Sublinear Algorithms	October 13-15, 2015

Lecture 13,14

Lecturer: Sofya Raskhodnikova *Scribe(s): Eunou Lee*

1 Introduction

This week, we will cover cover sampling based algorithms testing convexity of 2D images which is a problem to decide whether a given image is convex, or ϵ-far from convex. The algorithm has query complexity $O(\epsilon^{-4/3})$. Note that some of the text and all the images are directly adopted from the original paper[BMRar].

2 Preliminaries and Problem Definition

Figure representation: A figure (U, C) contains a compact convex universe $U \subseteq \mathbb{R}^2$, and a measurable subset $C \subseteq U$.
Input (U, C) is convex if C is convex, and for simplicity, we only consider cases where U is a square.

Sampling: By sampling, we mean that for uniformly random point $(x, y) \in U$, we get (x, y) and its color.

Problem: Given an image C and a parameter ϵ, a convexity tester should
1) accept the image with probability at least 2/3 if C is a convex image; and
2) reject the image with probability at least 2/3 if C is ϵ-far from the convexity property.

Results: Previously, query complexity $O(1/\epsilon^{3/2})$ was implied by learner of for 1-sided error. In this lecture, we present $O(1/\epsilon^{4/3})$ algorithm.

3 Algorithm

The algorithm is displayed in Algorithm 1.

Algorithm 1: A ϵ-convexity tester.

1 Sample $50\epsilon^{-4/3}$ points from uniformly and independently at random
2 Find convex hull of sampled black points.
3 **If** a white sample is in the convex hull, **then** *reject*;
4 **otherwise**, *accept*.

4 Proof Idea - Poissonization

Convex figures are always accepted. To analyze the case where images are $\epsilon-$far from convex, we use Poissonization. We want to think about disjoint parts of the image as being sampled from independently. But since this is not true under uniform sampling, we will define Poisson sampling that achieves this and approximates uniform sampling well.

Definition 1. *The* ***Poisson distribution*** *with parameter* $\lambda \geq 0$*, denoted* $Po(\lambda)$*, takes a value* $\mathbb{N}^{\geq 0}$ *with probability* $\frac{e^{-\lambda x}}{x!}$.

A random variable distributed according to $Po(\lambda)$ has expectation and variance equat to λ.
By Poissonization, we mean that we modify a probabilistic experiment to replace a fixed quantity (e.g., the number of samples) with a variable one that follows a Poisson distribution. This breaks up dependencies between different events, and makes the analysis tractable.

Definition 2. *A uniform algorithm is called Poisson-s if the number of sameples it takes is distributed as $Po(s)$.*

Lemma 3 (Poissonization lemma)**.**

- *Poisson algorithms can simulate uniform algorithms.*
- *Suppose the uniform distribution sampled by a Poisson-s algorithm* 1) *contains different disjoint types of outcomes;* 2) *one of these outcomes appears with probability q.*
  Then the number of samples of that type of algorithms is distributed as $Po(qs)$. ***Moreover, it is independent of the number of samples of other types of outcomes.***

Now, what is the smallest number of points to prove nonconvexity? We need 4 points if we assume the points are in general position: three black points forming a triangle and a white point inside it. An idea toward this end is that one witness point can be virtual.

Definition 4 (central point)**.** *The point of intersection of two lines is called central point if each open quadrant formed by the lines has black area $\geq \epsilon/4$. Let u denote a central point.*

Claim 5.

- *u exists, if the image is epsilon-far from convexity,*
- *Pr(u is in the convex hull of black samples)≥ 0.99.*

We can use a central point u as our virtual black point, which means u can perfrom as one of black point of nonconvexity witness.

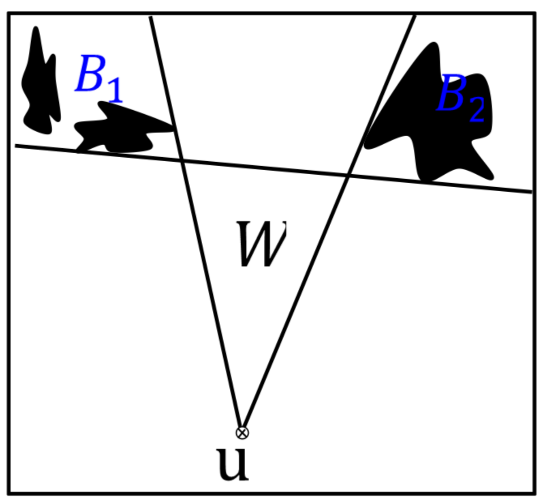

Figure 1: u in a witness

In the figure, if the black areas of B_1 and B_2 are both at least $t = 0.025\epsilon^{3/2}$, and if the white area of W is at least $0.025\epsilon^{4/3}$ we call it **many whites pattern**. If the white area of W is less than that, we call it **few-whites pattern**. We call (b_1, b_2, w) where $b_1 \in B_1$, $B_2 \in B_2$, $w \in W$ a witness triple.

For analyizing probability of sampling witness triple, we

5 Proof of correctness

Analysis consists of 2 phases: Recoloring-to-brown phase and Sweeping phase.

5.1 Recoloring-to-brown phase

Consider following mental experiment:
while there is a many-whites pattern,

- choose many-whites pattern, and
- recolor $[0.01\epsilon^{3/2}n^2]$ black pixels to brown in each B_1 and B_2.

Lemma 6. *If the number of iterations* $\geq 10\epsilon^{-1/3}$ *black pixels, the tester succeeds with probability* $\geq 2/3$.

Proof. Enumerate patterns. For pattern i, let E_i be event that black points are sampled from each of reagion B_1 and B_2. $\Pr(E_i) \geq 0.25\epsilon^{1/3}$. By the proposition below, $\Pr(\cup E_i) \geq 1 - e^{-0.25} \geq 0.91$. Also, Pr(a white is sampled in W' of j) ≥ 0.75 For some pattern j, let b_j its corresponding black central point. Then $\Pr(seeingaviolationofconvexity) \geq 0.91 * 0.75 * 0.99 \geq 2/3$. □

Proposition 7. *Let* $E_1, E_2, \ldots, E_k$ *be independent events such that* $\Pr(E_i) \geq q_i$. *Then* $\Pr(\cup E_i) \geq 1 - e^{-\sum q_i}$.

Proof. Use Markov's inequality and Taylor expansion. □

5.2 Sweeping phase

1. *Sweep with horizontal and vertical lines.* S_{l_i} *be set of "swept" black pixels by* l_i.

$$|S_{l_i}| = 4[0.01\epsilon^{3/2}n^2]$$

2. *Let* T_0 *be set of 4 gray triangles (see Figure 2.)*

3. *For* $i = 1$ *to* $m = \log \frac{2}{\sqrt{\epsilon}}$*:*
   Process each triangle T *in* T_{i-1} *to obtain triangles* T' *and* T'' *in* T_i*:*
   - *move line* l *parallel to the base of* T *until it hits* $b \in S_B$.
   - $W = (W \cup p_l)$
   - $B = B \cup S_l$, *where* p_l *is the anchor point of* S_l *on* l *(see Figure 4).*

Note that the sum of areas of all triangles in $T_m \leq 0.25\epsilon n^2$.

References

[BMRar] *Piotr Berman, Meiram Murzabulatov, and Sofya Raskhodnikova. Testing convexity of figures under the uniform distribution.* 32nd International Symposium on Computational Geometry (SoCG), *2016. To appear.*

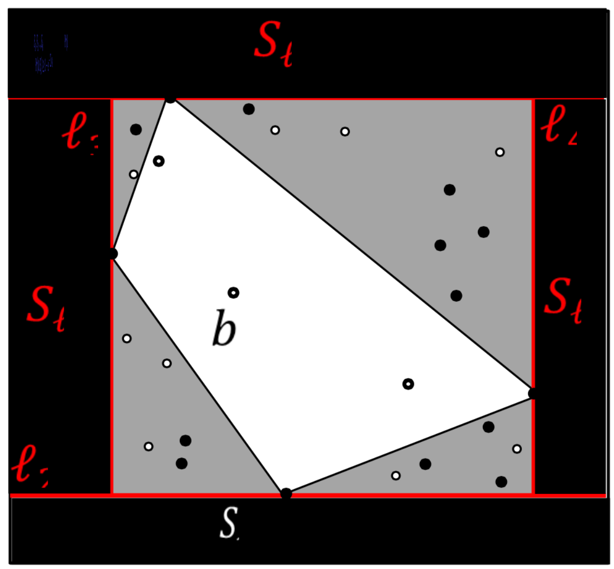

Figure 2: Sweeping phase

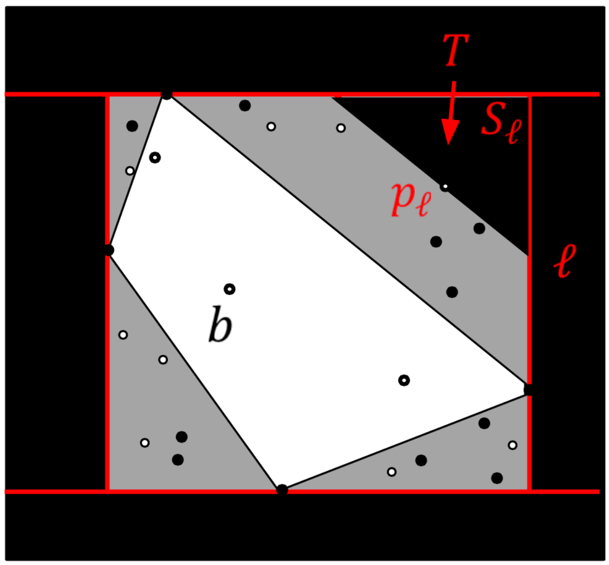

Figure 3: Sweeping phase