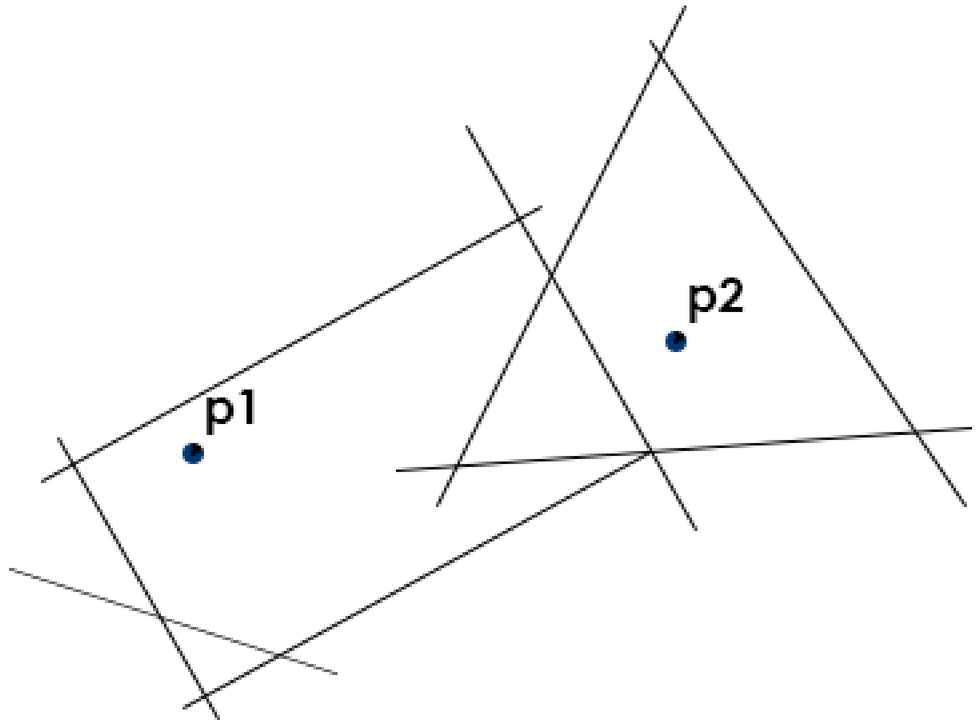
Vertical decomposition - motivation

We would like to build a data structure, which will make it easy to answer the following questions –

Are p1 and p2 in the same face?



Can one traverse from p1 to p2 without crossing any segment?



Vertical decomposition - definitions

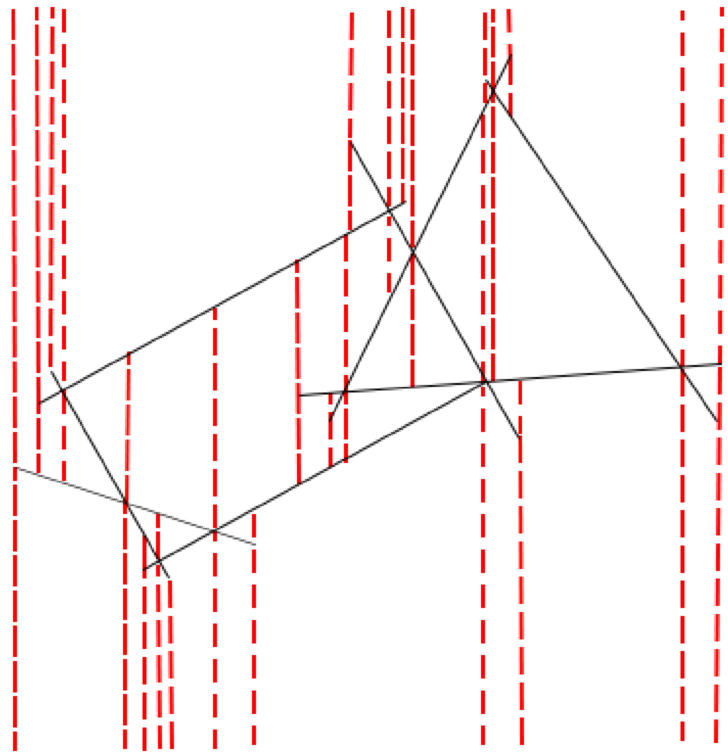
- S – set of segments (lines)

- A(S) – plane **arrangement** - Decomposition of R^2 by the segments into edges, vertices and faces.

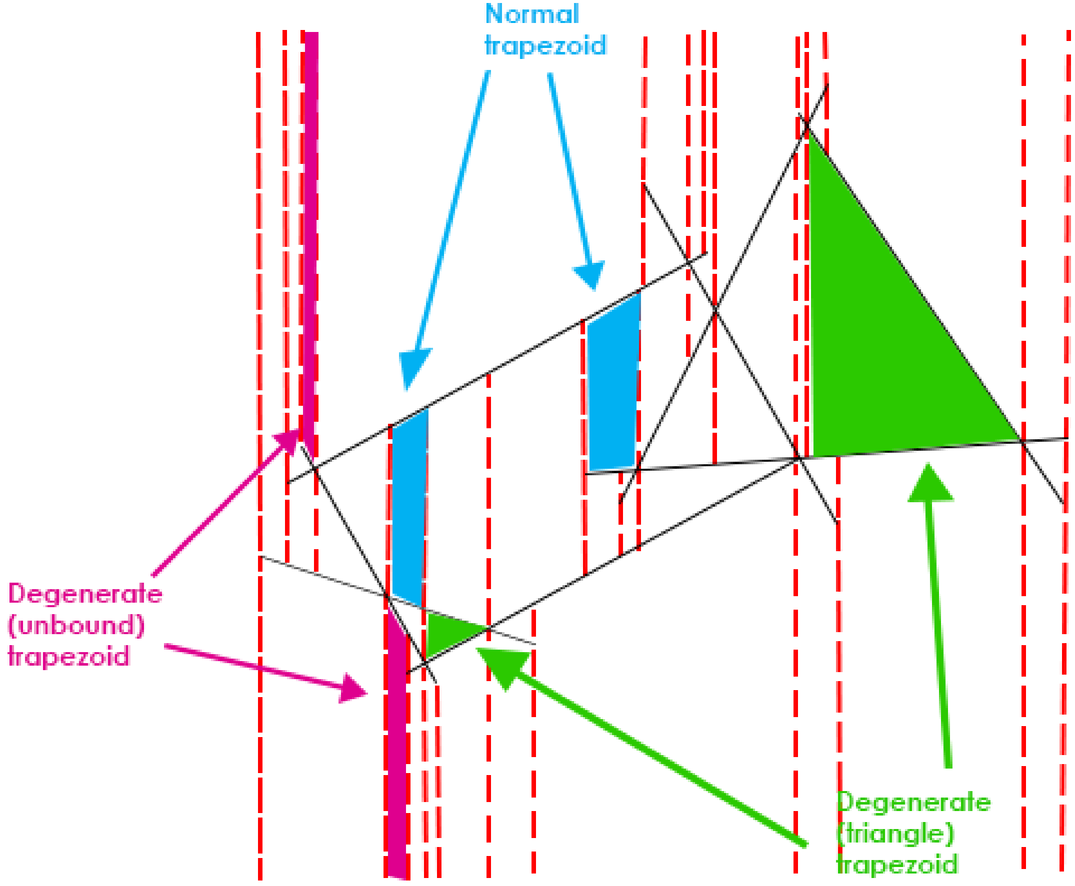
- A'(S) – the **data structure** that stores the arrangement A(S).

Vertical decomposition - algorithm

Draw a vertical line through each vertex in our arrangement (including endpoints), until it hits a segment or until infinity. The result is called **vertical decomposition**.



The vertical decomposition breaks the plane into **trapezoids**. Some of them might be degenerate.

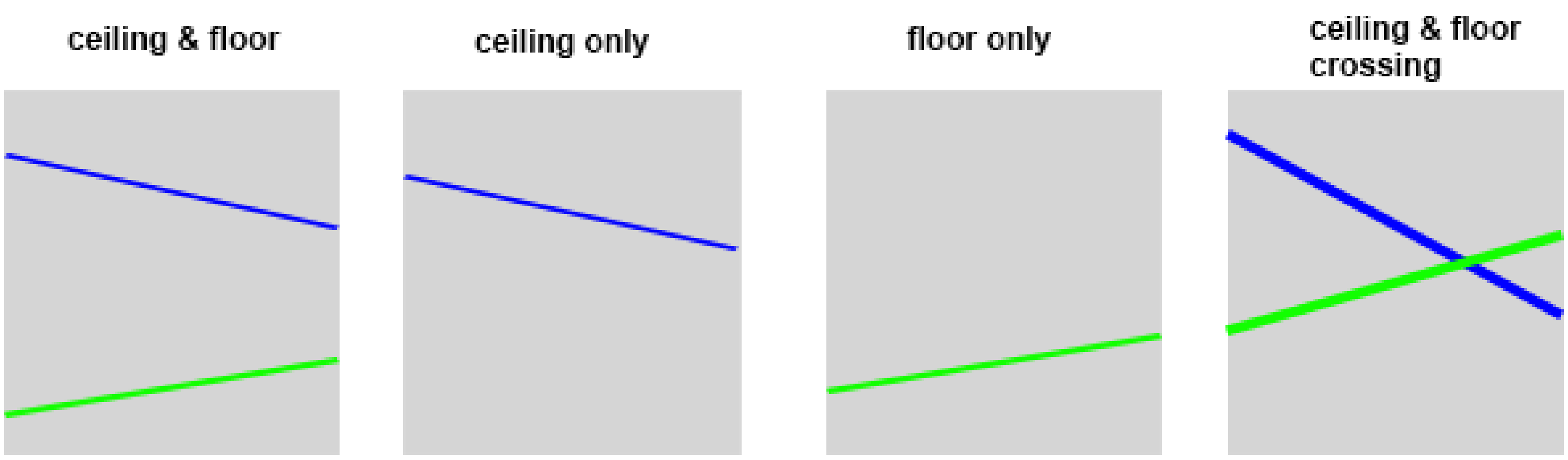


Trapezoids structure types

- Every trapezoid must have a ceiling or a floor (or both).

- If the ceiling touches the floor – we get a degenerate triangle trapezoid.

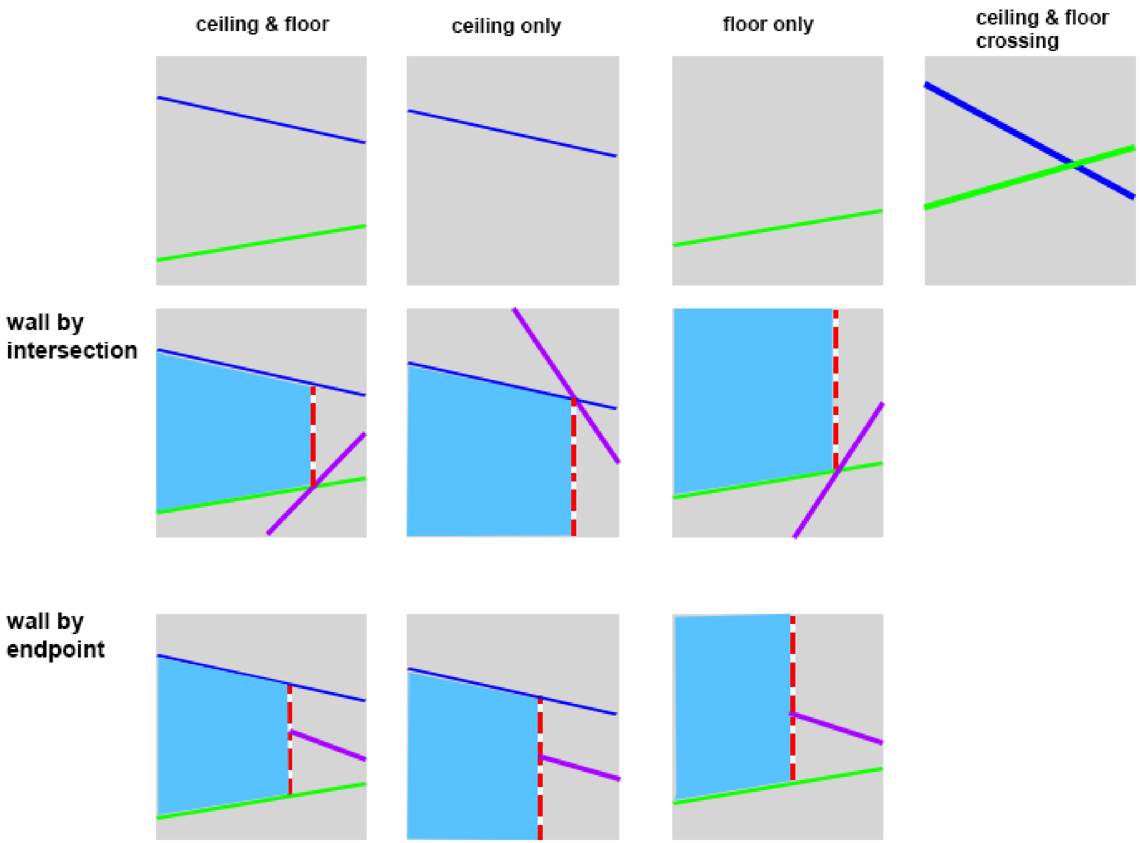
- If there is no ceiling or no floor, we get a degenerate unbounded trapezoid.



Trapezoids structure types - cont

- The left and right walls of the trapezoids are defined by one of the following –
  - A segment which crosses the ceiling or the floor
  - An endpoint of a segment
- In the case of a triangle, one wall is missing

Therefore, every trapezoid is defined by up to 4 segments.



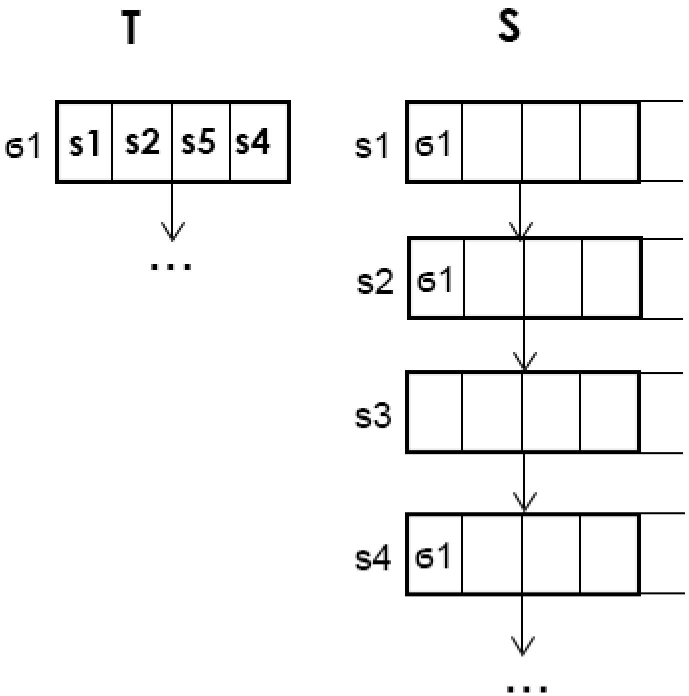
Data structure for A(S)

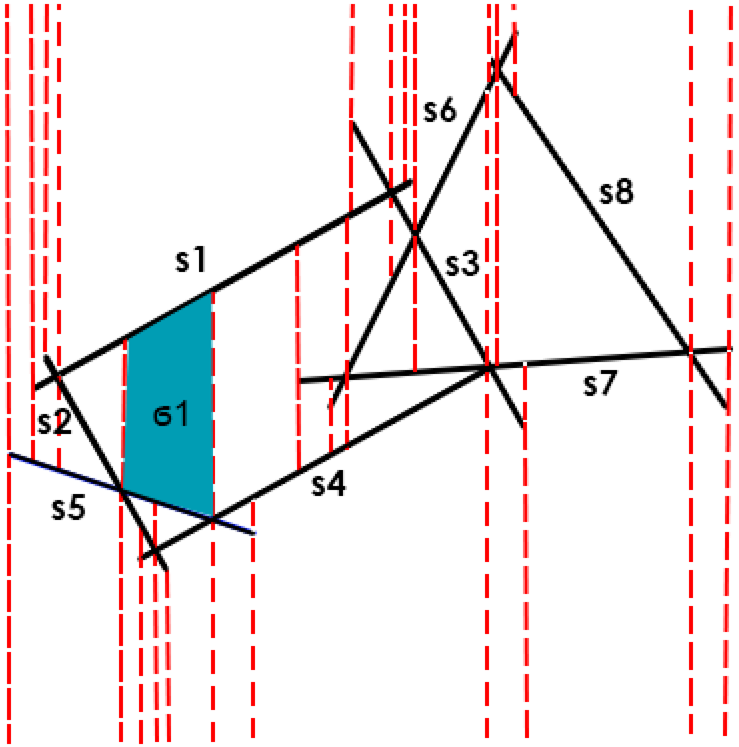
The data structure that represents A(S) will consists of –
T – a linked list of all trapezoids
S – a linked list of all segments

Each cell in T maintains up to 4 pointers to S, which represent the segments which define it.

We will call the data structure A'(S).

A'(S) - Example





Constructing A'(S) - Algorithm

- We take the group of segments S and apply a random permutation on it. Denote it as -

$$S =< s_1, s_2, \dots, s_n >$$

- Let S_i be the prefix of length i of S. $S_i =< s_1, s_2, \dots, s_i >$

- Before step 1, T and S are empty.

- On every step i we will add one segment S_i to the data structure.

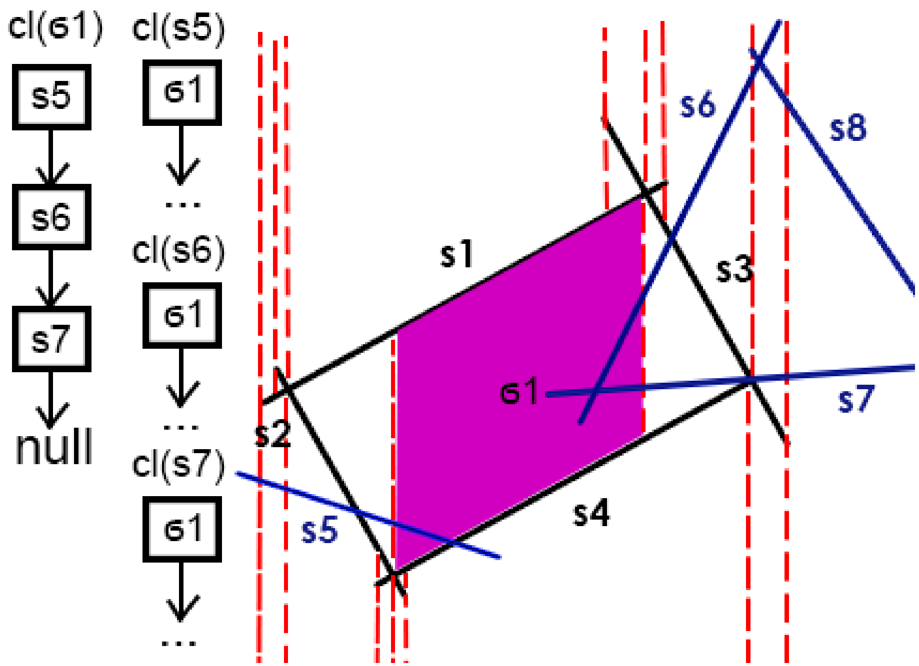
- $A'(S_i)$ - the data structure created after adding S_i

- $A'(S_n)$ – the final desire structure

Algorithm - continued

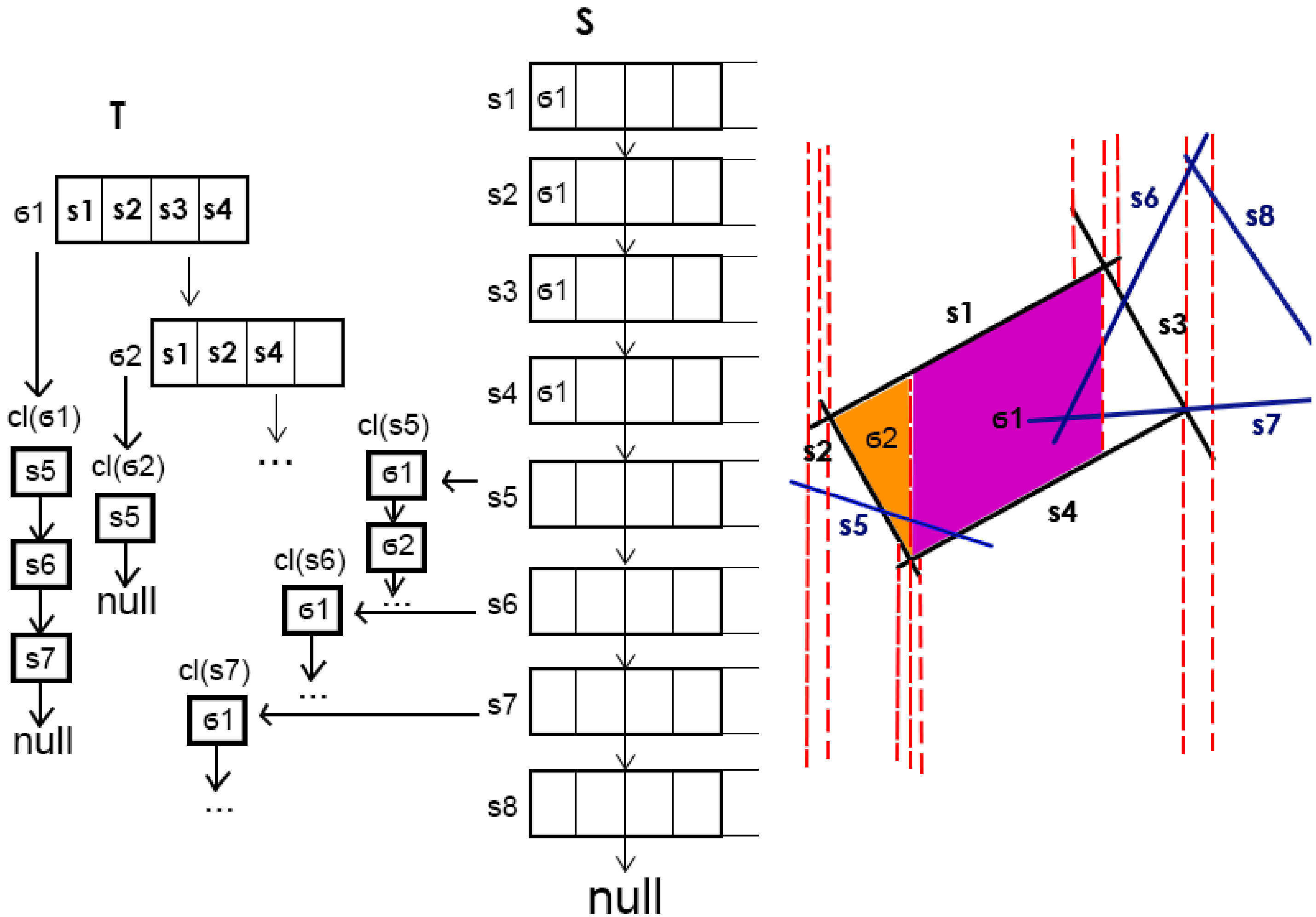
On each step we also maintain the following lists-

- $cl(\sigma)$ – contains the segments which intersect the trapezoid σ. We call it the **conflict list** of σ.
- $cl(s_i)$ – contains the trapezoids which intersect the segment s_i. We call it the **conflict list** of s_i.



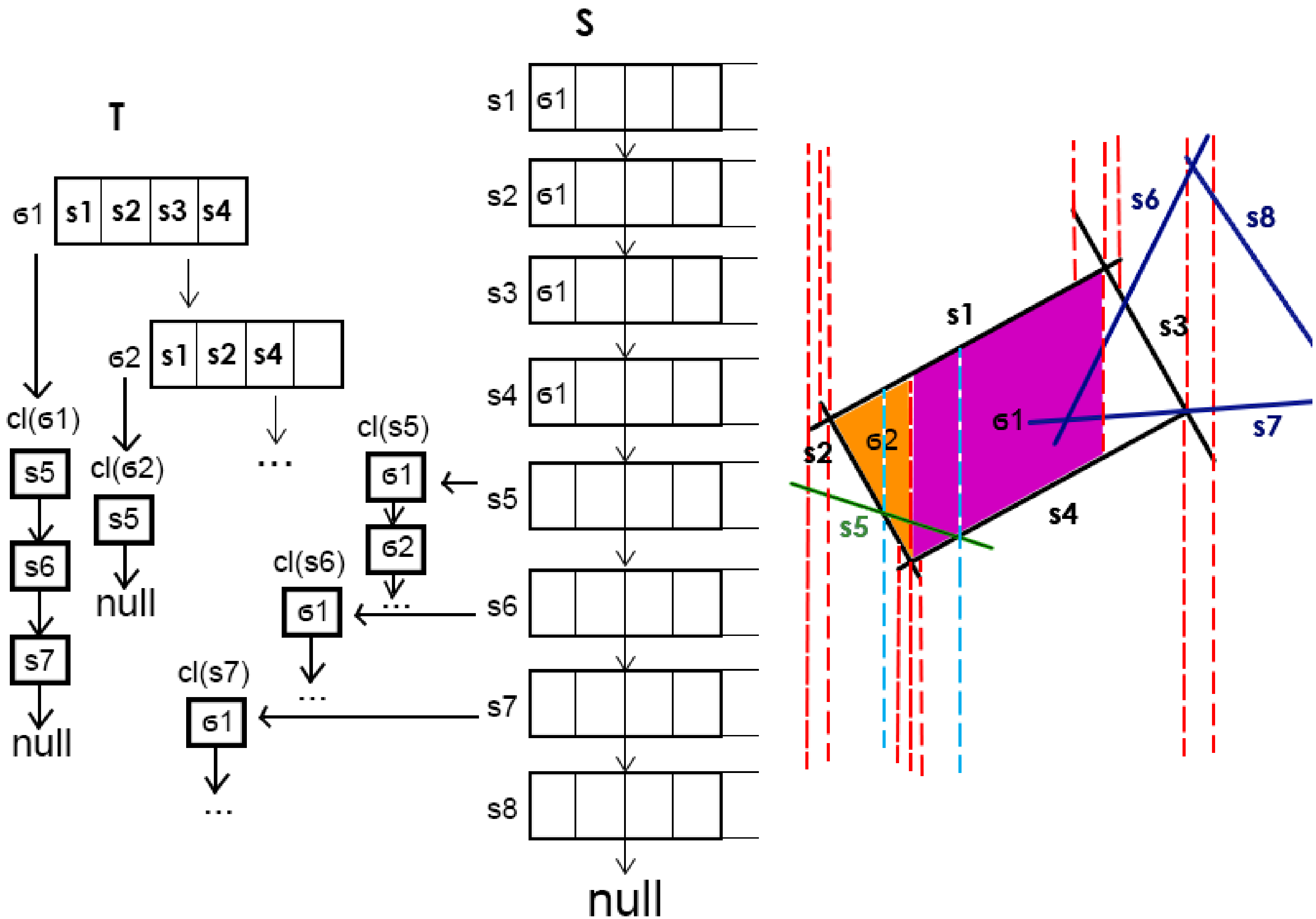
Example: After adding 4 segments

We've added s_1, s_2, s_3, s_4 and still have s_5, s_6, s_7, s_8 to add.



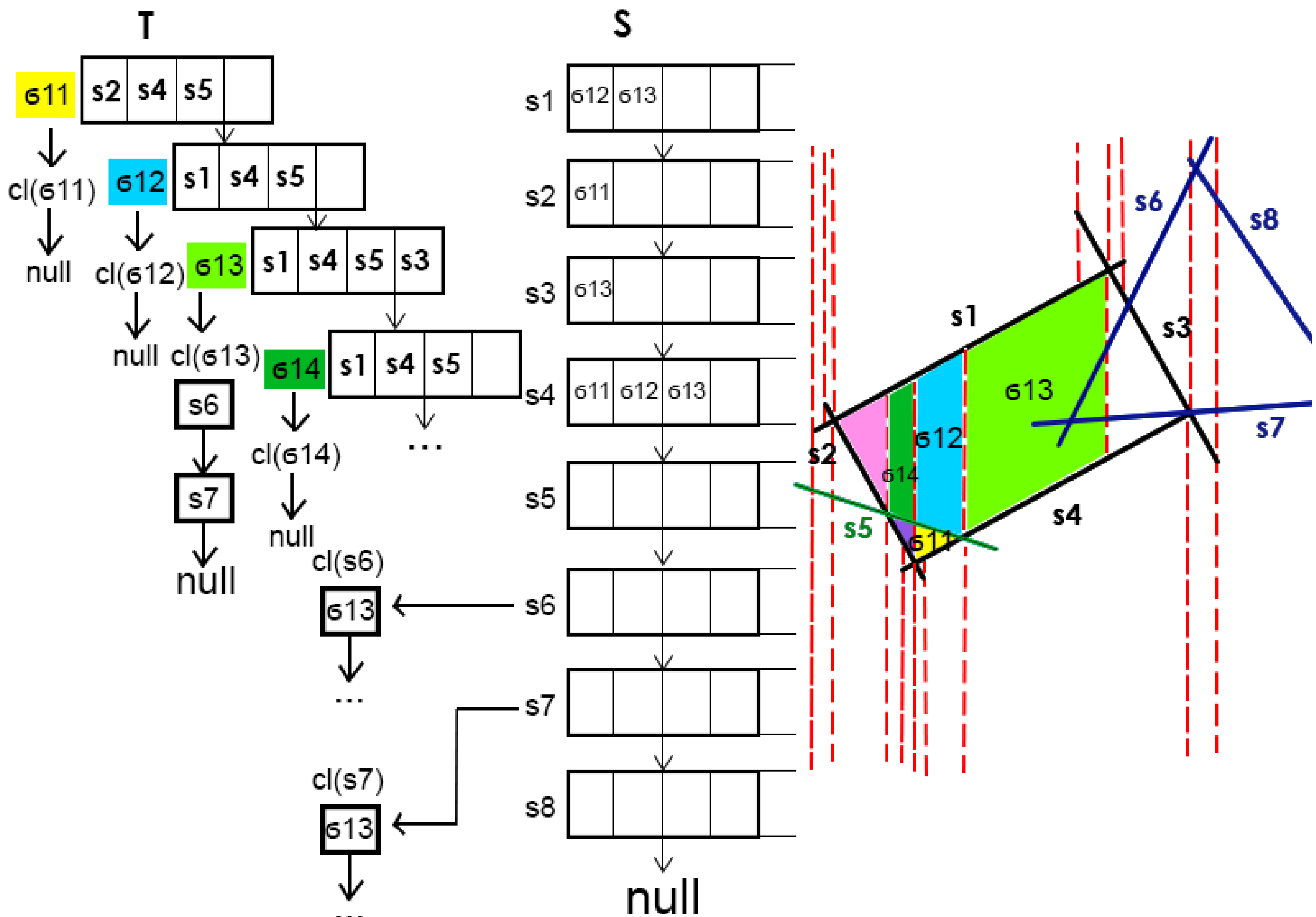
Example: Adding s_5

We want to add s_5. We go through the conflict list of s_5 and split every trapezoid in this list. There will be up to 4 new trapezoids created for each entry in the conflict list.

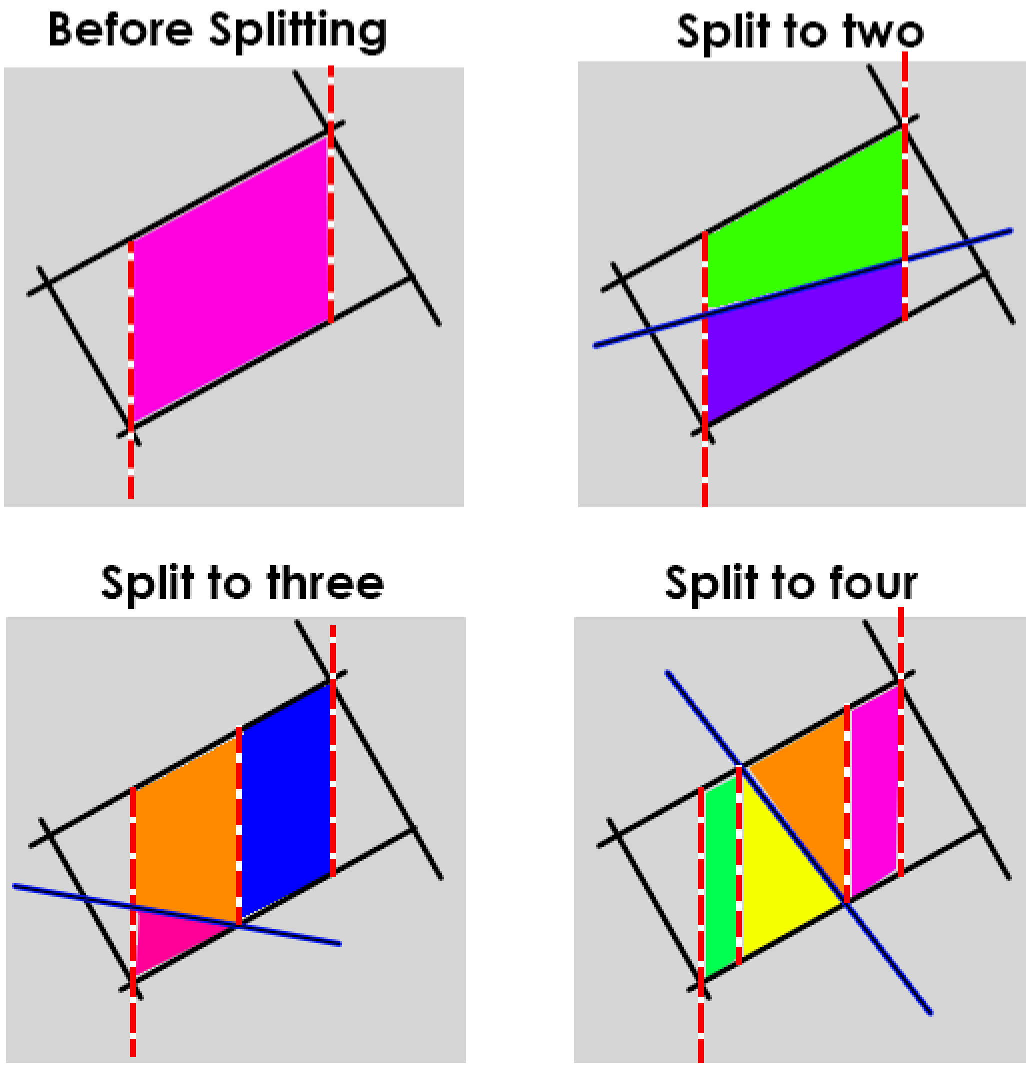


Example: Adding s_5 - continuation

When creating the new trapezoids, we construct their conflict lists out of the old trapezoid conflict list.

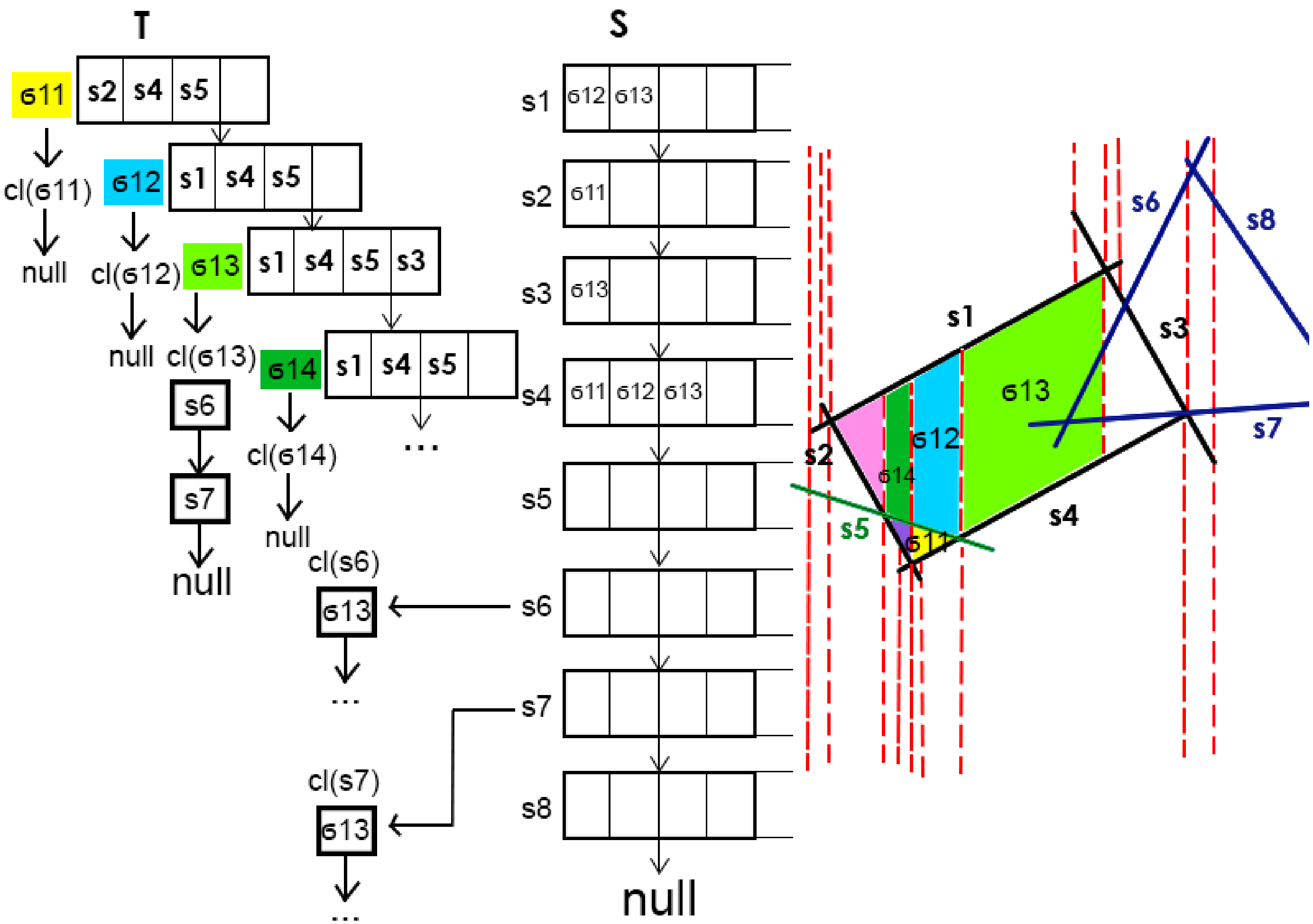


Possible splits of the trapezoid



Example: Adding s_5 - continuation

After creating the new trapezoids, some of them might be invalid. I.e if we would do a full decomposition, we would not get those trapezoids.
In the example below, σ_{12} and σ_{14} are invalid.



Merging invalid trapezoids

- To fix the problem of the invalid trapezoids, we need to perform the “merge” operation.

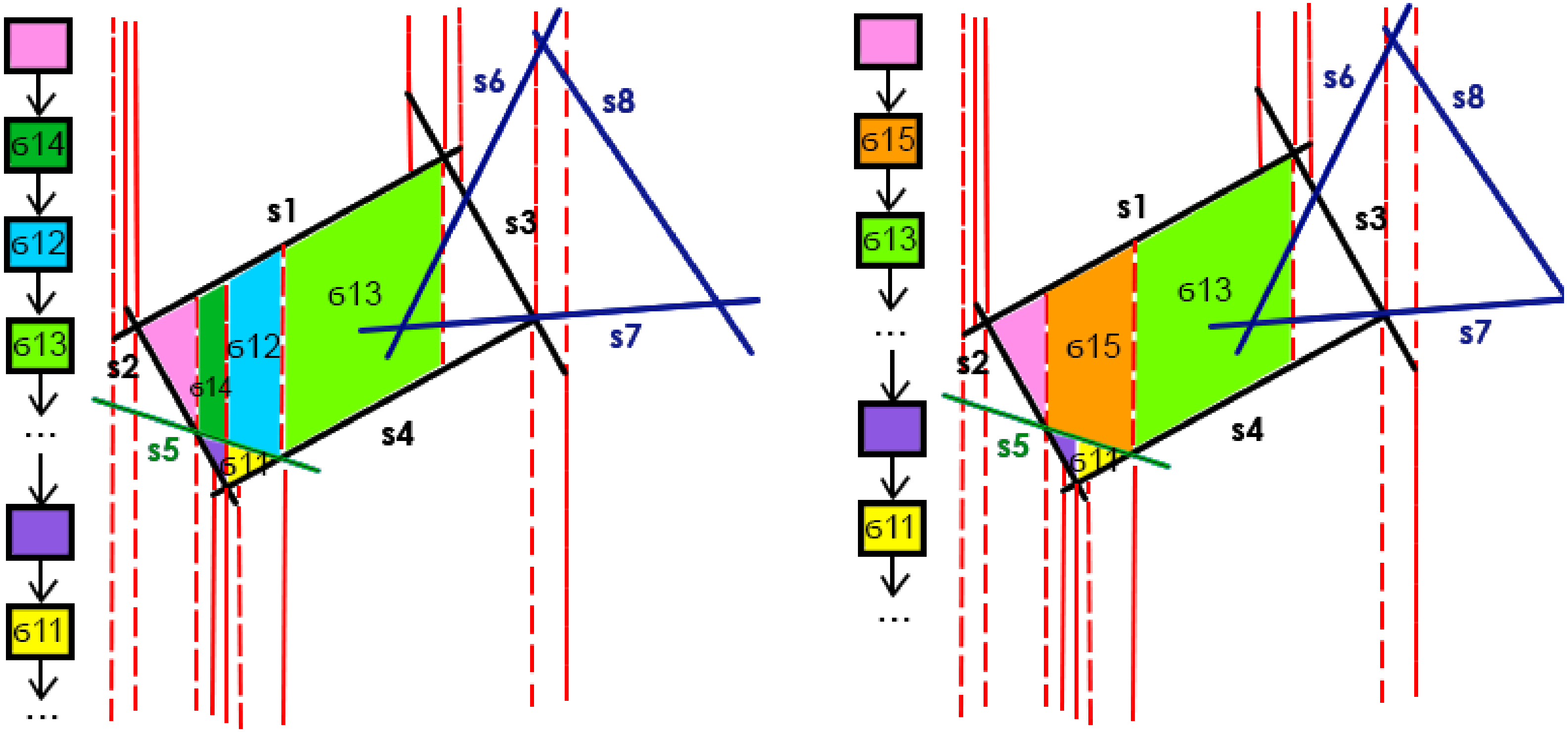
- Every invalid trapezoid has an adjacent trapezoid which has the same ceiling and floor.

- If we merge all the trapezoids with same ceiling and floor, we get rid of the invalid trapezoids and get a valid vertical decomposition.

- We maintain a list of adjacent trapezoids.

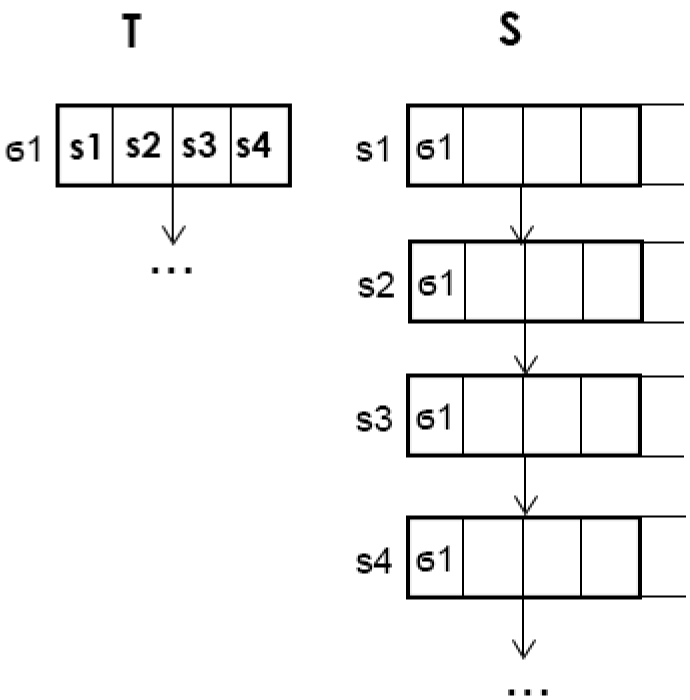
- After creation of new trapezoids, we go through adjacent trapezoids and merge them if they have same ceiling and floor.

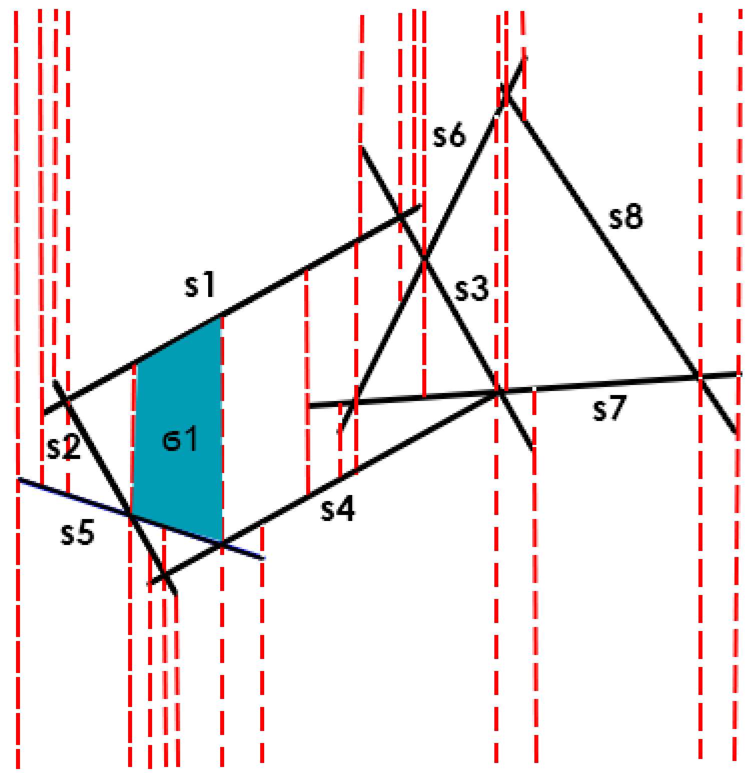
Merging invalid trapezoids – cont.



Example: A’(S)

This way we will proceed until we add all the segments. In the end, all the conflict lists will be empty (because the segments which already added can’t be in conflict list).





Running time

Claim 1: the amortized running time of constructing of $A'(S_i)$ is proportional to the size of the conflict lists of the trapezoids in $A'(S_i)\backslash A'(S_{i-1})$.

Proof:
Every time we create new trapezoids, we break an existing trapezoid. When we construct new trapezoids out of existing one, we do three things:

- Vertical decomposition of new trapezoids – for this we go through all 5 segments (4 old and one new) intersections. – up to 5^2 actions - O(1) per trapezoid

- Merging of new trapezoids – we go through all new trapezoids once (up to 4 new trapezoids from each old one) and merge them – O(1) per trapezoid

- We create the conflict list of the new trapezoids out of the old ones.

Running time – proof cont.

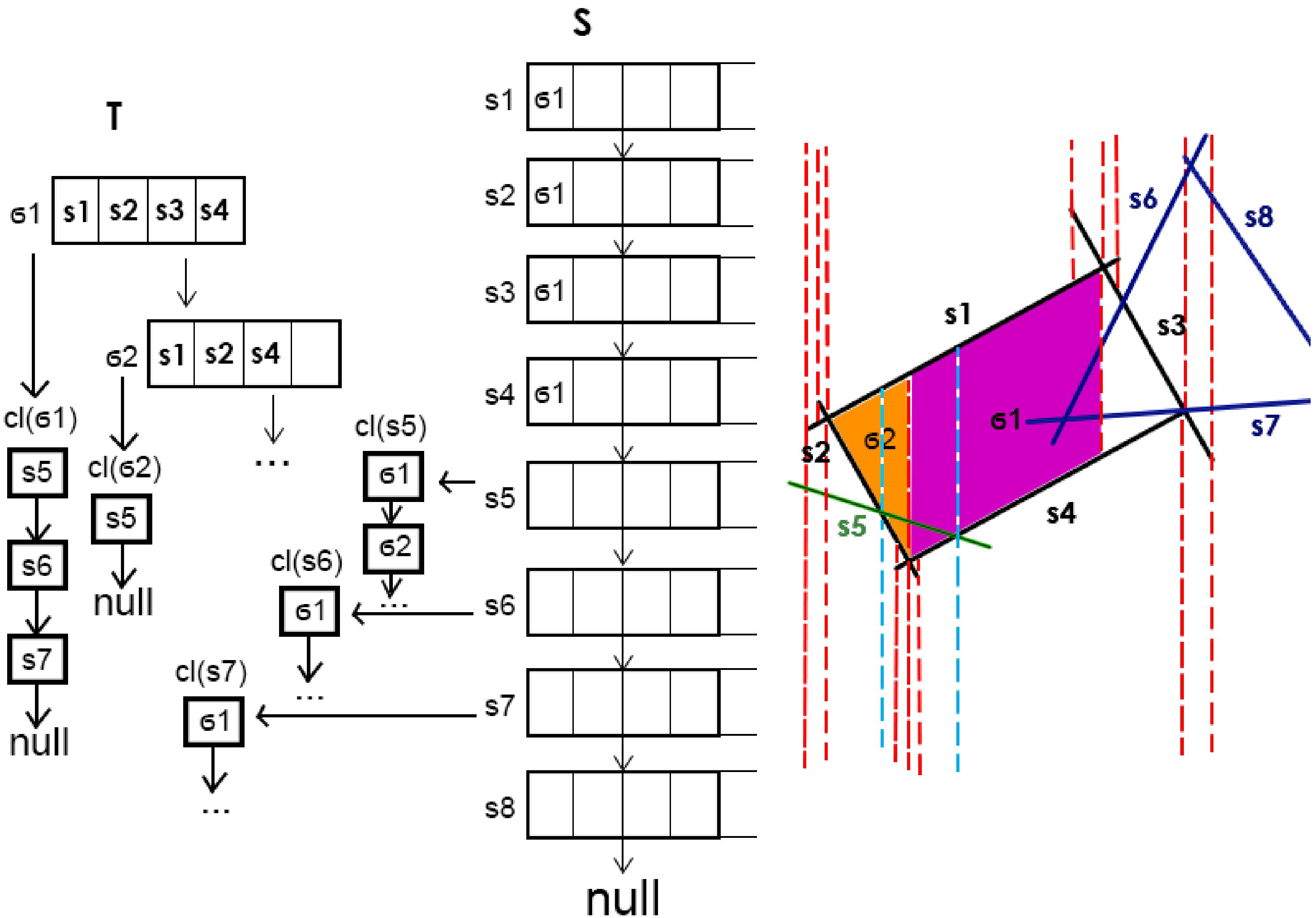
- Each old conflict list is used by at most 4 new conflict lists

- Each new conflict list is created out of the “ruins” of an old. So old destroyed lists pay for creation of new ones.

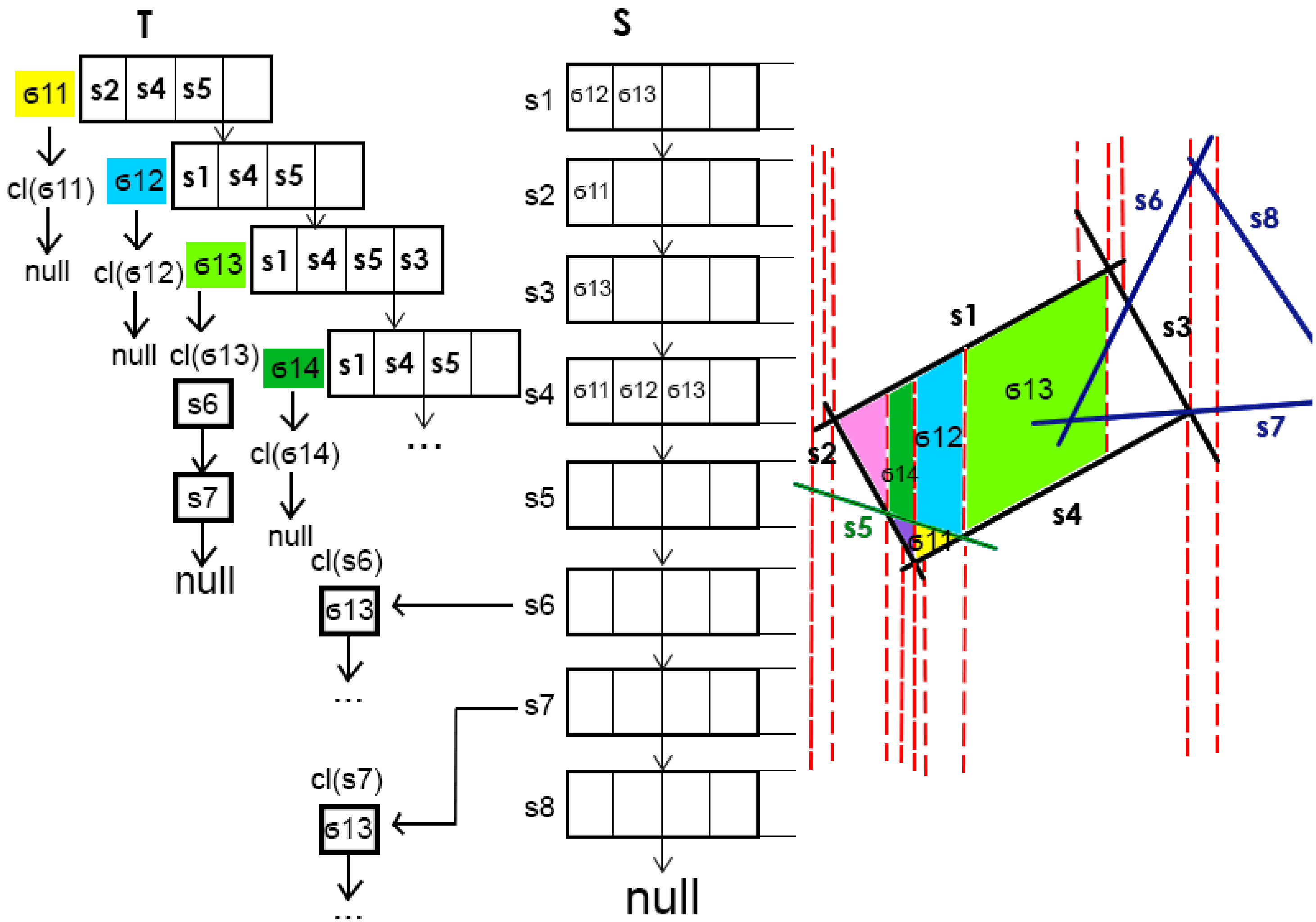
Therefore we can charge every time a conflict list is created. And the charges at step i are proportional to the size of the conflict lists of the trapezoids created at step i.

■

Running time – illustration



Running time – illustration



Running time of the algorithm

Therefore it is enough to bound the expected size of the conflict lists created in the i^{th} iteration. (Which is the size of the conflict lists in $A'(S_i)\backslash A'(S_{i-1})$)

We will analyze the running time in two steps:

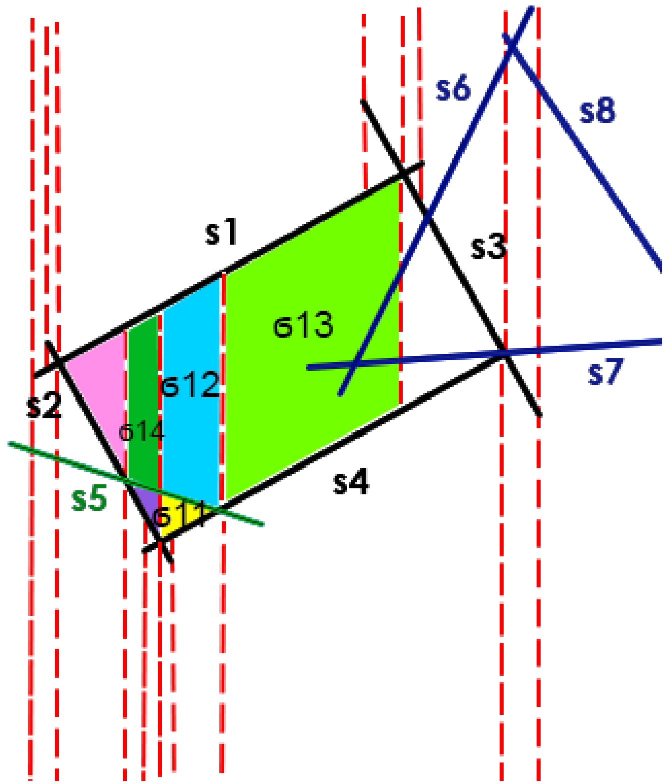
1) Find the expected size of $A'(S_i)$

2) Do **backward analysis** to compute the expected size of

$$A'(S_i)\backslash A'(S_{i-1})$$

Step 1 – the size of $A'(S_i)$

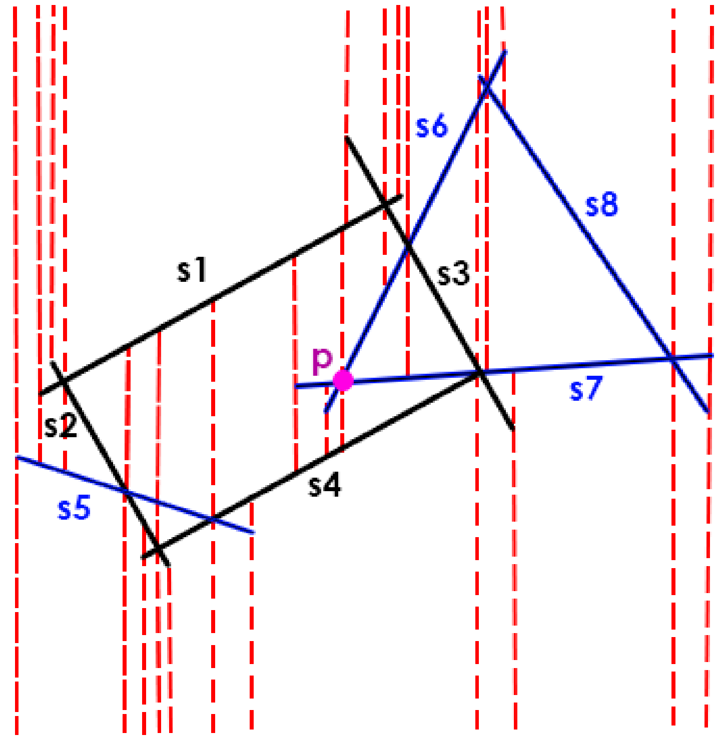
Lemma 1: Let S be a set of segments with k intersection points. Let S_i be the first i segments in the random permutation of S .The expected size of $A'(S_i)$ (i.e the number of trapezoids in $A'(S_i)$), denoted by $\tau(i)$, is $O\left(i + k\left(\frac{i}{n}\right)^2\right)$.



Proof: Consider an intersection point $p = s \cap s'$, where $s, s' \in S$. The probability that p is present in $A'(S_i)$ is the probability that both s and s' are in S_i.

$S_4 =< s_1, s_2, s_3, s_4 >$

$p = s_6 \cap s_7$



$$\alpha = \frac{\binom{n-2}{i-2}}{\binom{n}{i}} = \frac{(n-2)!}{(i-2)!\,(n-i)!} \cdot \frac{i!\,(n-i)!}{n!} = \frac{i(i-1)}{n(n-1)}.$$

Proof continuation

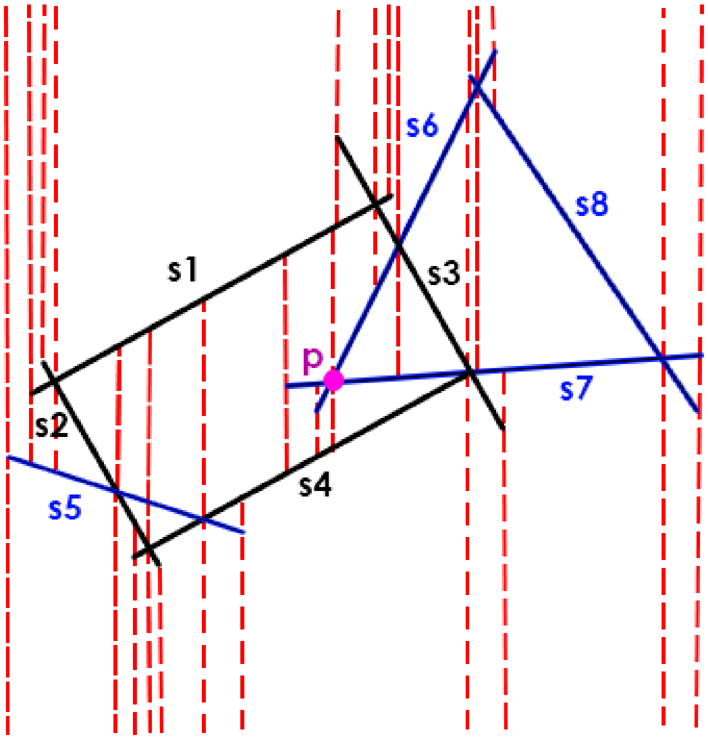
Now we define an indicator variable X_p which is 1 if the two defining segments of p are in S_i. 0 otherwise.

From before we have $E[X_p] = \alpha$.

Therefore, the expected number of the intersections in $A(S_i)$ is

$$\mathbf{E}\left[\sum_{p \in V} X_p\right] = \sum_{p \in V} \mathbf{E}[X_p] = \sum_{p \in V} \alpha = k\alpha,$$

where V is the set of k intersection points of A(S).

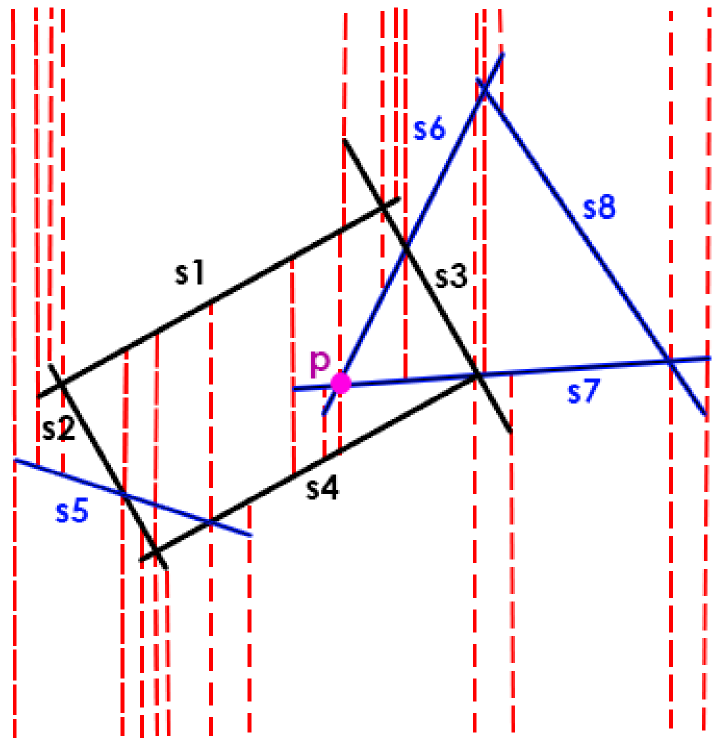


Proof continuation

Also, every end point of segment s of S_i contributes 2 endpoints to $A'(S_i)$
Thus, we get that the expected number of vertices in $A'(S_i)$ is

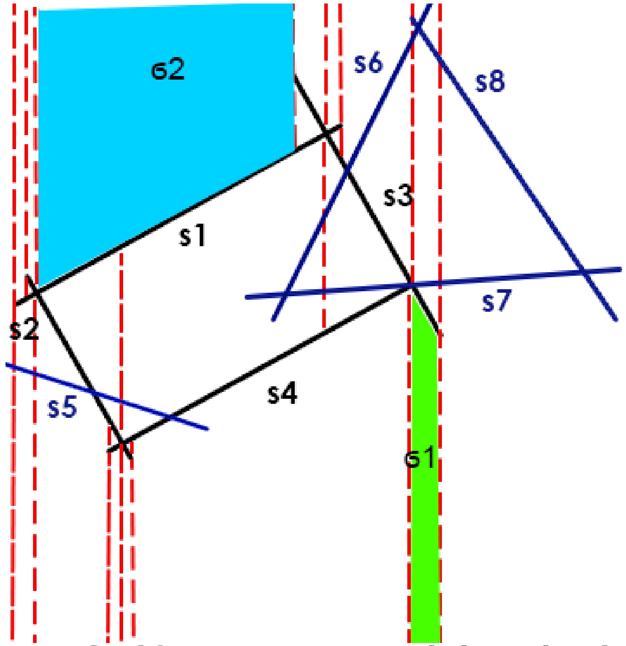
$$2\mathrm{i} + \mathrm{k}\alpha = 2i + \frac{i(i-1)}{n(n-1)}k$$

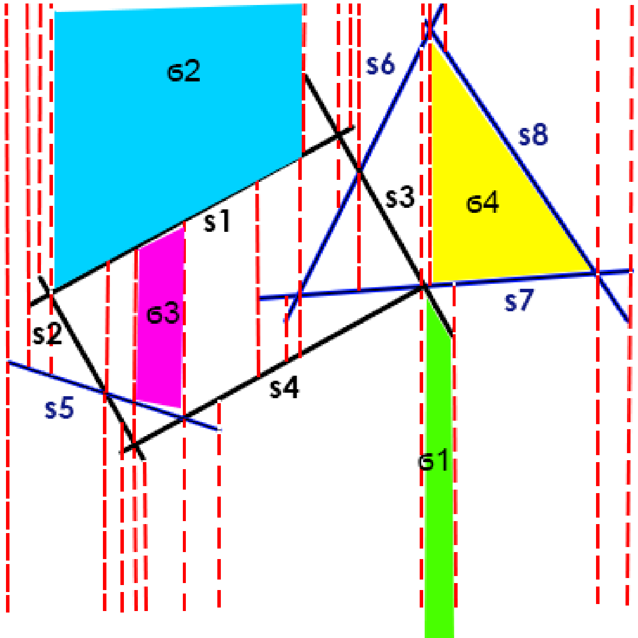
Since the number of trapezoids in $A'(S_i)$ is proportional to number of vertices in $A(S_i)$, we conclude that the expected number of trapezoids in $A'(S_i)$ is $O\left(i + k\left(\frac{i}{n}\right)^2\right)$ as desired. ■



Step 2 – backward analysis

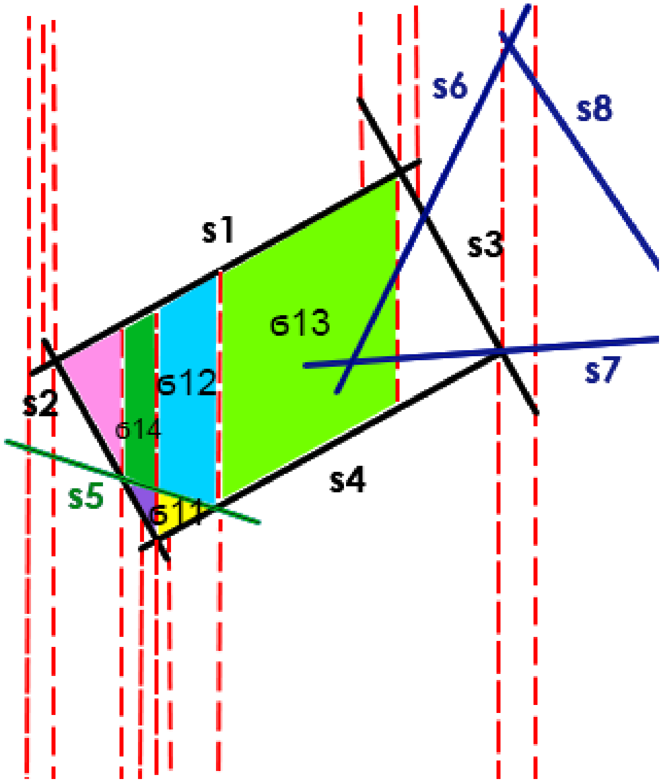
Claim 2: $\Pr[\sigma \in (A'(S_i) \backslash A'(S_{i-1}))] \leq \frac{4}{i}$





Proof: if a trapezoid σ is in $A'(S_i)$ but not in $A'(S_{i-1})$, that means that at least one of its defining segments s_i was added the last in S_i. The probability of a segment s_i to be the last in S_i is $\frac{1}{i}$. Therefore, the probability that at least one of the segments was added at S_i is at most $\frac{4}{i}$. ■

Running time analysis – summing up



Definitions:

- $B_i = A'(S_i)$
- $C_i = |cl(B_i \setminus B_{i-1})|$ - size of conflict lists introduced in step i.
- $W_i = |cl(B_i)|$ - total size of conflict lists in $A'(S_i)$
- By claim 2: $\Pr[\sigma \in (B_i \setminus B_{i-1})] \leq \frac{4}{i}$
- $W_i = \sum_{\sigma \in A'(S_i)} |cl(\sigma)|$

Therefore,

$$E[C_i | B_i] = \sum_{\sigma \in A'(S_i)} \Pr[\sigma \in (B_i \setminus B_{i-1})] \cdot |cl(\sigma)| \leq \sum_{\sigma \in A'(S_i)} \frac{4}{i} |cl(\sigma)| \leq \frac{4}{i} W_i$$

Intuition: The expected size of conflict lists added in step i is getting lower as i grows: the trapezoids become lighter and lighter.

Summing up - continuation

- $B_i = A'(S_i)$
- $W_i = \sum_{\sigma \in A'(S_i)} |cl(\sigma)|$
- By lemma 1: $|B_i| = O\left(i + k\left(\frac{i}{n}\right)^2\right)$
- **Guess:** the average size of the conflict list of a trapezoid of B_i is $O(\frac{n}{i})$.

Therefore

$$E(W_i) = |B_i| \cdot O\left(\frac{n}{i}\right) = O\left(i + k\left(\frac{i}{n}\right)^2\right) \cdot O\left(\frac{n}{i}\right) =$$
$$O\left(n + k\left(\frac{i}{n}\right)\right)$$

Running time analysis - continuation

- C_i - size of conflict lists introduced in step i.
- $E(W_i) = O\left(n + k\left(\frac{i}{n}\right)\right)$

Therefore

$$\mathbf{E}[C_i] = \mathbf{E}\left[\mathbf{E}\left[C_i \mid \mathcal{B}_i\right]\right] \le \mathbf{E}\left[\frac{4}{i} W_i\right] = \frac{4}{i}\mathbf{E}[W_i] = O\left(\frac{4}{i}\left(n + \frac{ki}{n}\right)\right) = O\left(\frac{n}{i} + \frac{k}{n}\right)$$

And finally, the overall expected running time of the algorithm is

$$\mathbf{E}\left[\sum_{i=1}^{n} C_i\right] = \sum_{i=1}^{n} O\left(\frac{n}{i} + \frac{k}{n}\right) = O(n \log n + k)$$

Intuition for the guess

We will now try to get some intuition for the guess from before –

On average, the size of the conflict list of a trapezoid of B_i *is about* $O(\frac{n}{i})$

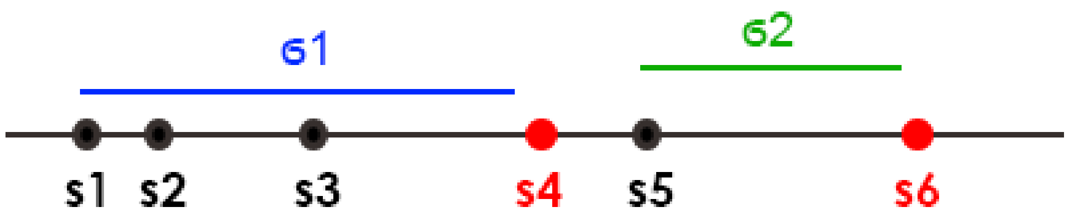
Intuition: In S_i we pick i out of n segments ≈ pick each segment with probability of $\frac{i}{n}$.

If $|cl(\sigma)| \gg \frac{n}{i}$, we expect to pick $\approx \frac{i}{n} \cdot |cl(\sigma)| \gg 1$ segments from it.
But we picked none!

Intuition cont.

Let's look on the one dimensional case.

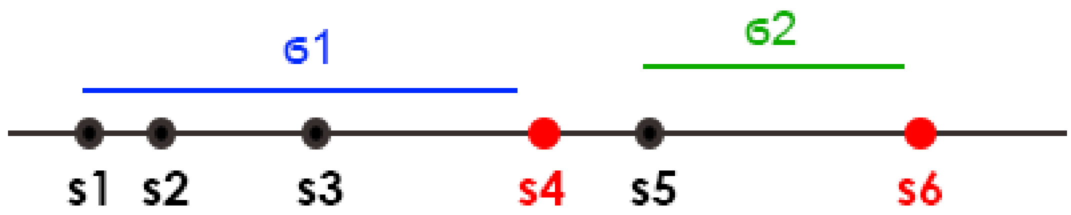
In this case we have a line instead of plane, interval I is a trapezoid, points s_i are the segments.



We choose i points $\mathrm{S_i} = \{s_{k_1}, \dots, s_{k_i}\}$, out of $S = \{s_1, \dots, s_n\}$ at random. Our trapezoids will be the biggest intervals we can draw that don't contain any $s \in \{s_{k_1}, \dots, s_{k_i}\}$ in their interior.

In the resulting decomposition, the number of the points which appear inside the intervals is the size of the conflict list of the trapezoid.

Intuition – cont.



We are interested in the expected size of conflict list of σ_i.
If we fix a point s and got to the right of it, while the probability of any point to be chosen to S_i is $\frac{i}{n}$, the random variable which is the number of the points in the interval (excluding the chosen points), acts like a geometric variable with probability $\frac{i}{n}$.
Therefore, the expected size of the conflict list of the trapezoid (ie number of points which fall into the interval) is $O\left(\frac{n}{i}\right)$.

Proof of the guess - preparation

As a main part of the proof, we first need to introduce and prove the "**Bounded moments Theorem**".

The Bounded moments theorem will give us some bound on the expected size of the conflict lists in step i.

To prove this theorem, we will need to introduce the following:

- **The sampling model** - how we sample the segments

- **General settings** – a framework for the analysis, more general than segments and trapezoids.

- **The exponential decay lemma** – a lemma which tells that the number of trapezoids with big conflict lists is dropping exponentially

The sampling model

In algorithms when we want to build a group of r randomly chosen objects out of n, we will usually implement it by first permuting the group and taking its r prefix.

For analysis, this sampling model is much harder to calculate than the model where we pick every object with probability r/n. We will use the “easier” model in our analysis.

General Settings

- Let S be a set of objects
- For a subset $R \subseteq S$, we define a collection of **regions** F(R).

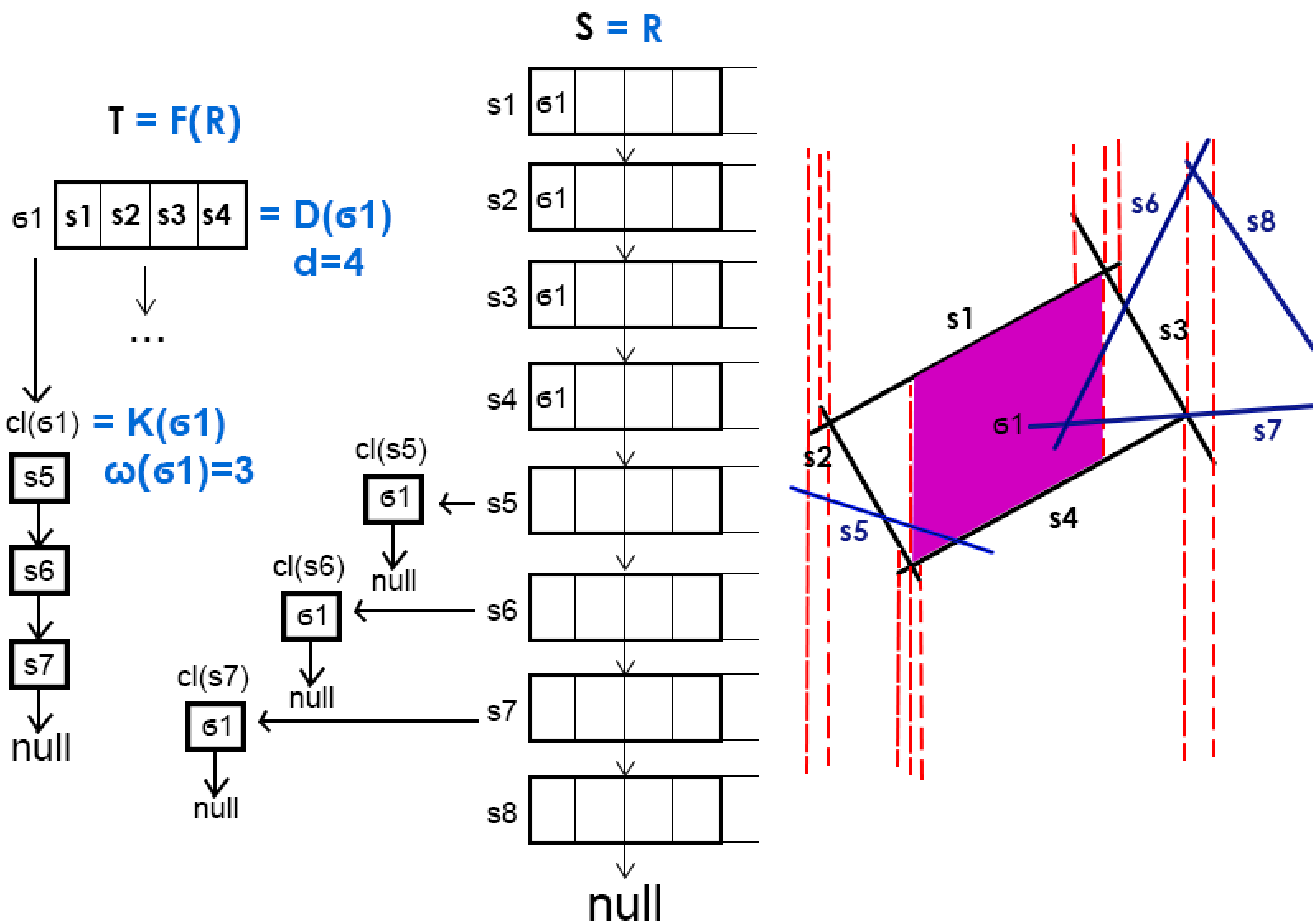
For the case of vertical decomposition, S will be the set of segments and F(R) will be the set of trapezoids.

- Let T be the set of all possible regions, defined by the subsets of S.

$$\mathcal{T} = \mathcal{T}(S) = \bigcup_{R \subseteq S} \mathcal{F}(R)$$

General Settings - continuation

- $D(\sigma)$ – is the **defining set** of σ. - *In the case of vertical decomposition $D(\sigma)$ is the set of segments which define σ.*
  - We assume that for every $\sigma \in T$, $|\boldsymbol{D}(\boldsymbol{\sigma})| \leq \boldsymbol{d}$ for a small constant d. *- In the case of vertical decomposition, each trapezoid is defined by at most 4 segments, therefore d=4.*
- $K(\sigma)$ - is the **stopping set** of σ. – In the case of vertical decomposition $K(\sigma)$ is the set of segments of S intersecting the interior of the trapezoid σ (its conflict list).
- $\omega(\sigma)$ - is the weight of σ. Defined to be $|K(\sigma)|$.


S = R
T = F(R)
σ1
s1
s2
s3
s4
= D(σ1)
d=4
...
cl(σ1)
= K(σ1)
ω(σ1)=3
s5
s6
s7
null
cl(s5)
σ1
null
cl(s6)
σ1
null
cl(s7)
σ1
null
s1 σ1
s2 σ1
s3 σ1
s4 σ1
s5
s6
s7
s8
null
s1
s2
s3
s4
s5
s6
s7
s8
σ1

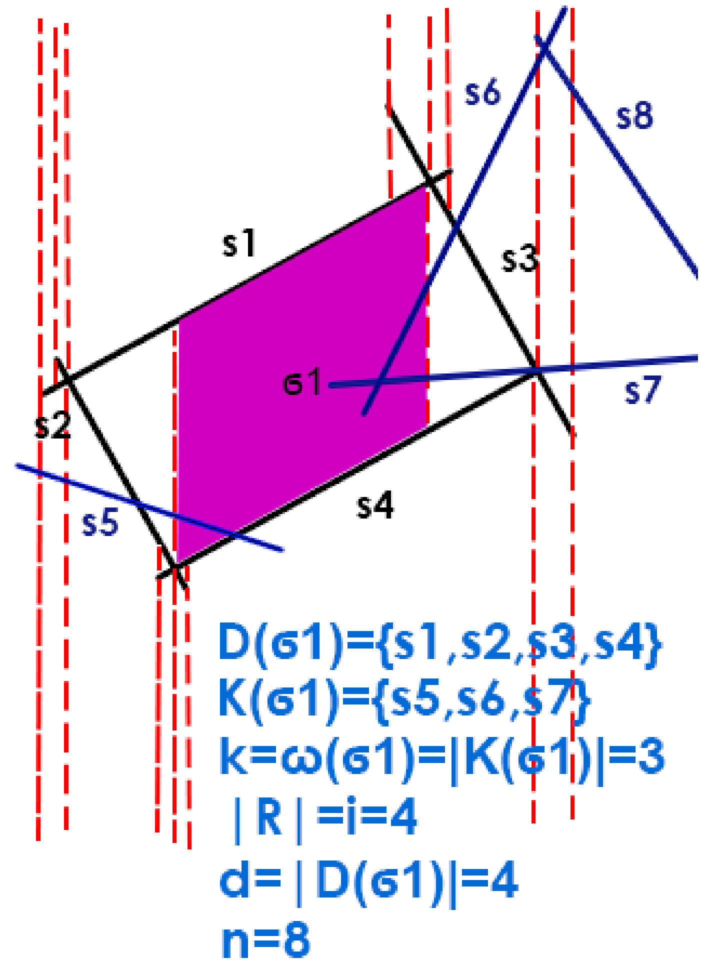
Axioms

Let S, F(R), $D(\sigma)$ and $K(\sigma)$ be such that for any subset $R \subseteq S$, the set F(R) satisfies the following axioms:

1) For any $\sigma \in F(R)$, we have $D(\sigma) \subseteq R$ and $R \cap K(\sigma) = \emptyset$.

   I.e: choose all defining segments. Don't choose any conflicting/stopping one.

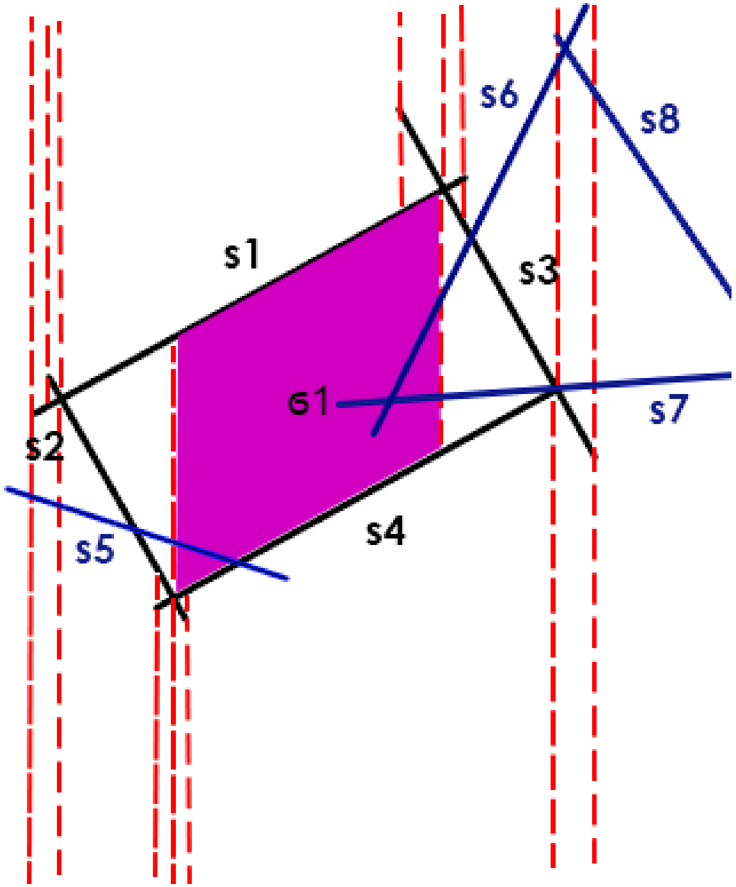
2) If $D(\sigma) \subseteq R$ and $K(\sigma) \cap R = \emptyset$, then $\sigma \in F(R)$



Probability of region to be created

Let S be a set complying with the axioms.

We denote by $\boldsymbol{\rho_{r,n}(d,k)}$ the probability that a region $\sigma \in T$ appears in F(R).
Where its defining set is of size **d**, its stopping set is of size **k**,
R is random sample of size **r** from S, and **n**=|S|.

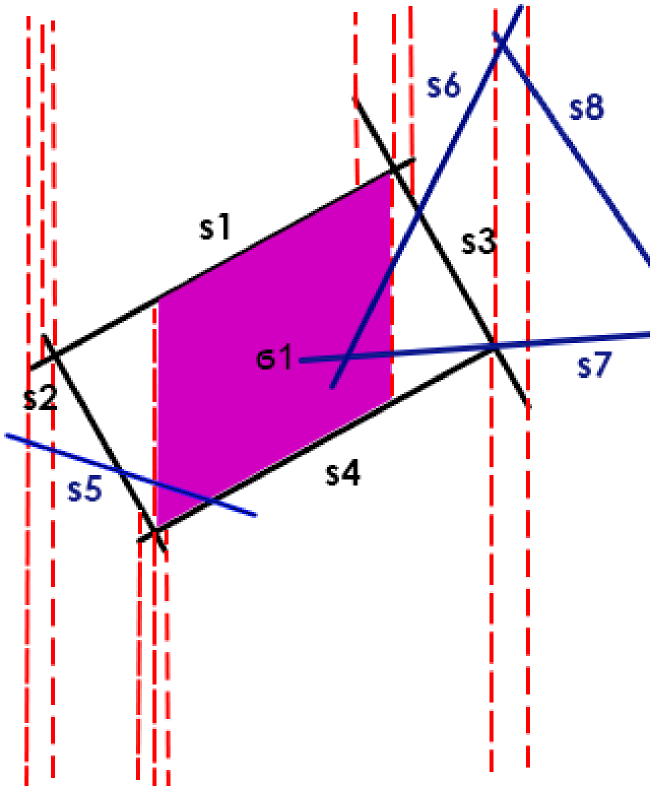


Claim 3:

$$\rho_{r,n}(d,k) \approx \left(1 - \frac{r}{n}\right)^k \left(\frac{r}{n}\right)^d$$

Proof of the claim

Claim 3: $\rho_{r,n}(d,k) \approx \left(1-\frac{r}{n}\right)^k \left(\frac{r}{n}\right)^d$



Proof in simpler sampling model:
If we assume that every segment is picked with the probability r/n, then the probability that the defining segments are chosen and that the stopping segments

aren't is indeed $\rho_{r,n}(d,k) \approx \left(1-\frac{r}{n}\right)^k \left(\frac{r}{n}\right)^d$

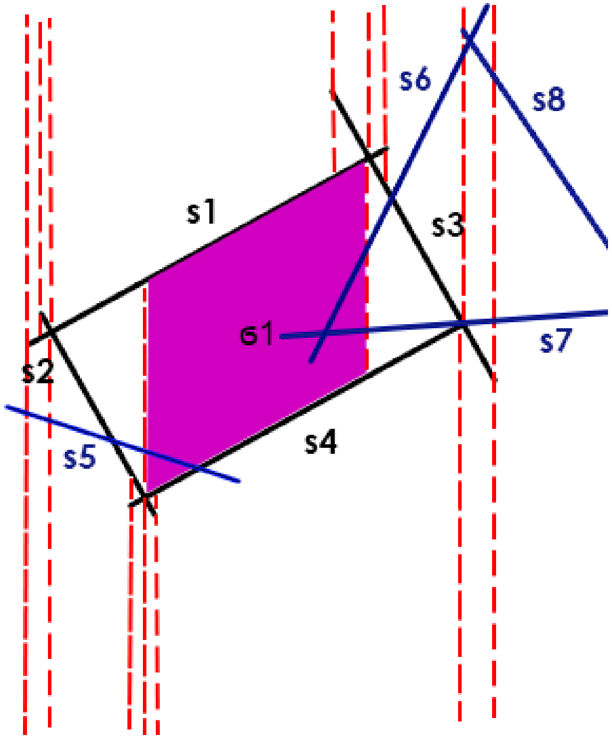
The exponential decay lemma

- S – set of objects
- $r \leq n$
- $1 \leq t \leq r/d$, where
  $d = \max_{\sigma \in T(S)} |D(\sigma)|$
- S comply to the axioms
- $Ef(\mathrm{r}) = \mathrm{E}[|F(\mathrm{R})|]$
- $\sigma \in F(R)$ is t-heavy if $\omega(\sigma) \geq t(\frac{n}{r})$
- $Ef_{\geq t}(\mathrm{r}) = \mathrm{E}[|\mathrm{F}_{\geq \mathrm{t}}(\mathrm{R})|]$

Then $$\mathbf{E}f_{\geq t}(r) = O\left(t^d \exp(-t/2)\, \mathbf{E}f(r)\right).$$

We will prove the lemma in steps.

The exponential decay intuition



- Consider R to be a random sample of size r from S without repetitions.
- A region $\sigma \in F(R)$ is t-heavy if $\omega(\sigma) \geq t\left(\frac{n}{r}\right)$
- $F_{\geq t}(R)$ - all t-heavy regions of F(R)

Intuition: the probability of creating a t-heavy trapezoid drops exponentially in t

- Indeed

$$\rho_{r,n}(d, t(n/r)) \approx \left(1 - \frac{r}{n}\right)^{t(n/r)} \left(\frac{r}{n}\right)^d \approx \exp(-t) \cdot \left(\frac{r}{n}\right)^d \approx \exp(-t+1) \cdot \left(1 - \frac{r}{n}\right)^{n/r} \left(\frac{r}{n}\right)^d$$

$$\approx \exp(-t+1) \cdot \rho_{r,n}(d, n/r)$$

The exponential decay - proof

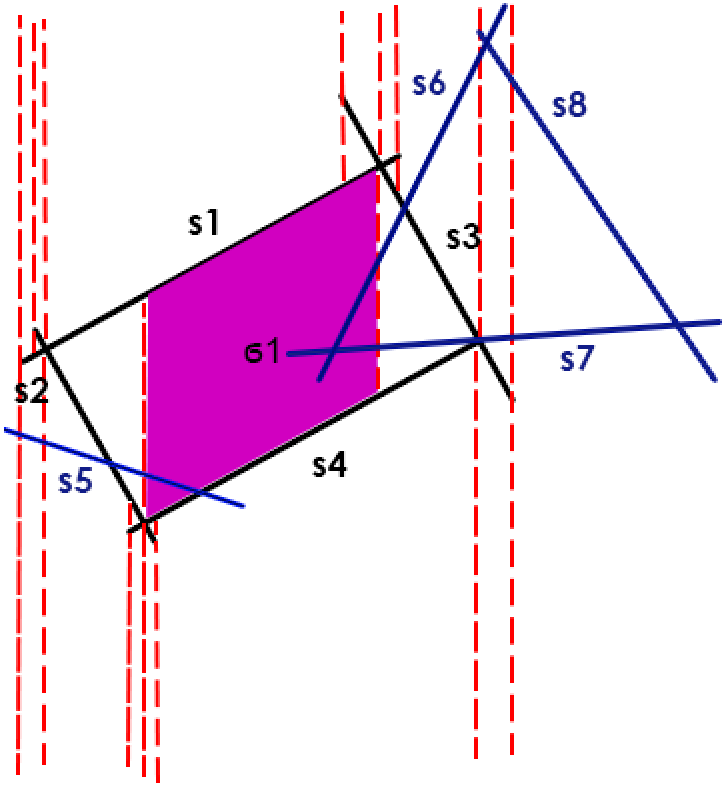
Lemma 2:

- $r \leq n$ and t, such that $1 \leq t \leq \frac{r}{d}$
- R - sample of size r
- R' - sample of size $r' = \left\lfloor \frac{r}{t} \right\rfloor$
- $\sigma \in T$ - trapezoid with weight $\omega(\sigma) \geq t\left(\frac{n}{r}\right)$

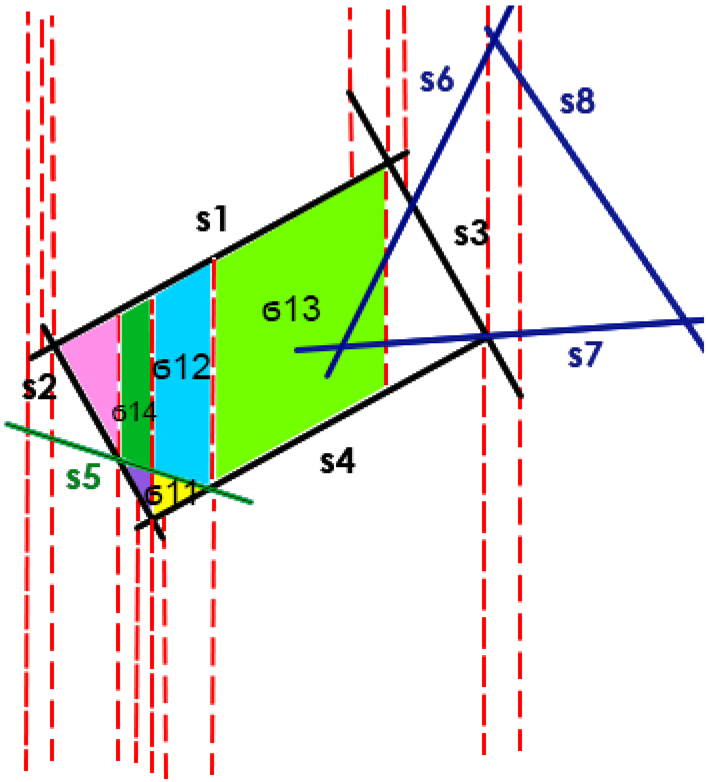
Then $\mathbf{Pr}[\sigma \in \mathcal{F}(R)] = O\left(\exp\left(-\frac{t}{2}\right) t^d \, \mathbf{Pr}[\sigma \in \mathcal{F}(R')]\right)$

Intuition: the probability that a heavy trapezoid will be created in the large sample R drops exponentially from its probability to be created in the small sample R'. (Because we are more likely to choose a conflicting segment in R).

Lemma 2 - proof - illustration



R={s1,s2,s3,s4}
|R|=4



R={s1,s2,s3,s4,s5}
|R|=5

Lemma 2 - proof – cont.

- $k = \omega(\sigma) = t(n/r)$
- $r' = \left\lfloor \frac{r}{t} \right\rfloor$
- By claim 3: $\rho_{r,n}(d,k) \approx \left(1 - \frac{r}{n}\right)^k \left(\frac{r}{n}\right)^d$

Therefore we get –

$$\frac{\mathbf{Pr}[\sigma \in \mathcal{F}(\mathrm{R})]}{\mathbf{Pr}[\sigma \in \mathcal{F}(\mathrm{R}')]} = \frac{\rho_{r,n}(d,k)}{\rho_{r',n}(d,k)} \leq \frac{2^{2d}\left(1 - \frac{1}{2} \cdot \frac{r}{n}\right)^k \left(\frac{r}{n}\right)^d}{\frac{1}{2^{2d}}\left(1 - 4\frac{r'}{n}\right)^k \left(\frac{r'}{n}\right)^d} \sim \frac{2^{2d}\left(e^{-\frac{rk}{2n}}\right)r^{\,d}}{\frac{1}{2^{2d}}\left(e^{-\frac{4r'k}{n}}\right)r'^{d}}$$

$$\sim \frac{2^{2d}\left(e^{-\frac{t}{2}}\right)}{\frac{1}{2^{2d}}\left(e^{-4}\right)} \mathrm{t}^{\mathrm{d}} = \mathrm{O}\left(\mathrm{e}^{-\frac{\mathrm{t}}{2}}\mathrm{t}^{\mathrm{d}}\right)$$

(The third transition is because $1 - x \sim e^{-x}$)

■

The exponential decay lemma

- S – set of objects
- $r \leq n$
- $1 \leq t \leq r/d$, where
  $d = \max_{\sigma \in T(S)} |D(\sigma)|$
- S complies to the axioms
- $Ef(\mathrm{r}) = \mathrm{E}[|F(\mathrm{R})|]$
- $Ef_{\geq t}(\mathrm{r}) = \mathrm{E}[|\mathrm{F}_{\geq \mathrm{t}}(\mathrm{R})|]$

Then $\mathbf{E}f_{\geq t}(r) = O\left(t^d \exp(-t/2)\, \mathbf{E}f(r)\right).$

The exponential decay lemma - proof

- R - sample of size r
- R’ - sample of size $r' = \left\lfloor \frac{r}{t} \right\rfloor$
- X_σ - indicator variable which is 1iff $\sigma \in F(R)$

$$\mathbf{E} f_{\geq t}(r) = \mathbf{E}\left[\, |\mathcal{F}_{\geq t}(\mathrm{R})| \,\right] = \mathbf{E}\left[\sum_{\sigma \in H} X_\sigma\right] = \sum_{\sigma \in H} \mathbf{E}\left[X_\sigma\right] = \sum_{\sigma \in H} \mathbf{Pr}\left[\sigma \in \mathcal{F}(\mathrm{R})\,\right]$$

$$= O\left(t^d \exp(-t/2) \sum_{\sigma \in H} \mathbf{Pr}[\sigma \in \mathcal{F}(\mathrm{R}')]\right) = O\left(t^d \exp(-t/2) \sum_{\sigma \in \mathcal{T}} \mathbf{Pr}[\sigma \in \mathcal{F}(\mathrm{R}')]\right)$$

$$= O\left(t^d \exp(-t/2)\, \mathbf{E} f(r')\right) = O\left(t^d \exp(-t/2)\, \mathbf{E} f(r)\right),$$

■

Bounded moments theorem

- $R \subseteq S$ a random sample of size r
- Denote $Ef(r) = E[|F(R)|]$
- $c \geq 1$ – arbitrary constant

Then $$\mathbf{E}\left[\sum_{\sigma \in \mathcal{F}(R)} (\omega(\sigma))^c\right] = O\left(\mathbf{E}f(r)\left(\frac{n}{r}\right)^c\right)$$

Intuition: if we want to sum up all the sizes of conflict lists after sample R (powered by some constant c), it would be similar to taking the expected number of trapezoids and multiplying it by $\left(\frac{n}{r}\right)^c$, the expected weight to the power c.

Bounded moments theorem

Sketch of the proof: By the exponential decay lemma, most regions have weight $\approx \frac{n}{r}$.

The very few that have large weight contribute little to the sum.

Applications

- Analyzing the running time of the vertical decomposition algorithm - proving the guess that the average size of the conflict list of the trapezoid of B_i is $O\left(\frac{n}{i}\right)$

- Showing an algorithm for creating a small size **(1/r)-cutting**

Proving the guess

- By **lemma 1**: the expected size of B_i (i.e the number of trapezoids in B_i) is $O\left(i + k\left(\frac{i}{n}\right)^2\right)$.

- By **bounded moments theorem** (plugging c=1), we have that the total expected size of the conflict lists computed at step i of the vertical decomposition algorithm is

$$\mathbf{E}[W_i] = \mathbf{E}\left[\sum_{\sigma \in \mathcal{B}_i} \omega(\sigma)\right] = O\left(\mathbf{E}f(i)\frac{n}{i}\right) = O\left(n + k\frac{i}{n}\right)$$

The Running Time of the Algorithm

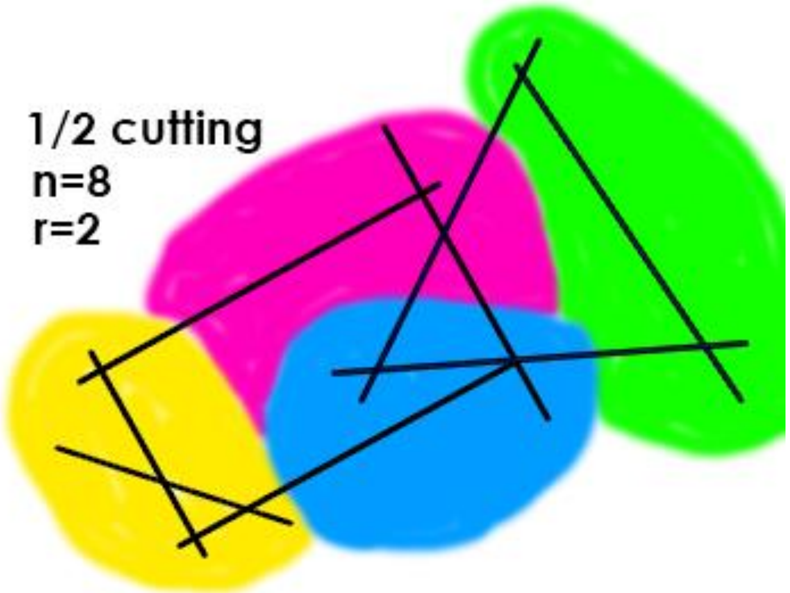
$$\mathbf{E}[W_i] = \mathbf{E}\left[\sum_{\sigma\in\mathcal{B}_i} \omega(\sigma)\right] = O\left(\mathbf{E}f(i)\,\frac{n}{i}\right) = O\left(n + k\frac{i}{n}\right)$$

And since the expected amortized work done by the algorithm in step i is $O(\frac{W_i}{i})$, we get that the total running time of the algorithm is -

$$\mathbf{E}\left[O\left(\sum_{i=1}^{n} \frac{W_i}{i}\right)\right] = O\left(\sum_{i=1}^{n} \frac{1}{i}\left(n + k\frac{i}{n}\right)\right) = O(n\log n + k)$$

(1/r)-cuttings

- S - set of n lines in the plane
- r – arbitrary parameter (<n)
- **(1/r)-cutting** of S is the partition of the plane into constant complexity regions, such that each region intersects at most n/r lines of S



Building (1/r)-cutting using vertical decomposition

- We want to show that using the vertical decomposition, we can build a (1/r)-cutting of size $O(r^2)$.

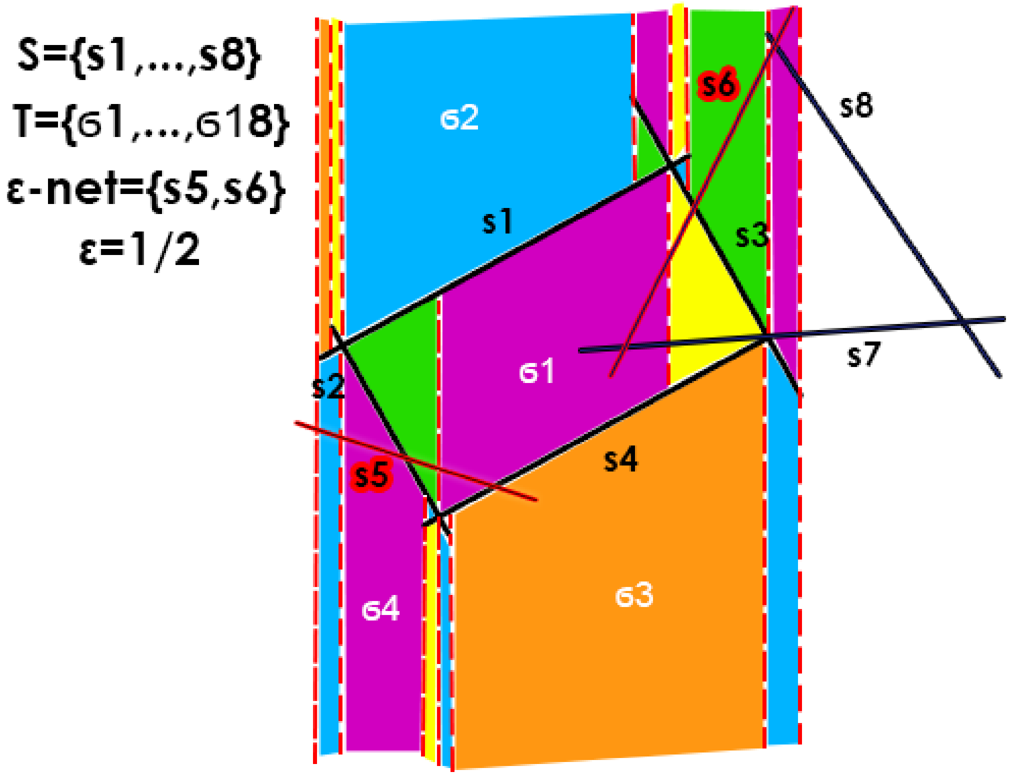
- We will show that $O(r^2)$ is the best (smallest) possible size.

- Let (S,T) be the range space, where S is the set of lines (the ground set)

- T are the trapezoids (ranges). The range of $\sigma \in T$: all the segments of S that intersect the interior of σ

- (S,T) has a VC dimension which is a constant

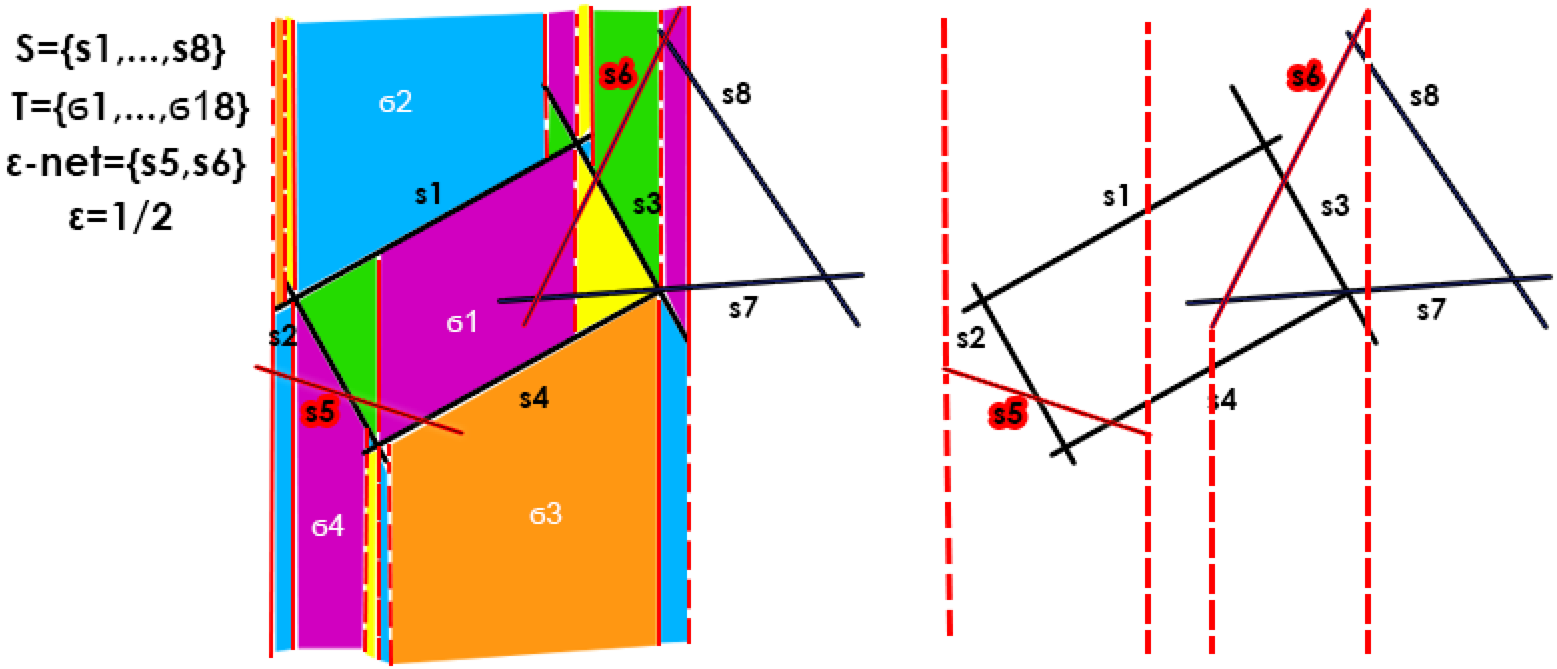
- $X \subseteq S$ – an ϵ-net for (S,T)



- By the ϵ-net theorem, there exists such an ϵ-net, of size

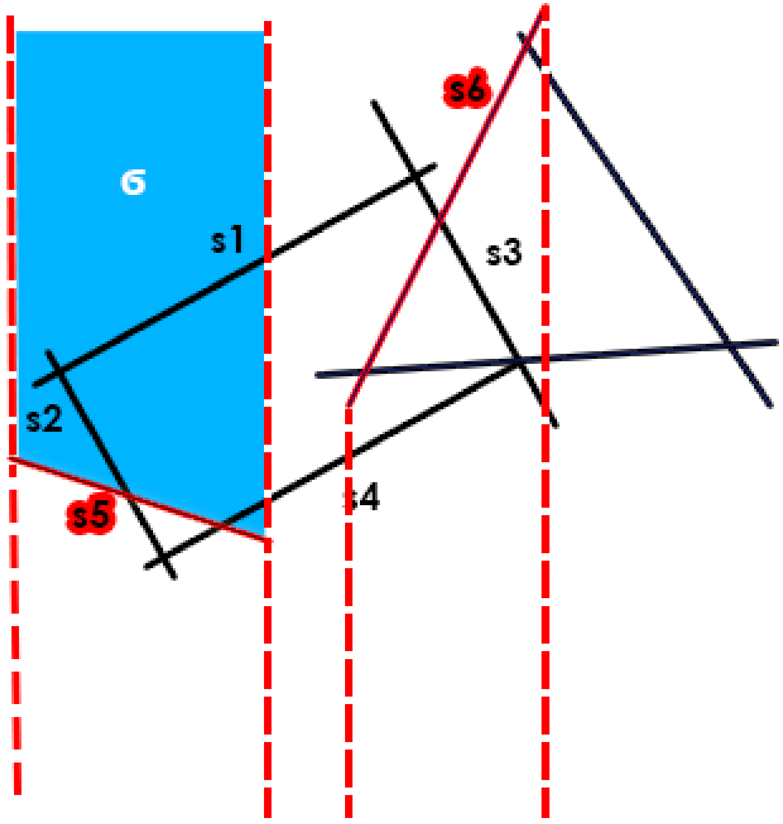
Lemma 4: There exists a (1/r)-cutting of a set of lines S in the plane of size $O\left((r\log r)^2\right)$.

Proof: consider the vertical decomposition A'(X) where X is as above (X is ϵ-net). Then, the collection of the trapezoids is the desired cutting.



Proof continuation:

The (1/r)-cutting is indeed of size $O((rlogr)^2)$, because the size of A'(X) (the number of trapezoids) is $O(|X|^2)$ and $|X| = O(rlogr)$.



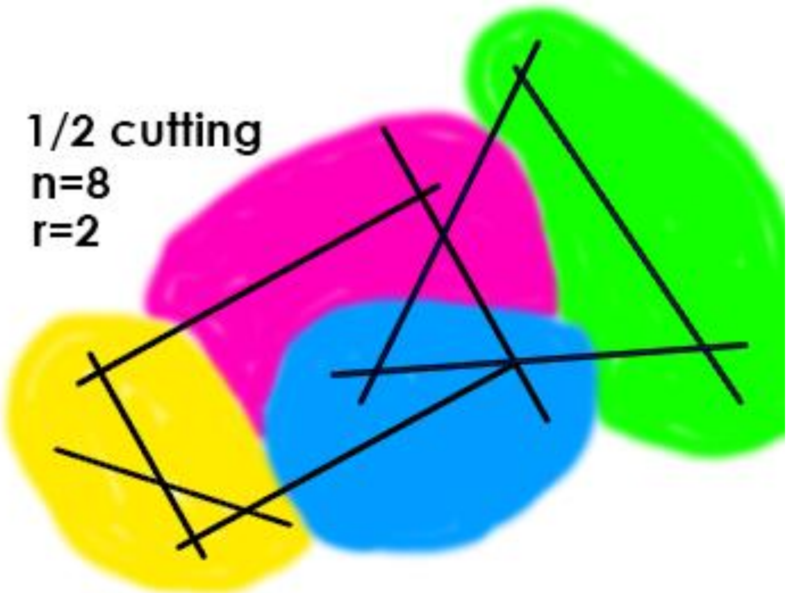
Correctness:

- Let $\sigma \in A'(S)$
- σ doesn't intersect any of the lines in X (s5, s6)
- If σ intersected more than n/r (8/2=4) lines of S in the interior, then σ intersects one of the lines in X, since X is an ϵ-net. Contradiction. ■

Claim 4: any (1/r)-cutting in the plane of n lines, contains at least $\Omega(r^2)$ regions.

Proof:

- Number of intersections in a region is at most $m = \binom{n/r}{2}$
- Number of all intersections of n lines is $M = \binom{n}{2}$



Therefore, number of regions in a cutting must be at least

$$M/m = \Omega\left(n^2/(n/r)^2\right) = \Omega\left(r^2\right).$$

■

Building (1/r)-cutting using vertical decomposition

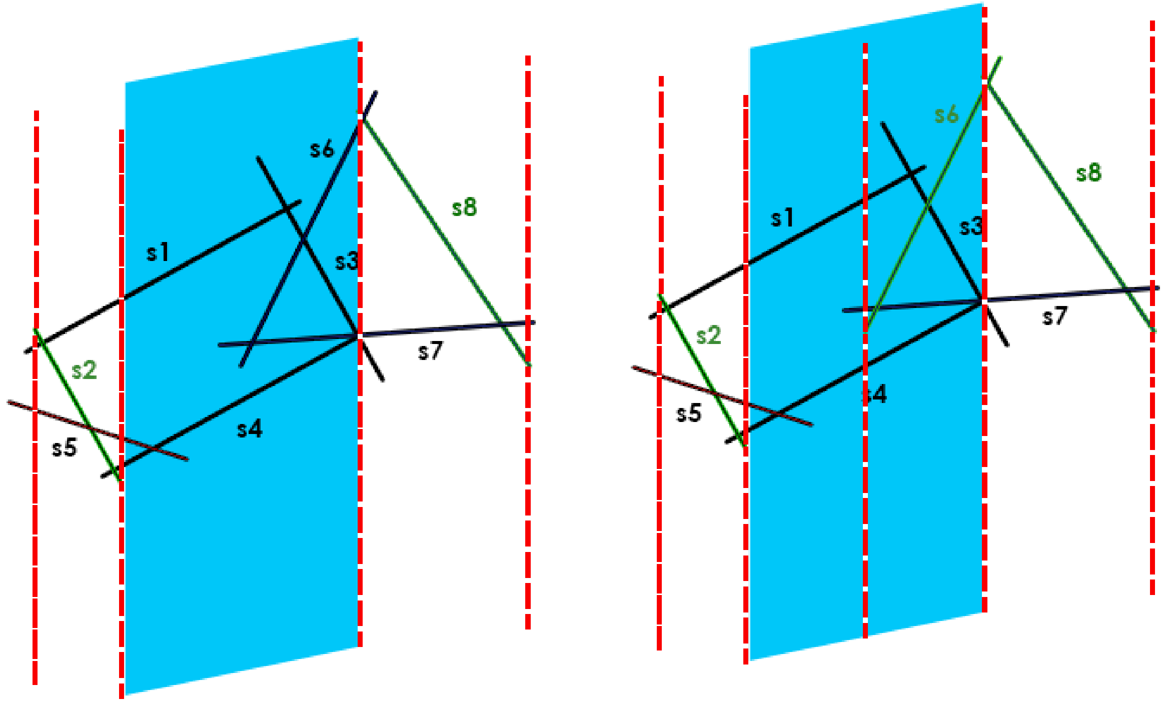
Theorem:

- S - set of lines in the plane
- r – arbitrary parameter

We can construct a (1/r)-cutting of size $O(r^2)$.

Theorem – proof

- Pick r random lines
- Build vertical decomposition



- If a trapezoid σ intersects at most n/r lines of S – add it to the cutting
- Otherwise, σ intersects t(n/r) lines of s (for some t>1) – apply a (1/t)-cutting on this trapezoid.
- Now, each trapezoid in this cutting intersects at most n/r lines in S.

Theorem – proof cont.

- The size of the cutting inside σ is $O(t^2 \log^2 t) = O(t^4)$
- By the bounded moments theorem, the expected size of the cutting is

$$O\left(\mathbf{E}f(r) + \mathbf{E}\left[\sum_{\sigma \in \mathcal{F}(R)} \left(2\frac{\omega(\sigma)}{n/r}\right)^4\right]\right) = O\left(\mathbf{E}f(r) + \left(\frac{r}{n}\right)^4 \mathbf{E}\left[\sum_{\sigma \in \mathcal{F}(R)} (\omega(\sigma))^4\right]\right)$$

$$= O\left(\mathbf{E}f(r) + \left(\frac{r}{n}\right)^4 \cdot \mathbf{E}f(r)\left(\frac{n}{r}\right)^4\right) = O(\mathbf{E}f(r)) = O(r^2)$$

∎