### CHAPTER 9 Depth Estimation via Sampling

Eitan-Hai Mashiah



#### Outline

#### Introduction

- The at most k-level
- The crossing lemma
  - On the number of incidences
  - On the number of k-sets
- A general bound for the at most k-weight

#### Outline

#### Introduction

- The at most k-level
- The crossing lemma
  - On the number of incidences
  - On the number of k-sets
- A general bound for the at most k-weight

#### Let 5 be a set of objects.

Let p be a point contained in some object. Definition: p's weight is the number of objects that contain p.

> For example: S is a set of halfplanes.



**P**'s weight is 2.

This chapter deals with p's weight/depth estimation by counting the weight of p in a random sample of objects.

The results in this chapter are not directly related to approximation algorithm.

However, the insights and general approach are useful for later results presented in the book.

#### Outline

#### Introduction

The at most k-level

#### The crossing lemma

- On the number of incidences
- On the number of k-sets

#### A general bound for the at most k-weight

Let L be a set of n lines in the plane.

**Definition:** A point p,  $p \in \bigcup_{l \in L} l$ , is of level k, if there are k lines strictly below p.





### Definition: The k-level is the closure of the set of points of level k.



#### The at most k-level

#### The number of vertices at the k-level

The O-level has at most n-1 vertices.



#### The at most k-level

#### The number of vertices at the k-level

The number of vertices at the k-level (k>0) is hard question!

1-level

**Each line might** contribute more than one segment to the k-level, and the polygon defined by the k-level is no longer convex.

Definition: The at most k-level is the closure of the set of points of at most level k; i.e. there are at most k lines below them.





The number of vertices of the AT MOST k-level

Theorem 1. The number of vertices of level at most k in an arrangement of n lines in the plane is O(nk). Theorem 1. O(nk) vertices of level at most k.

**Proof:** Let  $L_{\leq k}$  be the set of vertices of level at most k.

Let R be a random sample of L, where each line is picked with probability 1/k. In particular, E[|R|]=n/k.



#### The at most k-level

#### Theorem 1. O(nk) vertices of level at most k.

**Proof** cont.: assume that a vertex p is of level  $0 \le j \le k$ . Let  $X_p$  be an indicator that is 1 if p is in the O-level of R, then

$$P(X_p = 1) = \left(1 - \frac{1}{k}\right)^j \left(\frac{1}{k}\right)^2 \ge \left(1 - \frac{1}{k}\right)^k \left(\frac{1}{k}\right)^2 \ge \exp(-2\frac{k}{k}) \frac{1}{k^2} = \frac{1}{e^2k^2}$$



#### The at most k-level

#### Theorem 1. O(nk) vertices of level at most k.

Proof cont.:

$$P\left(X_{p}=1\right) \geq \frac{1}{e^{2}k^{2}}$$

#### On the other hand

$$\sum_{P \in L_{\leq k}} X_p \leq |\mathbf{R}| - 1 \implies \sum_{P \in L_{\leq k}} E[X_p] = E\left[\sum_{P \in L_{\leq k}} X_p\right] = E[|\mathbf{R}| - 1] \leq \frac{n}{k}$$
  
Hence,

$$\frac{n}{k} \ge \sum_{P \in L_{\le k}} E\left[X_p\right] = \sum_{P \in L_{\le k}} P\left(X_p = 1\right) \ge \frac{\left|L_{\le k}\right|}{e^2 k^2} \quad \Rightarrow \quad \left|L_{\le k}\right| \le e^2 kn$$

#### Outline

#### Introduction

The at most k-level

#### The crossing lemma

- On the number of incidences
- On the number of k-sets
- A general bound for the at most k-weight



Let G=(V(G),E(G)) be a graph with n vertices and m edges.

Definition: A graph G is planar if it can be drawn in the plane so that none of its pair of edges are crossing.



#### The crossing lemma

#### Theorem 2.(Euler's formula) For a connected planar graph G, one has f-m+n=2, where f,m and n are the number of faces, edges, and vertices in a planar drawing of G.

Face of a planar graph G is a region bounded by edges, including the outer, infinitely large region.





### Lemma 3. If G is a simple planar graph and n≥3 then m≤3n-6.



#### Lemma 3. If G is a simple planar graph and n≥3 then m≤3n-6.

**Proof:** Assume first that the number of edges of a planar graph G be maximal. Hence, each face in G is a triangle.

Notice, each edge in G is an edge in two such triangles. Therefore, 2m=3f.

Using Euler's formula, 2=f-m+n=2/3m-m+n=-m/3+n. In conclusion, m=3n-6.

Finally, if **m** is not maximal then m≤3n-6.



Lemma 3. If G is a simple planar graph and n≥3 then m≤3n-6.

Conclusion: The complete graph over 5 vertices  $K_5$  is not planar.  $10 = \binom{5}{2} = m \leq 3n - 6 = 9$ 





### Note, the bipartite complete graph with 3 vertices on each side, $K_{3,3}$ , is not a planar.



Kuratowski Theorem: A graph is a planar if and only if it does not contain either K<sub>5</sub> or K<sub>3,3</sub> induced inside it.



Definition: The crossing number of G, denoted by c(G), is the minimal number of edge crossings in any drawing of G in the plane.

For example, for a planar graph G, c(G)=0, while for  $K_5$ ,  $c(K_5)=1$ .





### Claim 4. Let G be a simple graph with n≥3, then c(G)≥m-3n+6.



Claim 4. Let G be a simple graph with n≥3, then <u>c(G)≥m-3n+6</u>.

**Proof:** If  $m-3n+6\leq 0$  then the claim holds trivially, since  $c(G)\geq 0$ .

Else, by Lemma 3, G is not planar. Draw G such that there are c(G) edge crossing.

Let H=(V(H), E(H)) be the graph induced by removing one edge from each edge crossing pair. Then,  $|E(H)| \ge m-c(G)$ . In addition, H is planar so  $|E(H)| \le 3|V(H)| - 6 = 3n - 6$ .

In conclusion,  $3n-6 \ge m-c(G)$ .

#### The crossing lemma

# Claim 5 (Crossing lemma). Let G be a simple graph. If $m \ge 6n$ then $c(G)=\Omega(m^3/n^2)$ .



#### Claim 5. Let G be a simple graph. If $m \ge 6n$ then $c(G)=\Omega(m^3/n^2)$ .

**Proof:** Let **D** be a drawing of G in the plane that has c(G) crossings.

Let U be a random set of vertices of V(G) where each vertex is picked with probability p=6n/m. Note that, since m  $\geq 6n$ , p satisfies 0 as necessary.

Denote by H=(U,E') such that  $E'=\{(u,v)\in E(G) \mid u,v\in U\}$ .



#### Claim 5. Let G be a simple graph. If $m \ge 6n$ then $c(G)=\Omega(m^3/n^2)$ .

- **Proof** cont.: Let  $X_v$  and  $X_e$  be the number of vertices and edges in H.  $E[X_v] = np$  and  $E[X_e] = mp^2$ .
  - Denote by  $X_c$  the number of crossing in  $D_H$ .  $E[X_c] = c(G)p^4$ By Claim 4,  $X_c \ge c(H) \ge X_e - 3X_v + 6$  Therefore,

$$c(G)p^{4} = E[X_{c}] \ge E[X_{e}] - 3E[X_{v}] = mp^{2} - 3np \implies$$

$$c(G) \ge \frac{m}{p^{2}} - \frac{3n}{p^{3}} \stackrel{\mathsf{p=6n/m}}{\ge} \frac{m^{3}}{72n^{2}} \qquad \blacksquare$$

#### Outline

#### Introduction

The at most k-level

#### The crossing lemma

- On the number of incidences
- On the number of k-sets
- A general bound for the at most k-weight

Let P be a set of n distinct points in the plane.

Let L be a set of m distinct lines in the plane.

Denote by I(P,L) the number of pairs  $(p,l) \in P \times L$  such that  $p \in l$ . I(P,L) is the number of incidences between lines of L and points of P.



Let  $I(n,m) = \max_{|P|=n,|L|=m} I(P,L)$  the maximal number of incidences between n points and m lines.

Lemma 6. The maximal number of incidences between n points and m lines is

$$I(n,m) = O(n^{2/3}m^{2/3} + n + m)$$

#### Lemma 6. $I(n,m) = O(n^{2/3}m^{2/3} + n + m)$ .

**Proof:** Let P and L be the set of n points and m lines, respectively, such that I(P,L)=I(n,m).

Define a graph G as follows: V(G) = P, and  $(p,p') \in E(G)$  iff p,p' lie consecutively on some line of L. Set e(G)=|E(G)| and  $v(G)=|V(\underline{G})|$ .



Lemma 6.  $I(n,m) = O(n^{2/3}m^{2/3} + n + m)$ .

**Proof** cont.: Note,  $e(G) \ge I(n,m) - m$ , v(G) = n, and  $c(G) \le m^2$ . By Lemma 5, If  $e(G) \ge 6v(G)$  then

$$m^{2} \ge c(G) \ge a \cdot e(G)^{3} / v(G)^{2} \ge a \cdot (I(n,m) - m)^{3} / n^{2} \Longrightarrow$$
$$I(n,m) = O\left(m^{2/3}n^{2/3} + m\right);$$
Otherwise,  $I(n,m) - m \le e(G) < 6v(G) = 6n \Longrightarrow$ 
$$I(n,m) = O\left(n + m\right)$$



Let  $I(n,m) = \max_{|P|=n,|L|=m} I(P,L)$  the maximal number of incidences between n points and m lines.

Lemma 7. The maximal number of incidences between n points and n lines is

$$I(n,n) = \Omega(n^{4/3})$$

**Lemma 7.** 
$$I(n,n) = \Omega(n^{4/3})$$

**Proof:** Assume that  $N = n^{1/3} / 2$  is an integer.

Let 
$$P = \{ (x, y) \mid x \in \{1, 2, ..., N\}, y \in \{1, 2, ..., 8N^2\} \}$$
.  
Let  $L = \{ y = ax + b \mid a \in \{1, 2, ..., 2N\}, b \in \{1, 2, ..., 4N^2\} \}$ .

Clearly |P| = |L| = n. In addition,  $x \in \{1, 2, ..., N\} \Rightarrow y = ax + b \le 2N \cdot N + 4N^2 \le 8N^2$ .

Hence, every line is incident to N points of P. Therefore,  $I(P,L) = |L|N = n \cdot n^{1/3} / 2 = n^{4/3} / 2.$ 



#### Outline

#### Introduction

The at most k-level

#### The crossing lemma

- On the number of incidences
- On the number of k-sets

#### A general bound for the at most k-weight

The crossing lemma On the number of k-sets

Let P be a set of n points in the plane such that no three points are collinear.

Definition: A pair of points  $p,q \in P$  form a k-set if there are k points in the closed halfplane below the line line(p,q) passing through p and q.



#### The crossing lemma On the number of k-sets

Via duality, the number of k-sets is exactly the complexity of the k-level in the dual arrangement.

Via duality

line  $l: y=cx+d \longrightarrow l^*: (c,-d)$  point A point p is below a line l, if and only if  $p^*$  is below  $l^*$ . So, every k-set in the original setting corresponds to a vertex on the (k-2)-level in the dual setting.



The k-set problem

Point p: (a,b)



The k-level problem

> p\*: y=ax-b line

Let G=(P,E) be a graph that has an edge (p,q) if they form a k-set.

Lemma 8 (Antipodality). Let (q,p) and (s,p) be two k-set edges of G, with q and s to the left of p. Then, there exists a point  $t \in P$  to the right of p such that (p,t) is a k-set, and line(p,t) lies between line(p,q) and line(p,s).



#### The crossing lemma On the number of k-sets

#### Lemma 8 (Antipodality).

**Proof:** Let  $f(\alpha)$  be the number of points below or on the line passing through p and having a slope  $\alpha$ . Set  $f_+(\alpha) = \lim_{\beta \to \alpha^+} f(\beta), \quad f_-(\alpha) = \lim_{\beta \to \alpha^-} f(\beta),$ 



#### The crossing lemma On the number of k-sets

#### Lemma 8 (Antipodality).

**Proof** cont.: Let  $\alpha_q$  and  $\alpha_s$  be the slope of the lines line(p,q) and line(s,p), respectively. Assume w.l.o.g. that  $\alpha_q < \alpha_s$ .

Note that 
$$f(\alpha_q) = k = f(\alpha_s), f_+(\alpha_q) = k - 1, f_-(\alpha_s) = k.$$

Let  $\beta$  be the minimal  $\alpha_q < \beta < \alpha_s$ , such that  $f(\beta) = k$ . In particular,  $f_-(\beta) = k - 1$ .

Hence, there must be a point  $t \in P$  such that the slope of line(t,p) is  $\beta$ . Furthermore, t must be to the right of p.



Lemma 9. Let p,q∈P such that q is of p's left and (p,q) is k-set edge that has the largest slope among all such edges.

Furthermore, assume that there are k-1 points of P to the right of p.

Then, there exists a point  $s \in P$ , such that (p,s) is k-set edge and it has larger slope than (p,q).



#### The crossing lemma On the number of k-sets

#### <u>Lemma 9.</u>

**Proof:** Let  $\alpha$  be the slope of line(p,q), then  $f(\alpha)=k$  and  $f_{+}(\alpha)=k-1$ .

Since there are k-1 points to the right of p,  $f_{-}(\infty) \ge k$ .

Hence, there must be a k-set, (p,s), that defines a line with a slope >  $\alpha$ .

However, since (p,q)has the maximal slope to the left, s is necessarily a point to the right of p with slope > $\alpha$ .



#### The crossing lemma On the number of k-sets

#### Forming chain of edges in G

Imagine, e=(q,p) is a k-set edge and that q is to the left of p.

Rotate the line around p (counterclockwise) till a k-set edge e'=(p,s) is found where s is to the right of p. Walk from e to e' and continue this way forming a chain of edges in G.

<u>Note</u>, each chain ended in one of the last k-1 points.

No two chains are merged using the same edges.



Lemma 10. The edges of G can be decomposed into k-1 convex chains  $C_1$ ,  $C_2$ ,...,  $C_{k-1}$ . Similarly, The edges of G can be decomposed into m=n-k+1 concave chains  $D_1$ ,  $D_2$ ,...,  $D_m$ .



Lemma 10. E(G) can be decomposed into k-1 convex chains  $C_1$ ,  $C_2$ ,...,  $C_{k-1}$ . Similarly, it can be decomposed into m=n-k+1 concave chains  $D_1$ ,  $D_2$ ,...,  $D_m$ .

**Proof:** The first part of the Lemma follows directly the process presented before.

For the second part one may rotate the plane by 180°. Each k-set is now an (n-k+2)-set. Hence, it can be decomposed into n-k+1 convex chains that can be interpreted as an n-k+1 concave chains in the original

plane.





The crossing lemma On the number of k-sets

#### Lemma 11. The number of k-sets defined by a set of n points in the plane is O(nk<sup>1/3</sup>).

Lemma 11. The number of k-sets defined by a set of n points in the plane is O(nk<sup>1/3</sup>).

Proof: Let G=(P,E) where E is the set of k-set edges. Then, |V(G)|=|P|=n and m=|E(G)| is the number of k-sets.

By Lemma 10, crossing of edges of G is an intersection point of one convex chain of  $C_1, C_2, ..., C_{k-1}$  with a concave chain of  $D_1, D_2, ..., D_{n-k+1}$ .



Lemma 11. The number of k-sets defined by a set of n points in the plane is O(nk<sup>1/3</sup>).

Proof cont: Therefore, there are at most 2(k-1)(n-k+1)
 crossing in G.

By the crossing lemma, if  $m \ge 6n$  then  $c(G) \ge a(m^3/n^2)$ . Hence,  $m^3/n^2 = O(nk)$  which implies that  $m = O(nk^{1/3})$ . Otherwise, m < 6n, m = O(n).



#### Outline

#### Introduction

- The at most k-level
- The crossing lemma
  - On the number of incidences
  - On the number of k-sets
- A general bound for the at most k-weight

### A general bound for at most k-weight Reminder

- Let S be a set of objects.
- For a subset R<sub>⊆</sub>S, let *F*(R) be the set of regions defined by R.
- Let  $\mathcal{T}=\mathcal{T}(S)$  be the set of all possible region defined by S.

#### For example:

- The set of objects, S, is a set of lines in the plane. Each line defines a closed halfplane below it.
- For every subset of lines R, the set of regions F(R) is the set of vertices in the polygon produced by the intersection of halfplanes defined by R.
- Tis the set of all the vertices defined by pair of intersecting lines in S.



Reminder

### A general bound for at most k-weight Reminder

For every region  $\sigma \in \sigma$  we associate two sets, first

Definition: The defining set of  $\sigma$ , denoted by D( $\sigma$ ), is the subset of S defining the region  $\sigma$ .

Assume  $|D(\sigma)| \le d$  for a small constant d, which is the combinatorial dimension.

#### In the example,

for every vertex  $\sigma$ , D( $\sigma$ ), is the set of the pair of lines that form  $\sigma$ .



Next, we associate for every region  $\sigma \in \sigma$ :

Definition: The conflicting set of  $\sigma$ , denoted by K( $\sigma$ ), is the set of objects of S such that if any object of K( $\sigma$ ) is in R then  $\sigma \notin \mathcal{F}(R)$ .

Reminder

In the example, for every vertex  $\sigma$ , K( $\sigma$ ) contains all the lines below  $\sigma.$ 



## A general bound for at most k-weight befinition: The weight of $\sigma$ , is $w(\sigma)=|K(\sigma)|$ .

In the example, for every vertex  $\sigma$ , the weight of  $\sigma$ , w( $\sigma$ ), is the number of lines below  $\sigma$ , namely, the level of  $\sigma$ .



### A general bound for at most k-weight Reminder

Axioms. Let S,  $\mathscr{F}(R)$ , D( $\sigma$ ) and K( $\sigma$ ) be such that for any subset R<sub>S</sub>, the set  $\mathscr{F}(R)$  satisfies the following axioms:

1) For any  $\sigma \in \mathscr{F}(\mathbb{R})$ , one has  $D(\sigma) \subseteq \mathbb{R}$  and  $K(\sigma) \cap \mathbb{R} = \phi$ . 2) If  $D(\sigma) \subseteq \mathbb{R}$  and  $K(\sigma) \cap \mathbb{R} = \phi$  then  $\sigma \in \mathscr{F}(\mathbb{R})$ ,



Let  $\Gamma_{\leq k}$  be the set of regions with weight at most k.

For a random sample R of size r from S, we denote  $f_0(r) = E[|\mathcal{F}(R)|].$ 

Theorem 12. Let S be a set of objects, with combinatorial dimension of d, and let k be a parameter. Let R be a random sample of S created by picking each element of S with probability 1/k. Then, for some parameter c we have

$$|\Gamma_{\leq k}(S)| \leq cE\left[k^d f_0(|R|)\right].$$

Theorem 12. 
$$|\Gamma_{\leq k}(S)| \leq cE\left[k^d f_0(|R|)\right]$$
.

**Proof:** Let **R** be a random sample of **S**, where each object is picked with probability **1/k**.

Assume that the weight of  $\sigma$  is j, where  $0 \le j \le k$ . By the axioms,  $\sigma$  is in the **O-weight** of the sample **R** iff  $\sigma \in \mathcal{F}(\mathbf{R})$ .

Theorem 12. 
$$\left|\Gamma_{\leq k}(S)\right| \leq cE\left[k^d f_0\left(|R|\right)\right]$$

**Proof** cont.: Let  $X_{\sigma}$  be an indicator that gets 1 iff  $\sigma$  is in the 0-weight of R iff  $\sigma \in \mathcal{F}(R)$ , then

$$P\left(X_{\sigma}=1\right) \geq \left(1-\frac{1}{k}\right)^{j} \left(\frac{1}{k}\right)^{d} \geq \left(1-\frac{1}{k}\right)^{k} \left(\frac{1}{k}\right)^{d} \geq \frac{1}{e^{2}k^{d}}.$$

The sample of size  $|\mathbf{R}|$  has equal probability of being picked to be R. Hence,  $f_0(r) = E[|\mathcal{F}(\mathbf{R})||R|=r]$ . Hence,  $E[f_0(|\mathbf{R}|)] = E[E[|\mathcal{F}(\mathbf{R})||R|=r]] =$ 

$$E\left[|\mathscr{F}(R)|\right] \ge \sum_{\sigma \in \Gamma_{\leq k}} E\left[X_{\sigma}\right] \ge \frac{\left|\Gamma_{\leq k}\right|}{k^{d} e^{2}}$$

Lemma 13. Let  $f_0(\bullet)$  be a monotone increasing function which is well behaved; namely, there exists a constant c, such that  $f_0(xr) \le c f_0(r)$ , for any r and  $1 \le x \le 2$ . Let Y be the number of heads in n coin flips where probability for head is 1/k. Then,

$$E\left[f_0(Y)\right] = O\left(f_0(n / k)\right).$$

Lemma 13. Let  $f_0(\bullet)$  be a monotone increasing function which is well behaved, and  $Y \sim Bin(n, 1/k)$ . Then,  $E[f_0(Y)] = O(f_0(n/k))$ .

**Proof:** Notice that E[Y]=n/k, and by Chernoff's Inequality,

$$P(Y \ge t(n/k)) \le 2^{-t(n/k)}$$

In addition,

$$f_0\left((t+1)n/k\right) \le cf_0\left(\frac{(t+1)}{2}n/k\right) \le c^{\lceil \log(t+1)\rceil}f_0\left(n/k\right).$$

Lemma 13. Let  $f_0(\bullet)$  be a monotone increasing function which is well behaved, and  $Y \sim Bin(n, 1/k)$ . Then,  $E[f_0(Y)] = O(f_0(n/k))$ .

**Proof** cont.: In conclusion,

$$\begin{split} &E\Big[f_0\left(Y\right)\Big] \leq \sum_i f_0\left(i\right) P(Y=i) \leq \\ &f_0\left(10\frac{n}{k}\right) + \sum_{t=10}^{k-1} f_0\left(\left(t+1\right)\frac{n}{k}\right) P\left(Y \geq t\frac{n}{k}\right) \leq \\ &O\left(f_0\left(n/k\right)\right) + \sum_{t=10}^{k-1} c^{\lceil \log(t+1) \rceil} f_0\left(n/k\right) 2^{-t\frac{n}{k}} = O\left(f_0\left(n/k\right)\right). \end{split}$$

#### A conclusion:

Theorem 14. Let S be a set of n objects, with combinatorial dimension of d, and let k be a parameter. Assume that the number of regions formed by a set of m objects is bounded by a well behaved function  $f_0(m)$ . Then

$$\Gamma_{\leq k}(S) \Big| = \mathcal{O}\Big(k^d f_0(n / k)\Big).$$

In particular, if  $f_{\leq k}(n) = \max_{|S|=n} \Gamma_{\leq k}(S)$  be the maximum number of regions of weight at most k that can be defined by any set of n objects, then

$$f_{\leq k}(n) = O\left(k^d f_0(n/k)\right).$$

