The effect of noise on the number of extreme points

Dominique Attali - GIPSA Grenoble

Olivier Devillers - INRIA Sophia-Antipolis

Xavier Goaoc - INRIA Nancy

Introduction

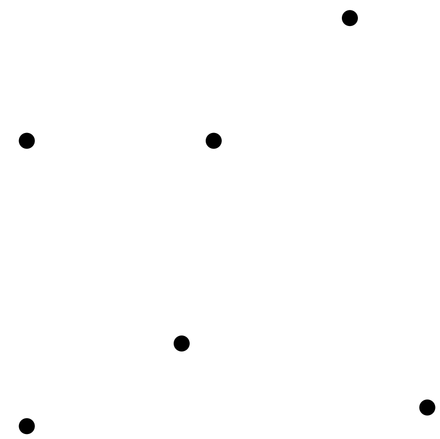
Geometric data structure

Complexity analysis

Limits: structure and precision

Geometric data structures

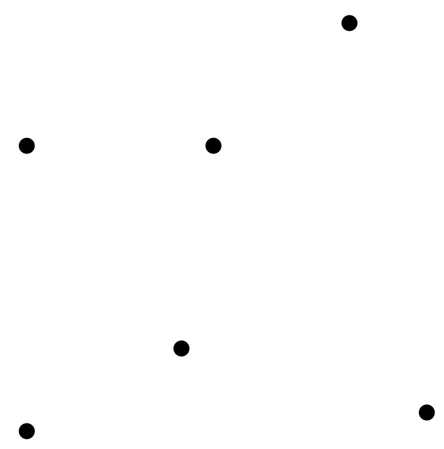
Let P be a set of n points...



Geometric data structures

Let P be a set of n points...

P induces many structures.

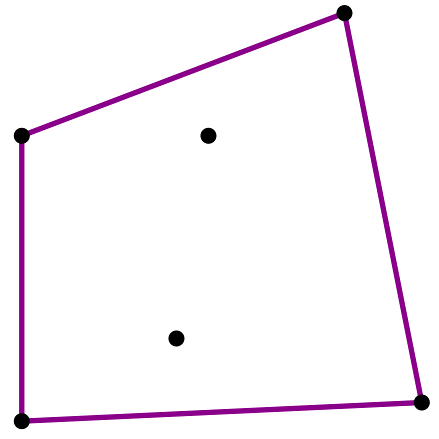


Geometric data structures

Let P be a set of n points...

P induces many structures.

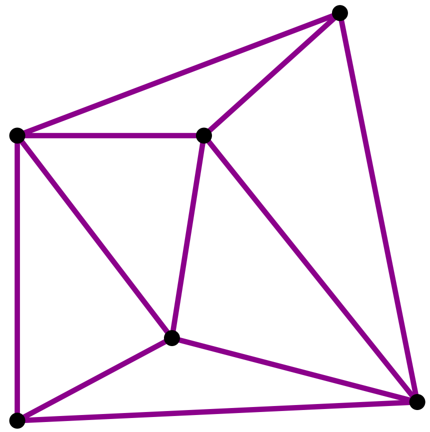
- $\star$ the *convex hull* of P



Geometric data structures

Let P be a set of n points...

P induces many structures.

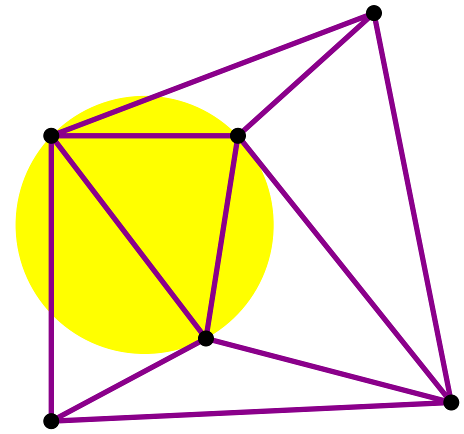


$\star$ the *convex hull* of P

$\star$ the *Delaunay triangulation* of P

Geometric data structures

Let P be a set of n points...



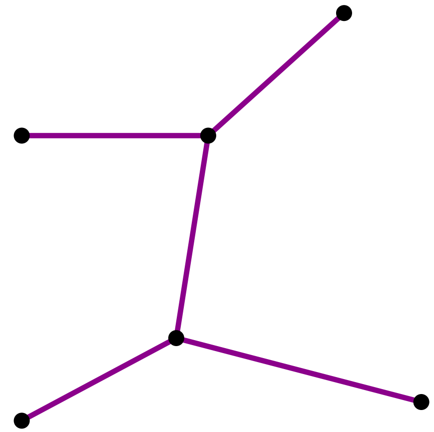
P induces many structures.

- ⋆ the *convex hull* of P
- ⋆ the *Delaunay triangulation* of P

Geometric data structures

Let P be a set of n points...

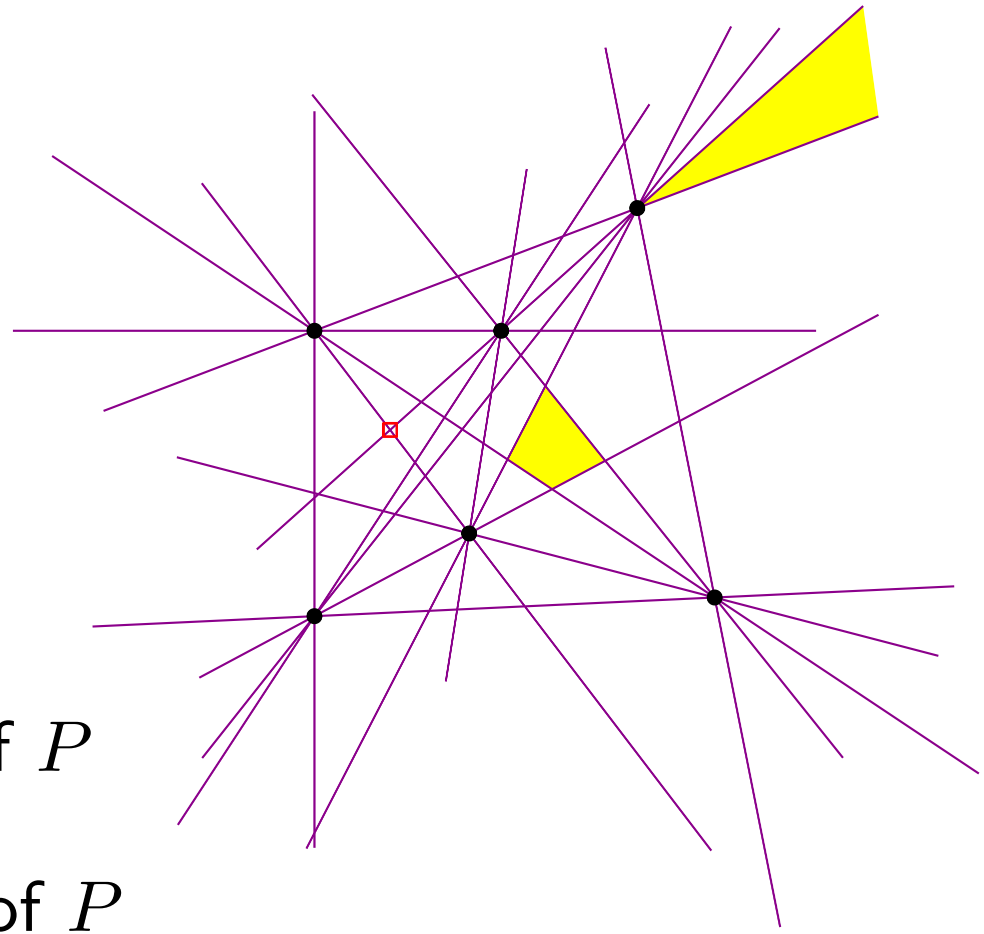
P induces many structures.



- $\star$ the *convex hull* of P
- $\star$ the *Delaunay triangulation* of P
- $\star$ the *minimum spanning tree* of P

Geometric data structures

Let P be a set of n points...



P induces many structures.

- $\star$ the *convex hull* of P
- $\star$ the *Delaunay triangulation* of P
- $\star$ the *minimum spanning tree* of P
- $\star$ the *arrangement* of the lines spanned by P

Complexity analysis

Predict the size of a data structure.

Complexity analysis

Predict the size of a data structure.

in the worst-case

function of n (# of points), d (dimension)...

"Input points" considered over $\mathbb{R}^d$

Complexity analysis

Predict the size of a data structure.

- in the worst-case
- function of n (# of points), d (dimension)...
- "Input points" considered over $\mathbb{R}^d$

Convex hull: $\Theta\left(n^{\lfloor d/2 \rfloor}\right)$

Delaunay triangulation: $\Theta\left(n^{\lceil d/2 \rceil}\right)$

Minimum spanning tree: $n-1$

Arrangement of induced hyperplanes: $\Theta\left(n^{d^2}\right)$

First limit: structure

Worst-case bounds often pessimistic in practical situations, when data is structured.

First limit: structure

Worst-case bounds often pessimistic in practical situations, when data is structured.

Ex: Near-linear behavior of 3D Delaunay triangulation in reconstruction applications.

First limit: structure

Worst-case bounds often pessimistic in practical situations, when data is structured.

Ex: Near-linear behavior of 3D Delaunay triangulation in reconstruction applications.

Refine the model to make it more realistic.

Find properties that rule out standard lower bounds.

First limit: structure

Worst-case bounds often pessimistic in practical situations, when data is structured.

Ex: Near-linear behavior of 3D Delaunay triangulation in reconstruction applications.

Refine the model to make it more realistic.

Find properties that rule out standard lower bounds.

Often difficult and application-specific.

First limit: structure

Worst-case bounds often pessimistic in practical situations, when data is structured.

Ex: Near-linear behavior of 3D Delaunay triangulation in reconstruction applications.

Refine the model to make it more realistic.

Find properties that rule out standard lower bounds.

Often difficult and application-specific.

Still no good model of "computer graphic scene"...

First limit: structure

Example of 3D Delaunay triangulation:

First limit: structure

Example of 3D Delaunay triangulation:

Spread-sensitive bounds
[E 03], [E 05]

First limit: structure

Example of 3D Delaunay triangulation:

Spread-sensitive bounds

[E 03], [E 05]

"*Nice*" *sample* of a "nice" surface

[GN 02], [GN03], [ABL 03], [DEG 08], [AAD 09]

First limit: structure

Example of 3D Delaunay triangulation:

Spread-sensitive bounds

[E 03], [E 05]

"*Nice*" *sample* of a "nice" surface

[GN 02], [GN03], [ABL 03], [DEG 08], [AAD 09]

Size of DT $\leftrightarrow$ "dimension" of the point set.

Second limit: precision

Input points given in finite-precision arithmetic.

Second limit: precision

Input points given in finite-precision arithmetic.

Can lower bounds be realized over $\mathbb{Q}$ (rather than $\mathbb{R}$)?

Second limit: precision

Input points given in finite-precision arithmetic.

Can lower bounds be realized over $\mathbb{Q}$ (rather than $\mathbb{R}$)?

Low complexity forbids certain configurations from arising.

Second limit: precision

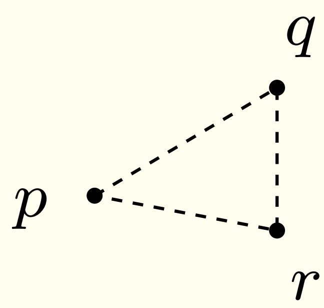
Input points given in finite-precision arithmetic.

Can lower bounds be realized over $\mathbb{Q}$ (rather than $\mathbb{R}$)?

Low complexity forbids certain configurations from arising.

The order-type of n points $\simeq$ the map $f : \binom{n}{3} \to \{L, R\}$

$$f(p, q, r) = R \text{ and } f(p, r, q) = L.$$



There exists an order type s. t. all its realization have exponential bit complexity [PS 89].

Second limit: precision

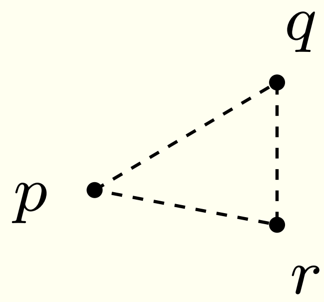
Input points given in finite-precision arithmetic.

Can lower bounds be realized over $\mathbb{Q}$ (rather than $\mathbb{R}$)?

Low complexity forbids certain configurations from arising.

The order-type of n points $\simeq$ the map $f : \binom{n}{3} \to \{L, R\}$

$$f(p, q, r) = R \text{ and } f(p, r, q) = L.$$



There exists an order type s. t. all its realization have exponential bit complexity [PS 89].

Data is often noisy, lower-bounds are often carefully designed.

Problems

Geometric structures defined by finite-precision data?

Robustness of lower/upper bounds to noise?

Smoothed complexity analysis

General principles

Geometric probabilities

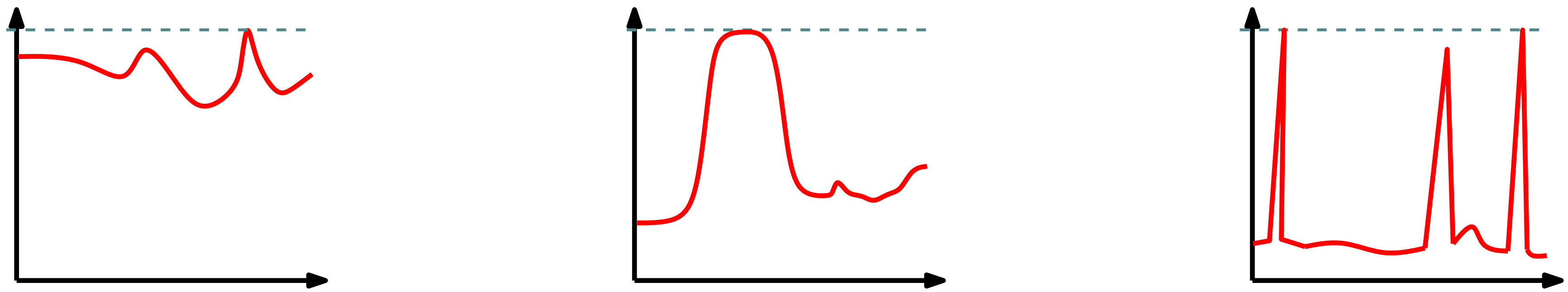
Shape matters

Smoothed complexity

"Worst-case complexity" = max. of the complexity function.

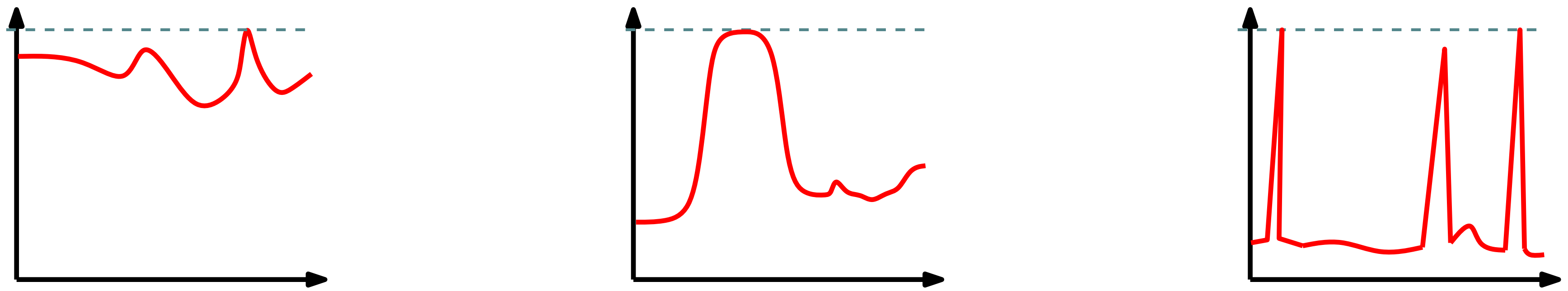
Smoothed complexity

"Worst-case complexity" = max. of the complexity function.



Smoothed complexity

"Worst-case complexity" $=$ max. of the complexity function.

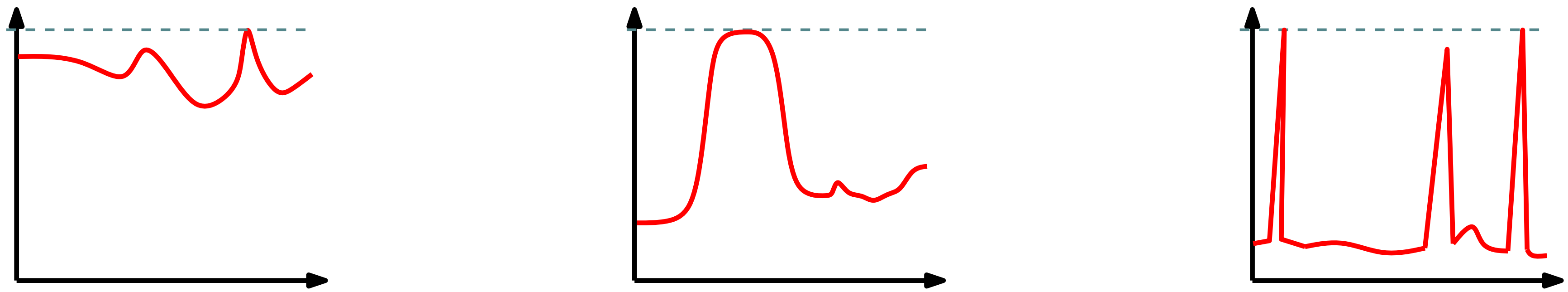


"Smoothed complexity(x)" $=$ Expected complexity over small perturbations of x.

"Smoothed complexity" $= \max_x$ smoothed complexity(x)

Smoothed complexity

"Worst-case complexity" $=$ max. of the complexity function.



"Smoothed complexity(x)" $=$ Expected complexity over small perturbations of x.

"Smoothed complexity" $= \max_x$ smoothed complexity(x)

$\simeq$ Convolution with a distribution concentrated near the origin.

Smoothed complexity

Introduced by Spielman and Teng [ST 04].

Smoothed complexity

Introduced by Spielman and Teng [ST 04].

Smoothed complexity of the simplex alg. is polynomial.

Smoothed complexity

Introduced by Spielman and Teng [ST 04].

Smoothed complexity of the simplex alg. is polynomial.

Increasing number of problems analyzed

`http://www.cs.yale.edu/homes/spielman/SmoothedAnalysis/`

Smoothed complexity

Introduced by Spielman and Teng [ST 04].

Smoothed complexity of the simplex alg. is polynomial.

Increasing number of problems analyzed

`http://www.cs.yale.edu/homes/spielman/SmoothedAnalysis/`

Gödel prize in 2008

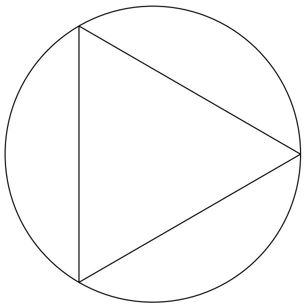
Geometric probabilities

Probabilities over geometric objects must be defined carefully.

Geometric probabilities

Probabilities over geometric objects must be defined carefully.

Bertrand's paradox

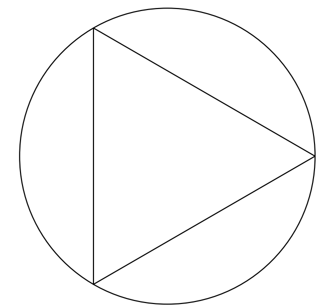


Prob. that "random" chord longer than $\sqrt{3}$?

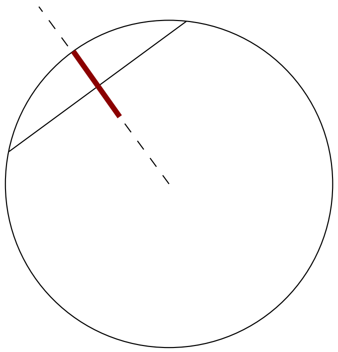
Geometric probabilities

Probabilities over geometric objects must be defined carefully.

Bertrand's paradox



Prob. that "random" chord longer than $\sqrt{3}$?

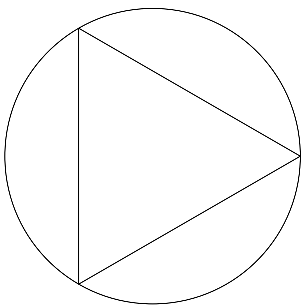


$1/2$

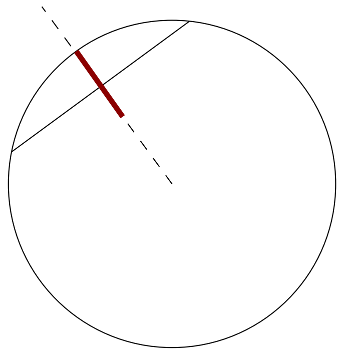
Geometric probabilities

Probabilities over geometric objects must be defined carefully.

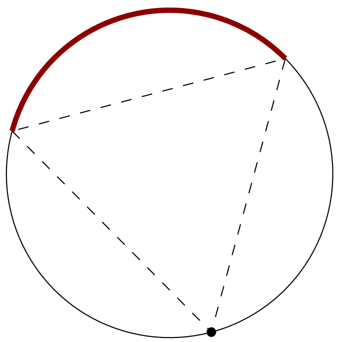
Bertrand's paradox



Prob. that "random" chord longer than $\sqrt{3}$?



$1/2$



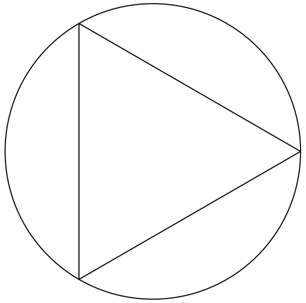
$1/3$

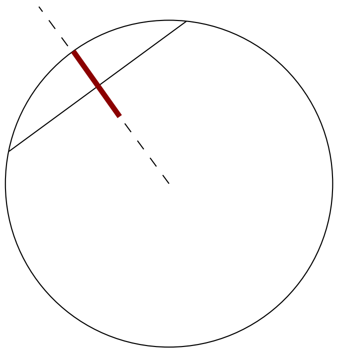
Geometric probabilities

Probabilities over geometric objects must be defined carefully.

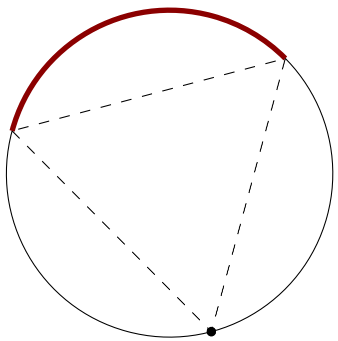
Bertrand's paradox

Prob. that "random" chord longer than $\sqrt{3}$?

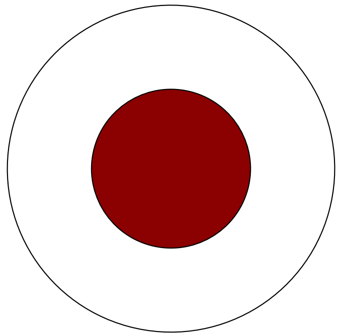




$1/2$



$1/3$



$1/4$

Shape matters!

X a set of n points chosen randomly, indep. with uniform distribution on a domain D.

Expected $\#$ of vertices on the convex hull of X?

Shape matters!

X a set of n points chosen randomly, indep. with uniform distribution on a domain D.

Expected # of vertices on the convex hull of X?

D is a disk in the plane: $\Theta\left(n^{1/3}\right)$.

[RS 63-64][R 70]

Shape matters!

X a set of n points chosen randomly, indep. with uniform distribution on a domain D.

Expected # of vertices on the convex hull of X?

D is a disk in the plane: $\Theta\left(n^{1/3}\right)$.

[RS 63-64][R 70]

D is a convex k-gon in the plane: $\Theta(k \log n)$.

Shape matters!

X a set of n points chosen randomly, indep. with uniform distribution on a domain D.

Expected # of vertices on the convex hull of X?

D is a disk in the plane: $\Theta\left(n^{1/3}\right)$.

[RS 63-64][R 70]

D is a convex k-gon in the plane: $\Theta(k \log n)$.

Intuition: points chosen "close to corners" dominate many points.

Number of extreme points

The *shuffled convex hull* game

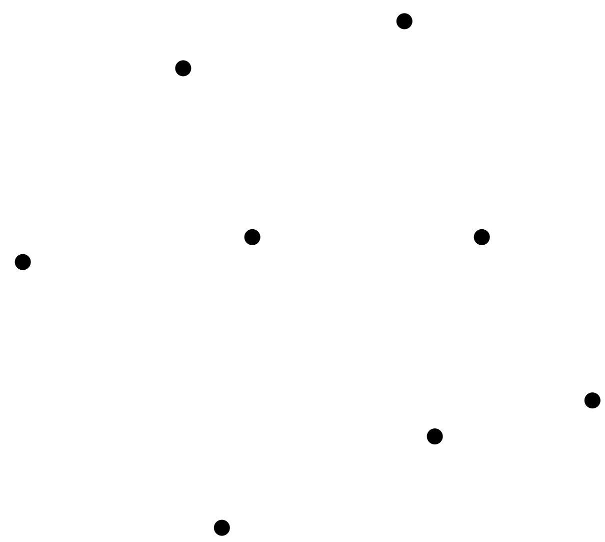
Our results

Comparison to "experimental" data

Shuffled convex hull game

Let X be a set of n points in some fixed domain D.

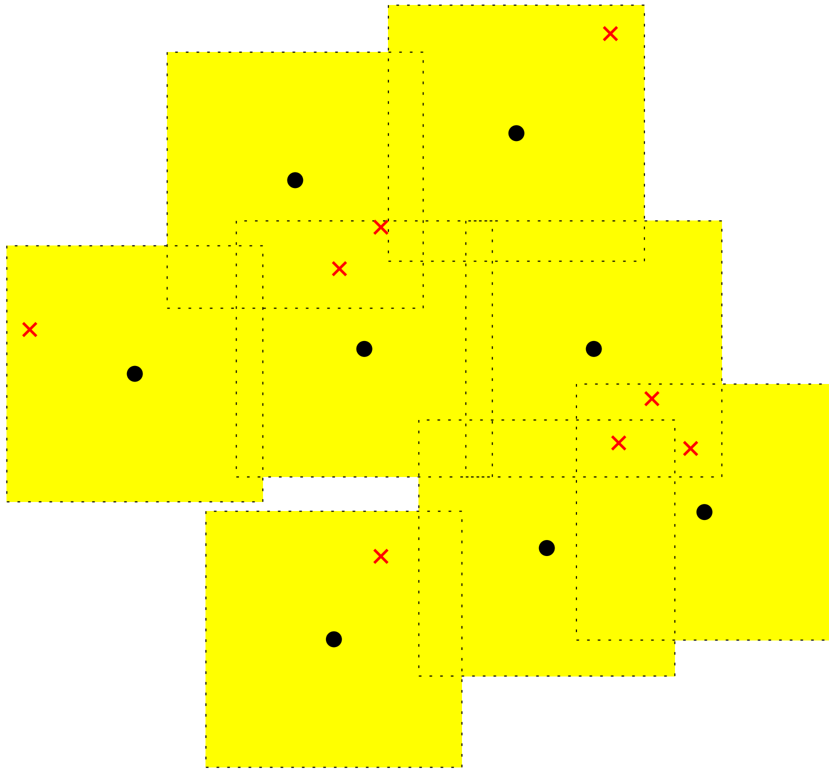
Y a L^p perturbation of X of amplitude δ.



Shuffled convex hull game

Let X be a set of n points in some fixed domain D.

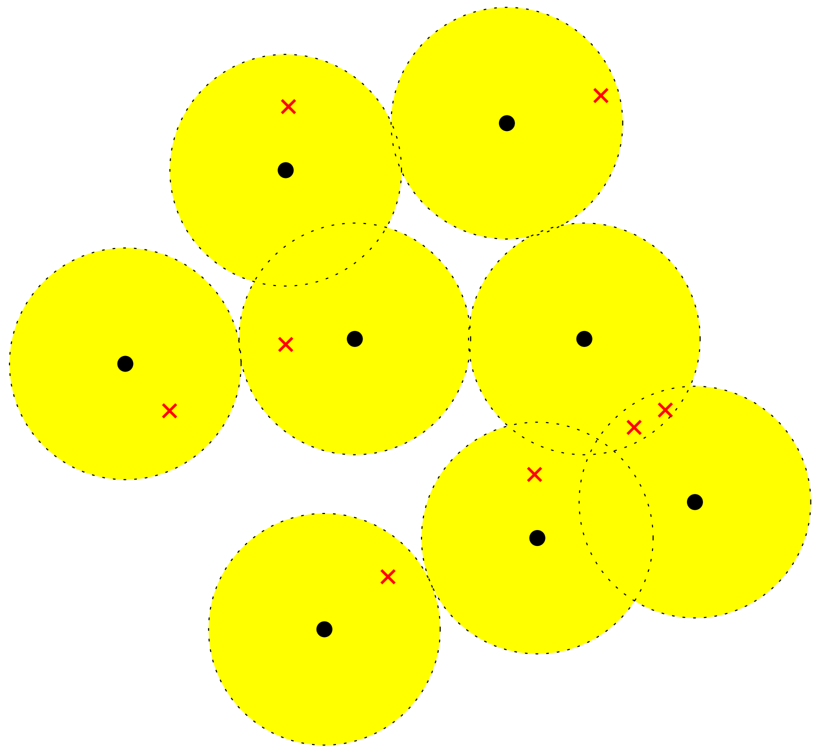
Y a L^p perturbation of X of amplitude δ.



Shuffled convex hull game

Let X be a set of n points in some fixed domain D.

Y a L^p perturbation of X of amplitude δ.

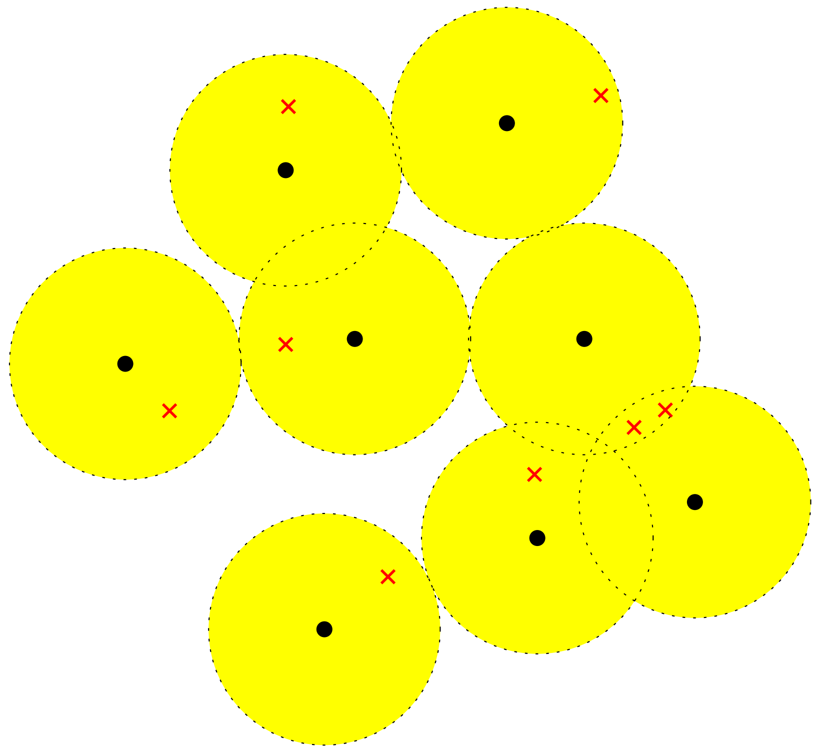


Shuffled convex hull game

Let X be a set of n points in some fixed domain D.

Y a L^p perturbation of X of amplitude δ.

Given ϵ and p, find X such that $E[\#CH(Y)]$ is max.



Different D, different answers?

Our results

X = regular n-gon inscribed in the unit circle.

Our results

$X =$ regular n-gon inscribed in the unit circle.

Near-tight estimates of $E[\#CH(Y)]$ for $p = 1, 2$ and ∞.

Our results

X = regular n-gon inscribed in the unit circle.

Near-tight estimates of $E[\#CH(Y)]$ for $p = 1, 2$ and ∞.

perturbation	$E[\#CH(Y)]$	range of δ
L^1, L^∞	$\tilde{\Theta}\left(n^{1/5}\delta^{-2/5}\right)$	$\delta \in \left(\tilde{\Omega}(1/n^2), O(1)\right)$
L^2	$\tilde{\Theta}\left(n^{1/4}\delta^{-3/8}\right)$	

Comparison to exp. data

Uniform noise simulated by pseudo-random number generators.

Comparison to exp. data

Uniform noise simulated by pseudo-random number generators.

regular n-gon → perturbed point set

1. rounding to the *double* grid
2. random perturbation of the coordinates

Comparison to exp. data

Uniform noise simulated by pseudo-random number generators.

regular n-gon → perturbed point set

1. rounding to the *double* grid
2. random perturbation of the coordinates

$n = 10^i$ for $i = 3, \ldots, 7$

$\delta = 10^j$ for $j = -7, \ldots, 5$

average over 1000 - 100 trials

Comparison to exp. data

Uniform noise simulated by pseudo-random number generators.

regular n-gon ⟶ perturbed point set

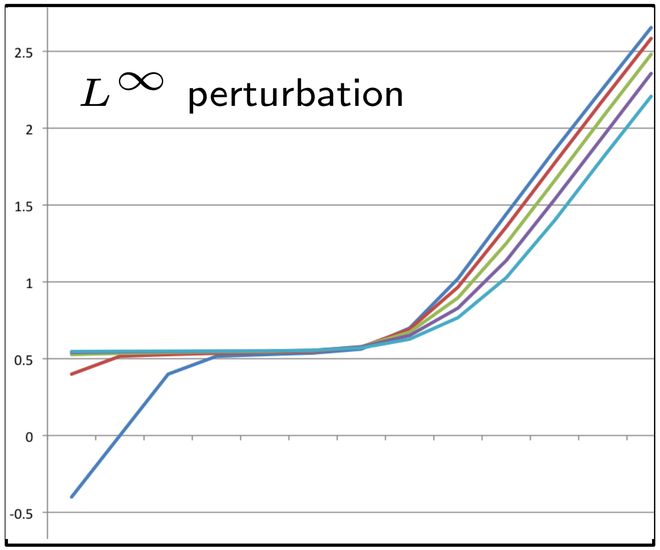
1. rounding to the *double* grid
2. random perturbation of the coordinates

$n = 10^i$ for $i = 3, \ldots, 7$

$\delta = 10^j$ for $j = -7, \ldots, 5$

average over 1000 - 100 trials

Plot $\log_{10} \#CH(Y) - \log_{10}(n^{1/5}\delta^{-2/5})$



Comparison to exp. data

Uniform noise simulated by pseudo-random number generators.

regular n-gon ⟶ perturbed point set

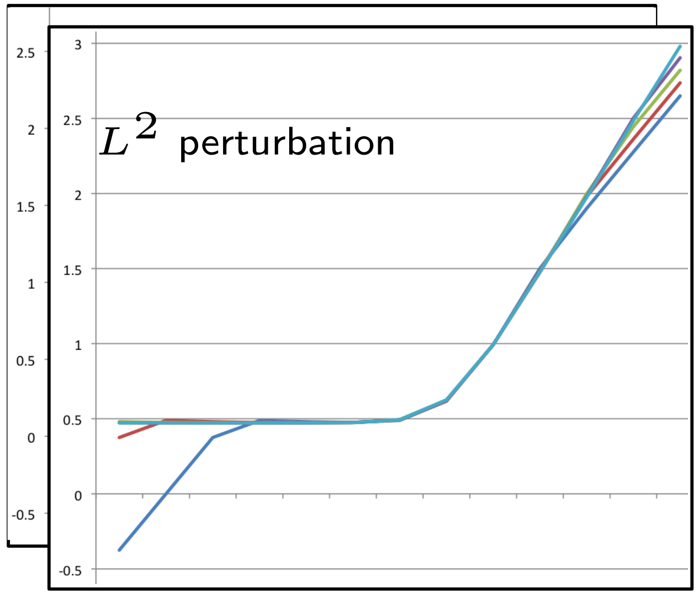
1. rounding to the *double* grid
2. random perturbation of the coordinates

$$n = 10^i \text{ for } i = 3, \dots, 7$$

$$\delta = 10^j \text{ for } j = -7, \dots, 5$$

average over 1000 - 100 trials

Plot $\log_{10} \#CH(Y) - \log_{10}(n^{1/4}\delta^{-3/8})$



Comparison to exp. data

Effect of rounding coordinates to a coarse grid.

regular n-gon ⟶ rounded point set

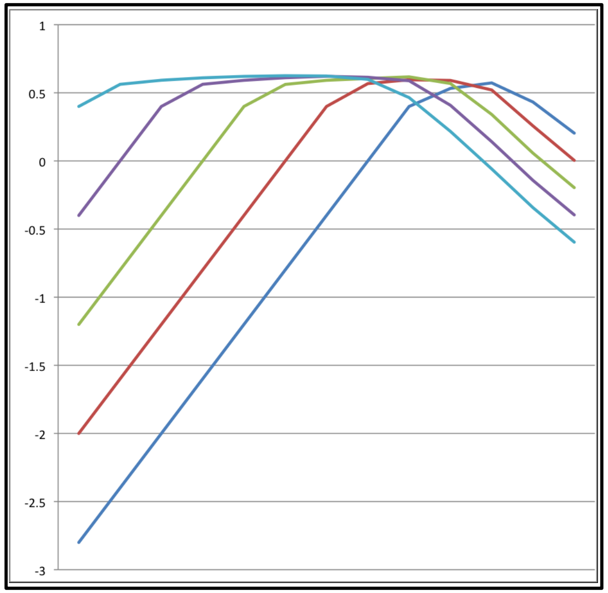
1. rounding to the *double* grid
2. rounding to the *float* grid.

$$n = 10^i \text{ for } i = 3, \ldots, 7$$

$$\delta = 10^j \text{ for } j = -7, \ldots, 5$$

average over 1000 - 100 trials

Plot $\log_{10} \#CH(Y) - \log_{10}(n^{1/5}\delta^{-2/5})$



To summarize

Near-tight bounds for L^2 and L^∞.

Predicts the behavior of regular rounding quite accurately.

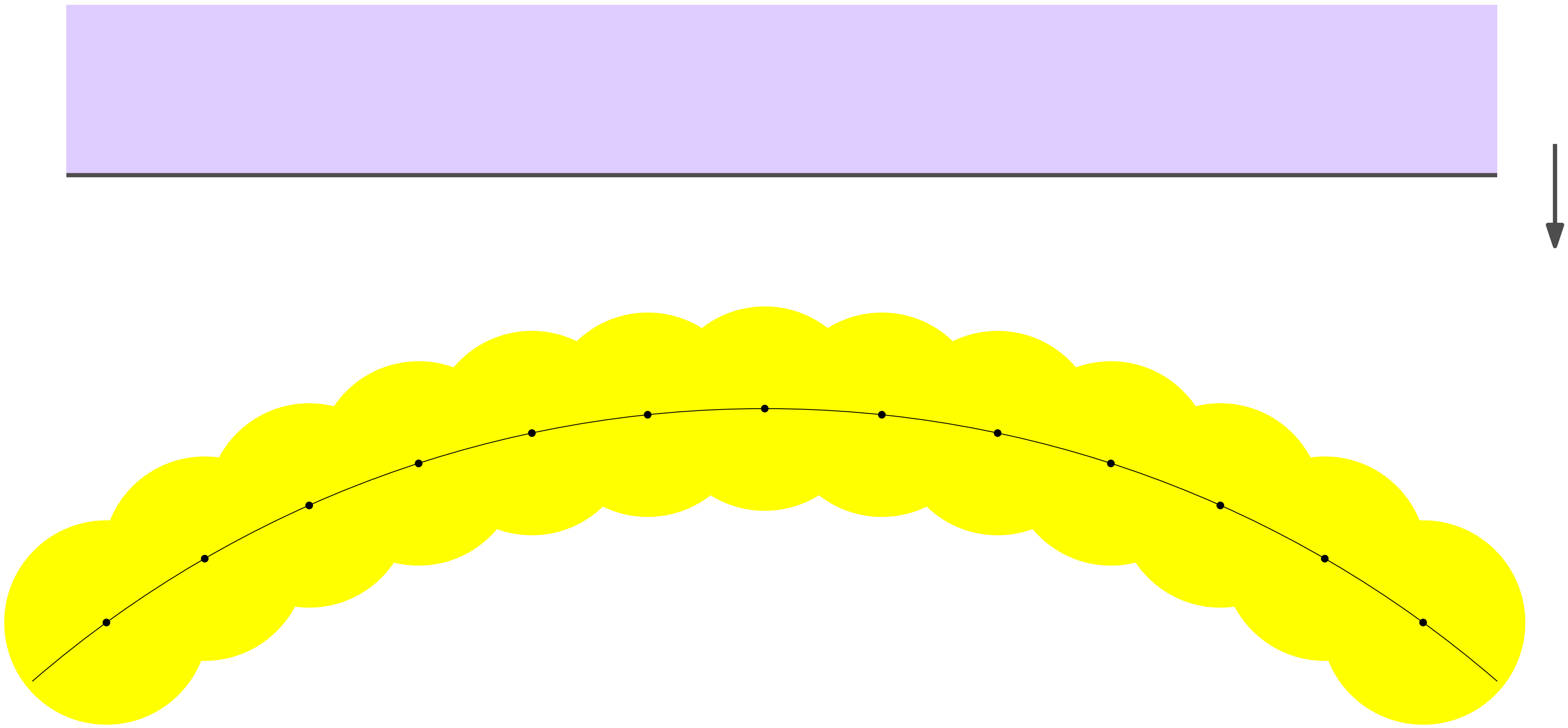
Generalizes to arbitrary dimension for L^2 perturbations.

A look under the hood

Witnesses and collectors

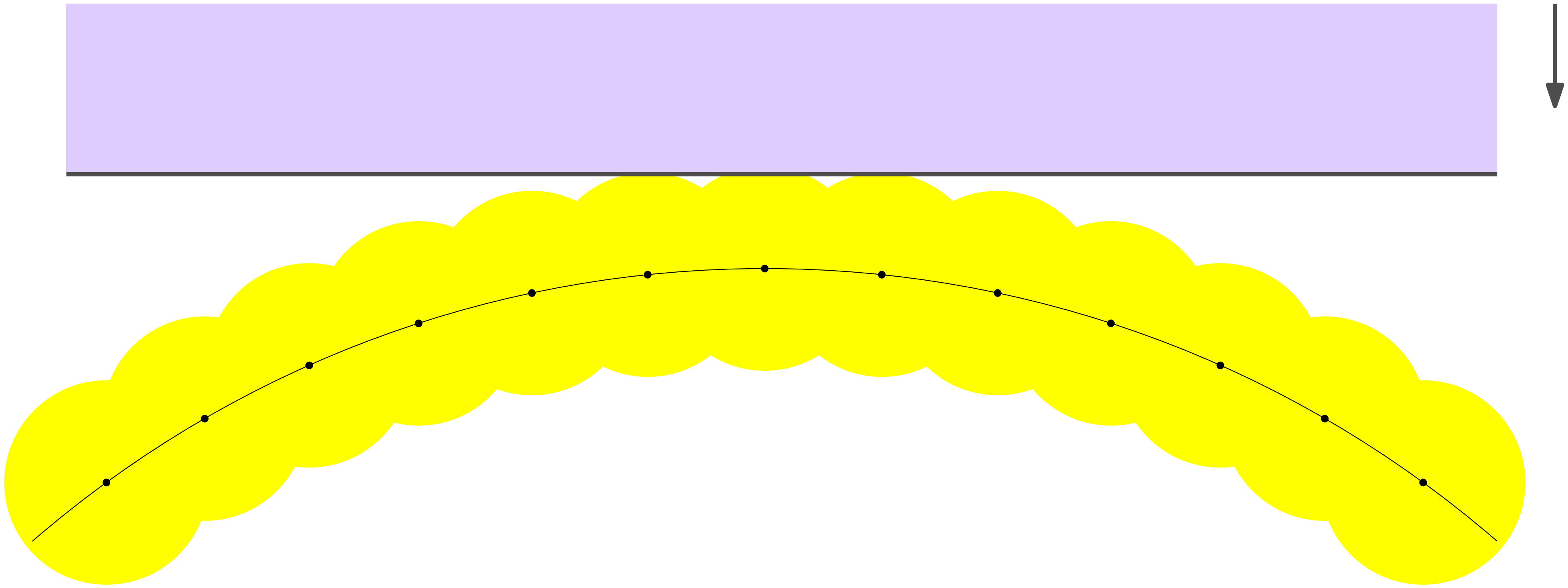
Shape matters (bis)

Principle



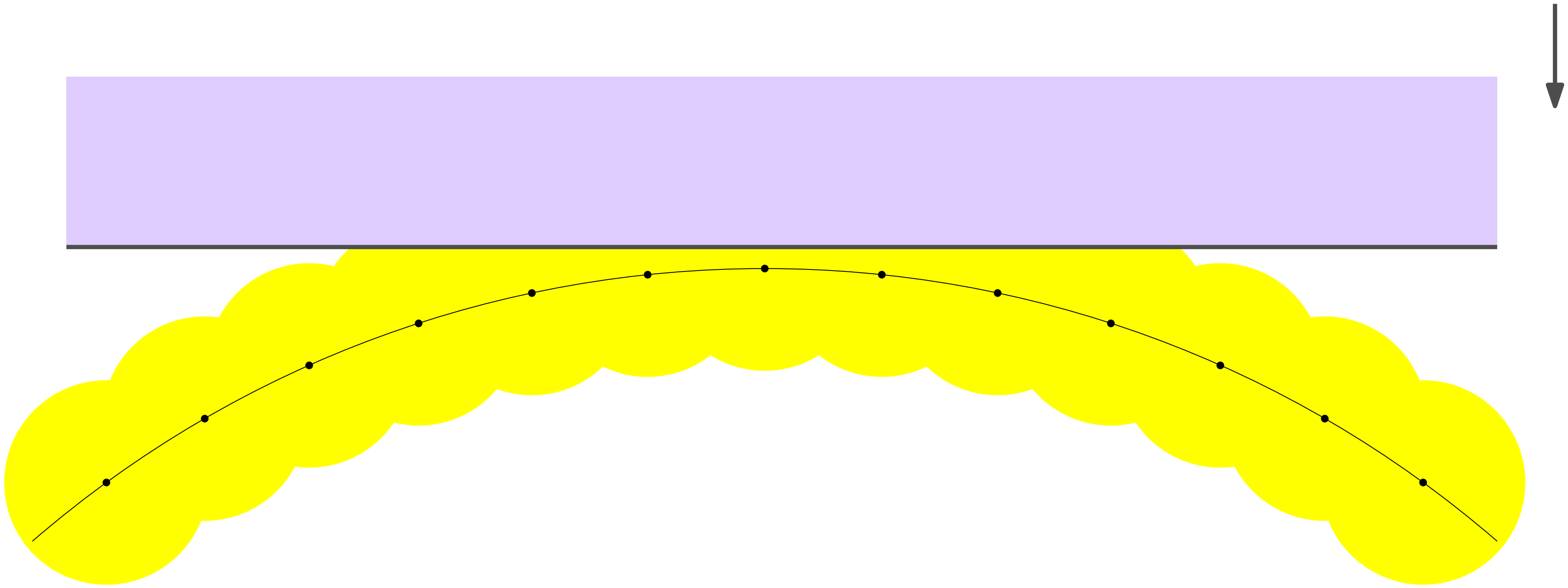
Push a halfplane and count the average # of points it contains.

Principle



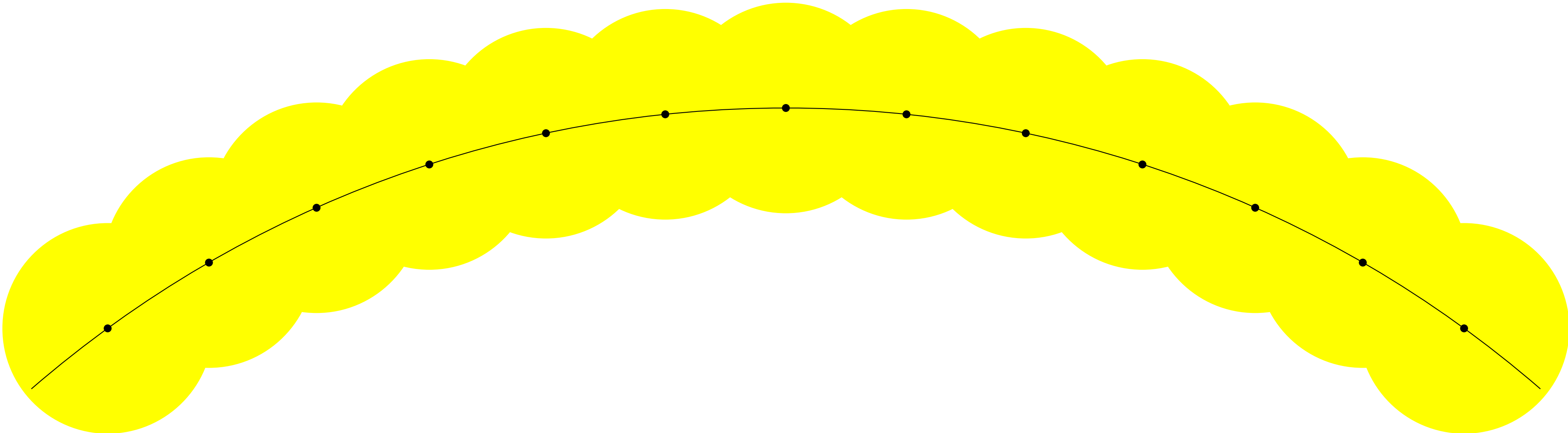
Push a halfplane and count the average # of points it contains.

Principle

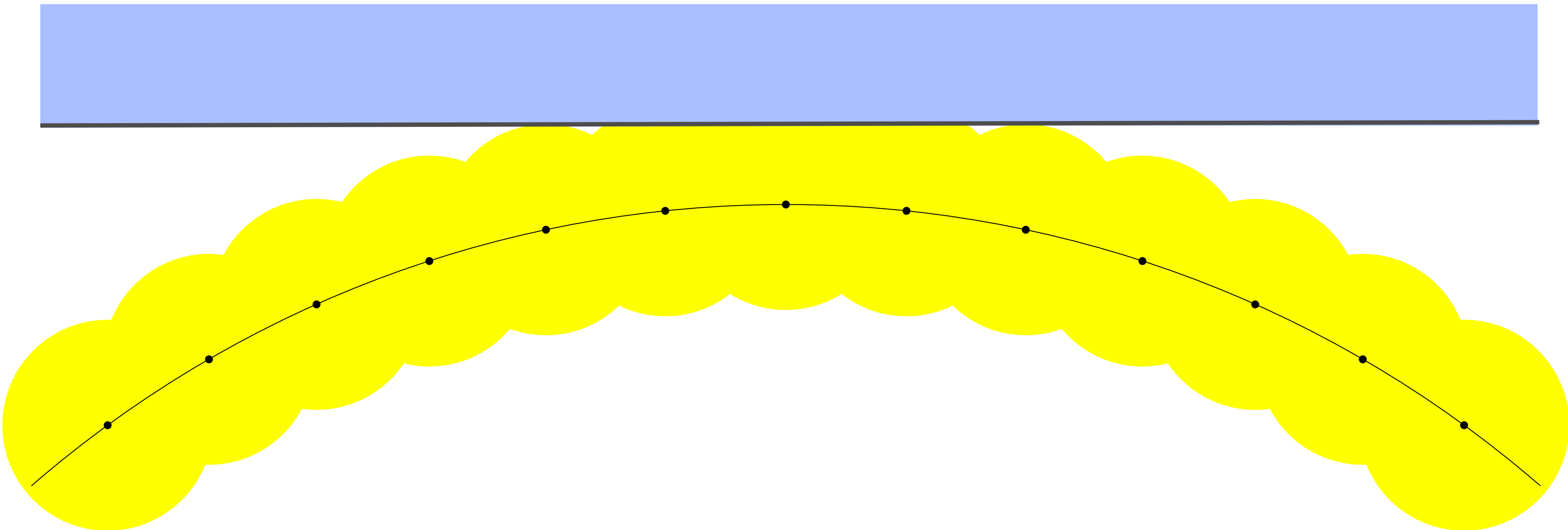


Push a halfplane and count the average # of points it contains.

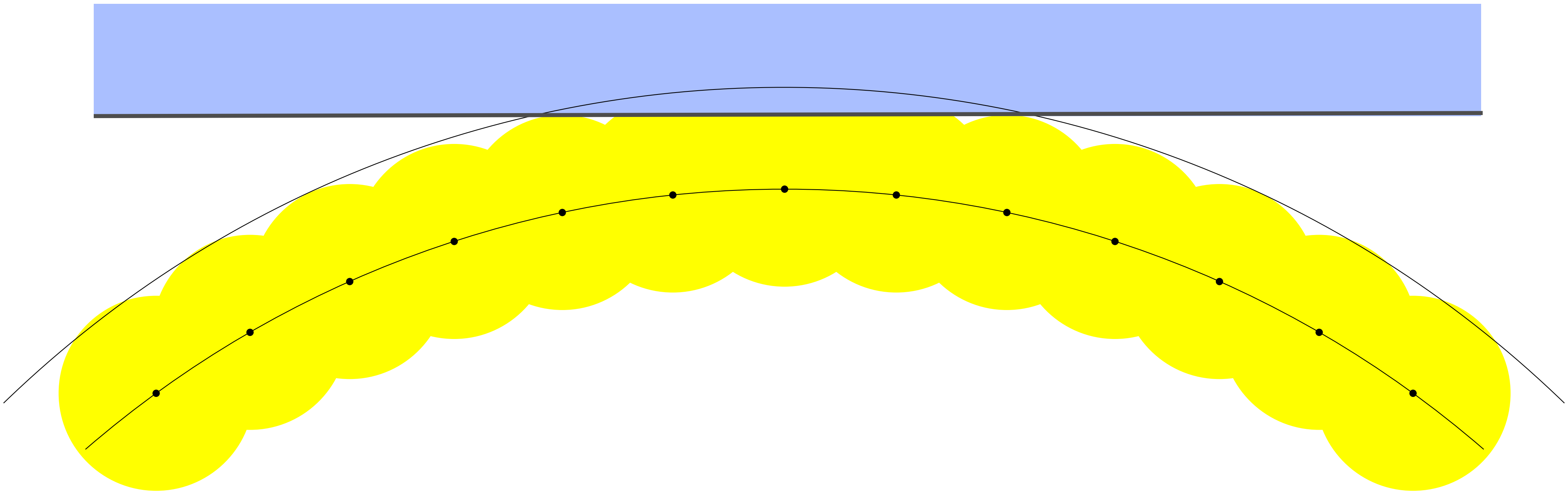
Principle



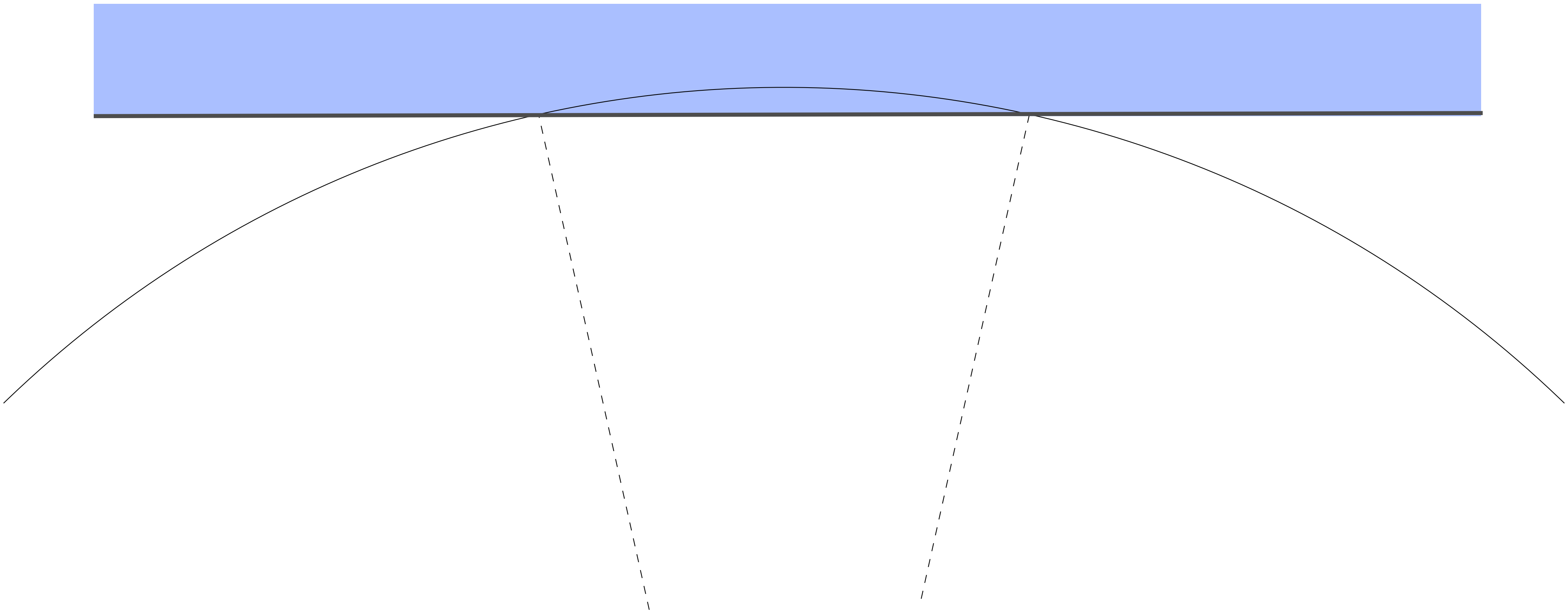
Principle



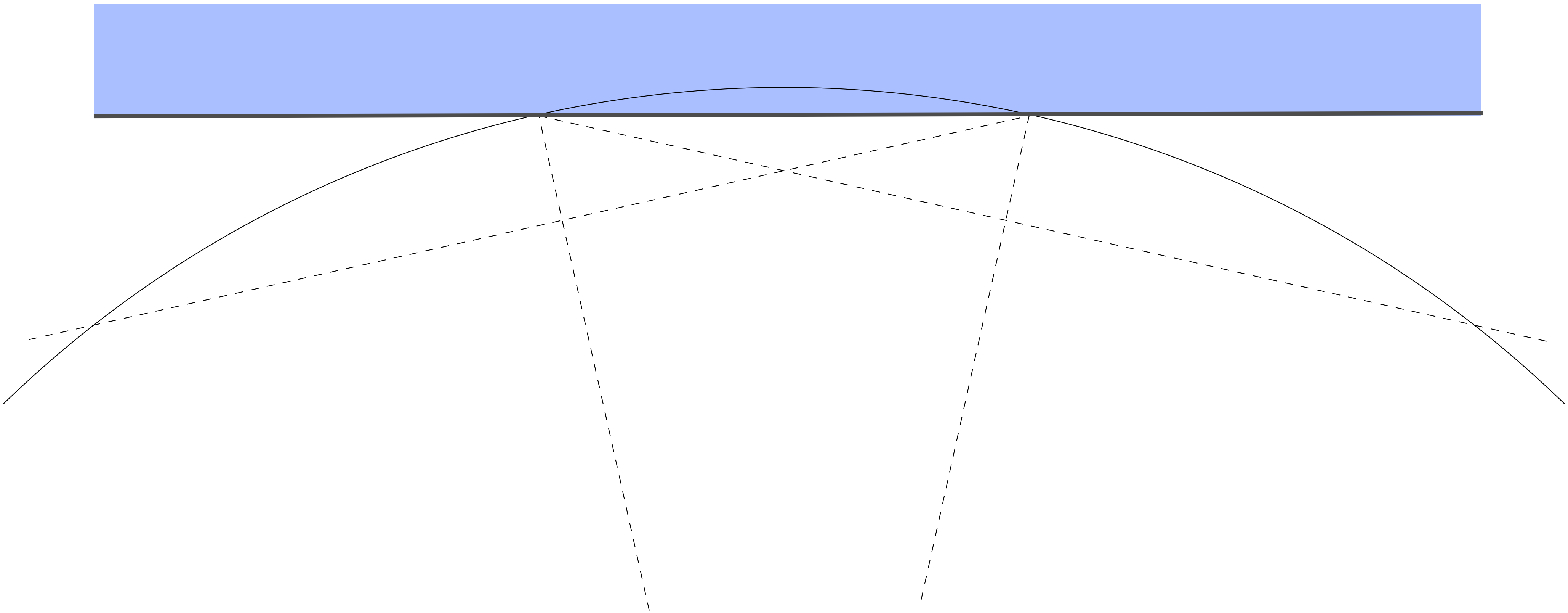
Principle



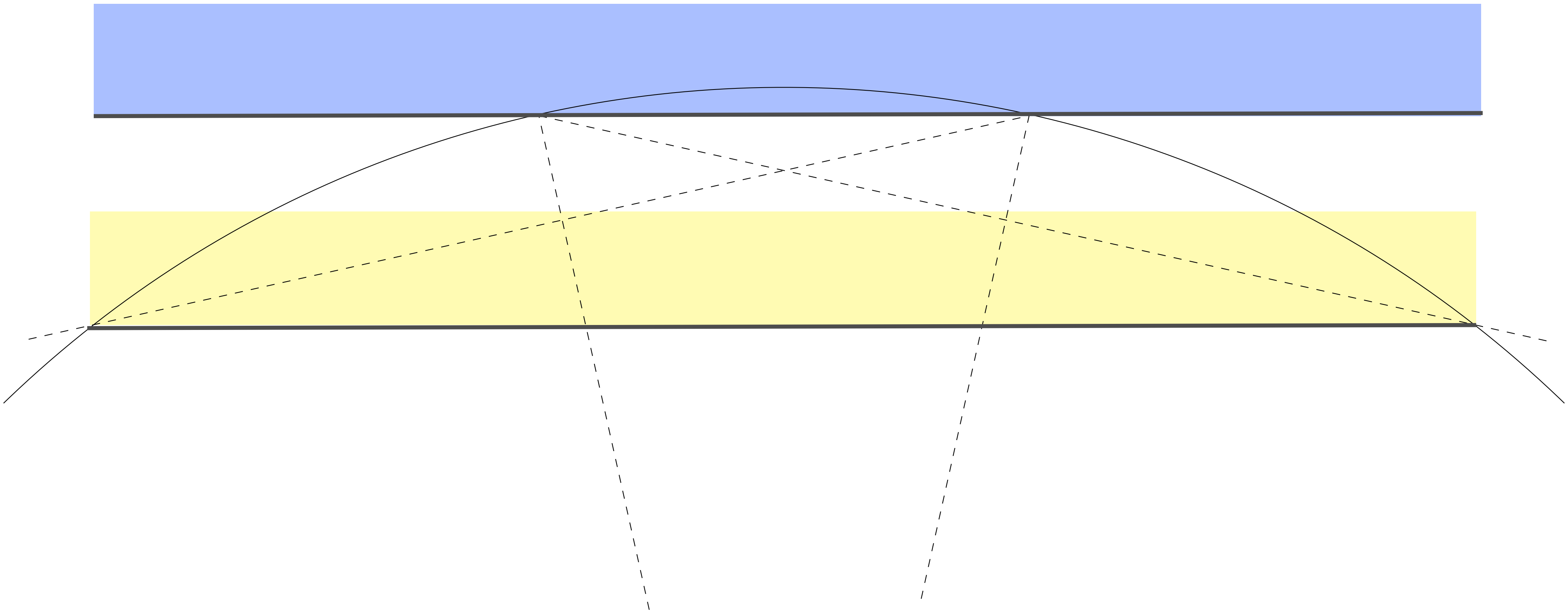
Principle



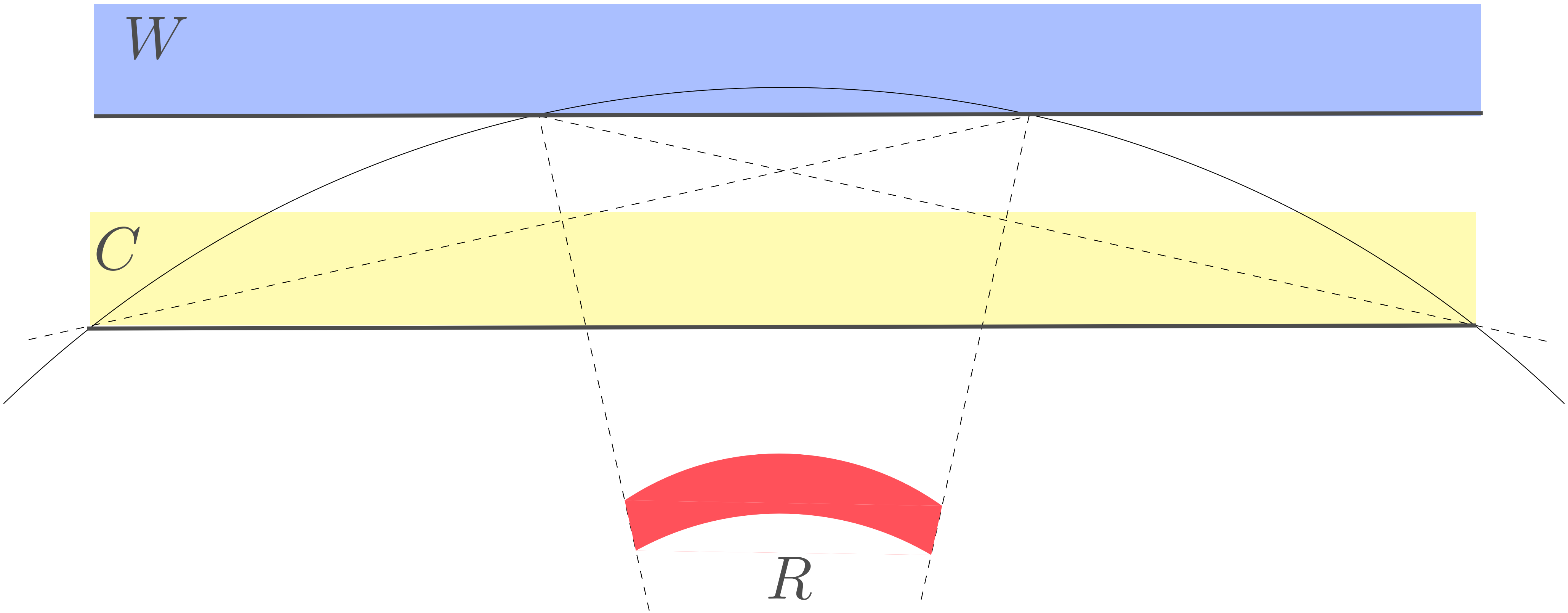
Principle



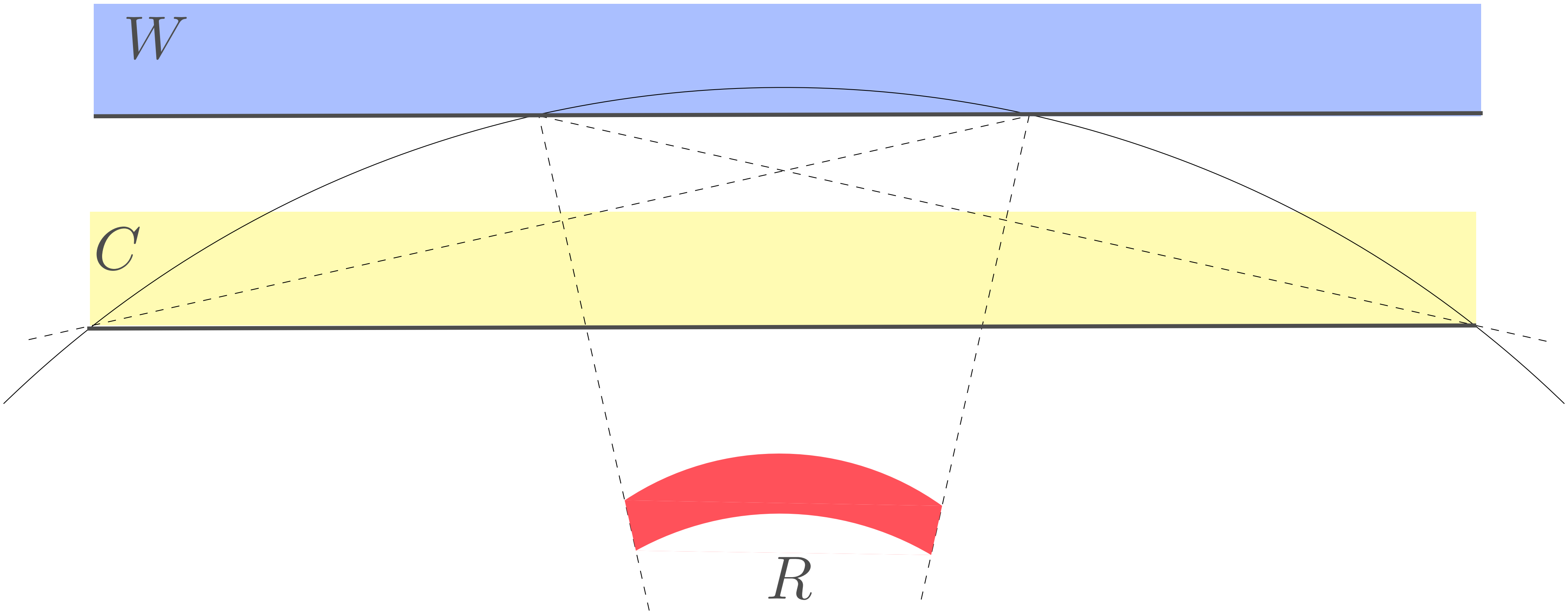
Principle



Principle



Principle



Any point in W dominates any point outside C for any direction in R.

Witnesses & collectors

Place enough pairs (C, W) so that the R's cover $\mathbb{S}^1$...

Witnesses & collectors

Place enough pairs (C, W) so that the R's cover $\mathbb{S}^1$...

$$E[\#CH(Y)] \leq ne^{-\Sigma E[W_i \cap Y]} + \Sigma E[C_i \cap Y]$$

Witnesses & collectors

Place enough pairs (C, W) so that the R's cover $\mathbb{S}^1$...

$$E[\#CH(Y)] \leq ne^{-\Sigma E[W_i \cap Y]} + \Sigma E[C_i \cap Y]$$

... while having any perturbation disk meet $O(1)$ W's.

Witnesses & collectors

Place enough pairs (C, W) so that the R's cover $\mathbb{S}^1$...

$$E[\#CH(Y)] \leq ne^{-\Sigma E[W_i \cap Y]} + \Sigma E[C_i \cap Y]$$

... while having any perturbation disk meet $O(1)$ W's.

$$\Omega\left(\Sigma(1 - e^{-E[W_i \cap Y]})\right) \leq E[\#CH(Y)]$$

Witnesses & collectors

Place enough pairs (C, W) so that the R's cover $\mathbb{S}^1$...

$$E[\#CH(Y)] \leq n e^{-\Sigma E[W_i \cap Y]} + \Sigma E[C_i \cap Y]$$

... while having any perturbation disk meet $O(1)$ W's.

$$\Omega\left(\Sigma(1 - e^{-E[W_i \cap Y]})\right) \leq E[\#CH(Y)]$$

Make sure that W is likely to be nonempty...

... while keeping $C \cap Y$ small.

Witnesses & collectors

Place enough pairs (C, W) so that the R's cover $\mathbb{S}^1$...

$$E[\#CH(Y)] \leq ne^{-\Sigma E[W_i \cap Y]} + \Sigma E[C_i \cap Y]$$

... while having any perturbation disk meet $O(1)$ W's.

$$\Omega\left(\Sigma(1 - e^{-E[W_i \cap Y]})\right) \leq E[\#CH(Y)]$$

Make sure that W is likely to be nonempty...

... while keeping $C \cap Y$ small.

Both $E[W_i \cap Y]$ and $E[C_i \cap Y]$ of order $\Theta(\log n)$

Witnesses & collectors

Place enough pairs (C, W) so that the R's cover $\mathbb{S}^1$...

$$E[\#CH(Y)] \leq ne^{-\Sigma E[W_i \cap Y]} + \Sigma E[C_i \cap Y]$$

... while having any perturbation disk meet $O(1)$ W's.

$$\Omega\left(\Sigma(1 - e^{-E[W_i \cap Y]})\right) \leq E[\#CH(Y)]$$

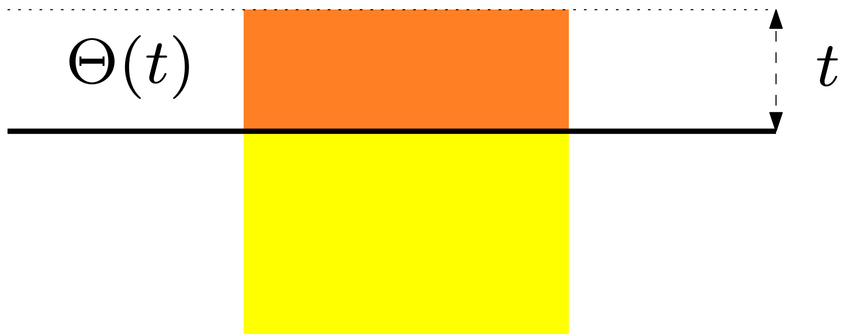
Make sure that W is likely to be nonempty...

... while keeping $C \cap Y$ small.

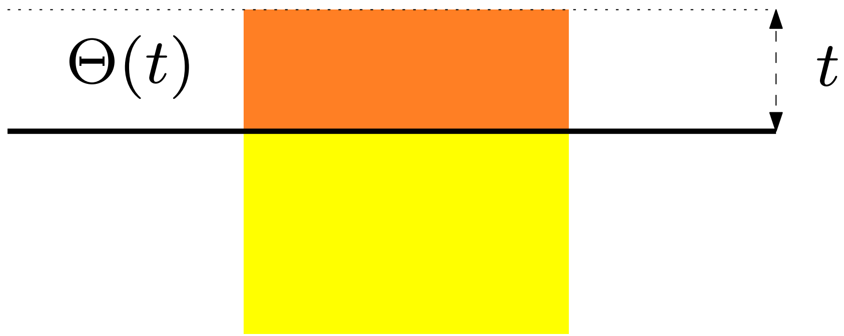
Both $E[W_i \cap Y]$ and $E[C_i \cap Y]$ of order $\Theta(\log n)$

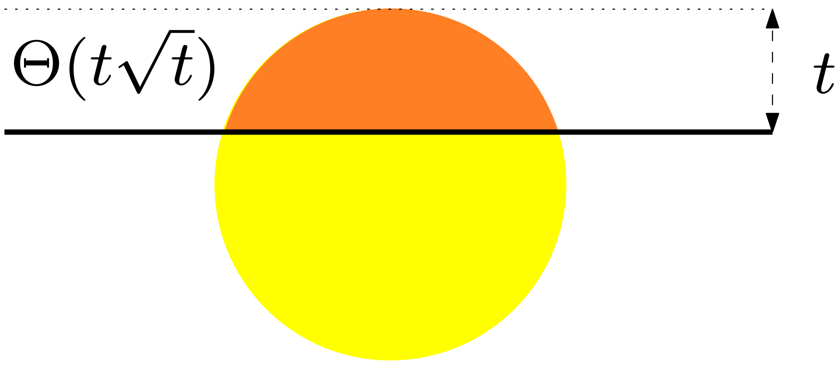
$$\Rightarrow E[\#CH(Y)] = \Theta\left(\#\text{ pairs }(C, W)\right)$$

Shape matters! (bis)

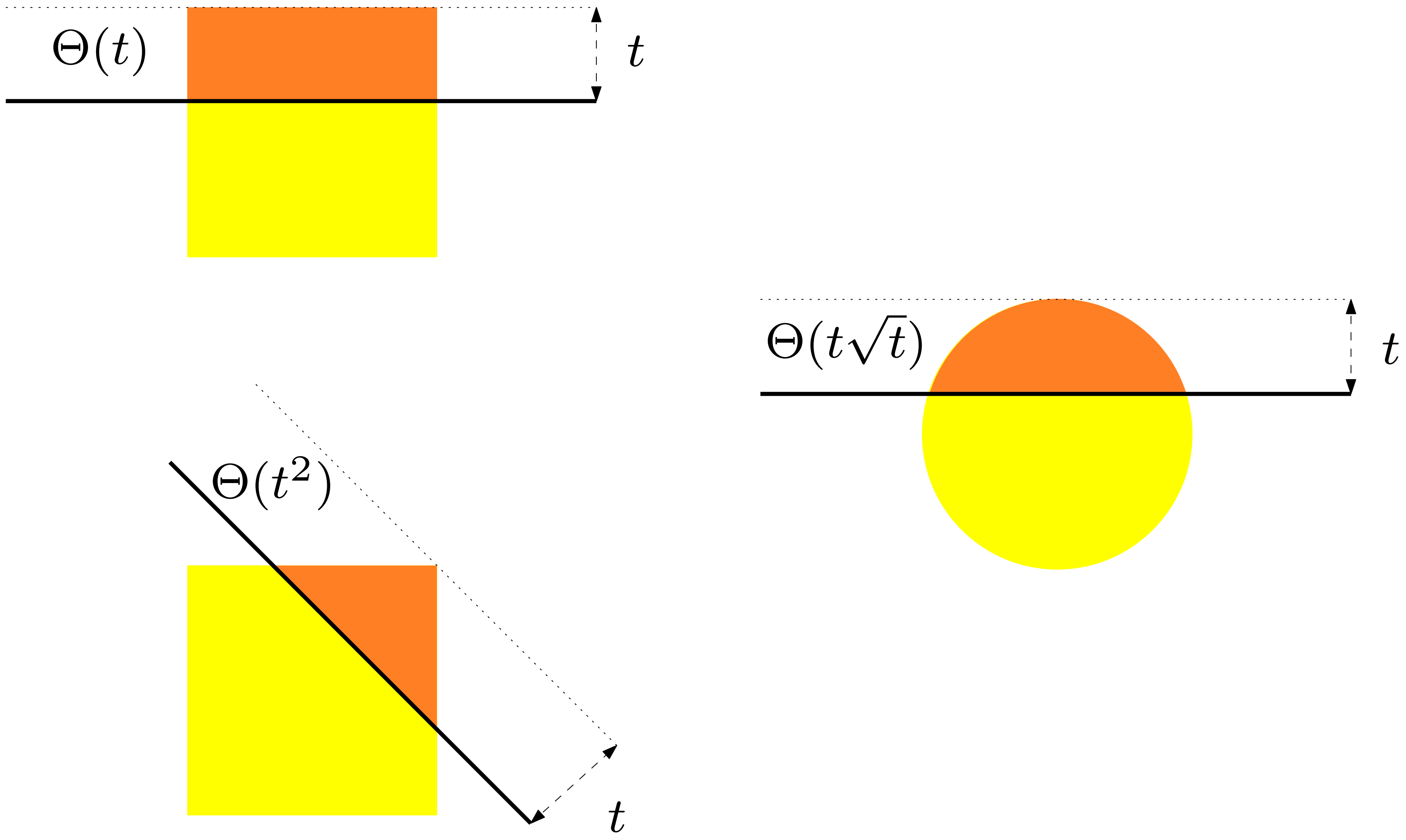


Shape matters! (bis)





Shape matters! (bis)



Conclusion

Conclusion

At circular shuffled convex hull game, for $\epsilon \gg n^{-2/9}$...

... all points at the origin win over the regular n-gon.

Conclusion

At circular shuffled convex hull game, for $\epsilon \gg n^{-2/9}$...

... all points at the origin win over the regular n-gon.

Are there other "extremal" configurations?

Conclusion

At circular shuffled convex hull game, for $\epsilon \gg n^{-2/9}$...

... all points at the origin win over the regular n-gon.

Are there other "extremal" configurations?

Smoothed analysis of other geometric data structures?

Conclusion

At circular shuffled convex hull game, for $\epsilon \gg n^{-2/9}$...

... all points at the origin win over the regular n-gon.

Are there other "extremal" configurations?

Smoothed analysis of other geometric data structures?

$\square$