

# Dispersion of Mass and the Complexity of Randomized Algorithms II

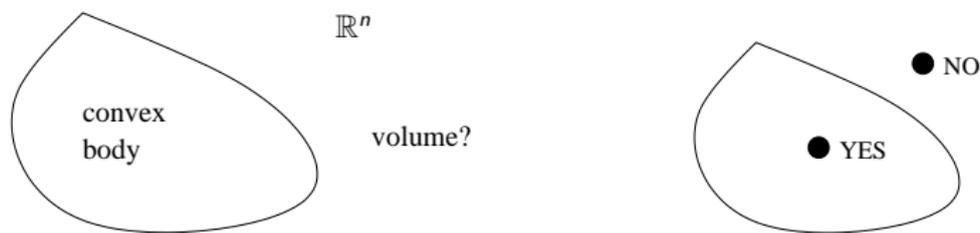
## ICMS

Luis Rademacher  
MIT

Santosh Vempala  
Georgia Tech and MIT

April 21, 2007

# Volume computation



- ▶  $\#P$ -hard for polytopes given as a list of vertices or halfspaces.
- ▶ There is a family of polytopes having rational halfspaces such that the volume is a rational  $a/b$  with  $b$  of length exponential in the input length.
- ▶ Standard model: membership oracle. Randomization provably helps.
- ▶ Because of hardness results, we forget exact computation and consider approximations up to a multiplicative factor.

## Known bounds for volume

- ▶ Deterministic lower bound:  
Elekes; Báráni, Füredi: polynomial number of queries  $\implies$   
error of a factor of  $\sim n^{n/2}$ .
- ▶ Upper bound: random walks:

Dyer, Frieze, Kannan 1991	$n^{23}$
Lovász, Simonovits 1990	$n^{16}$
Applegate, Kannan 1990	$n^{10}$
Lovász 1990	$n^{10}$
Dyer, Frieze 1991	$n^8$
Lovász, Simonovits 1993	$n^7$
Kannan et al. 1997	$n^5$
Lovász, Vempala 2003	$n^4$

This and the deterministic lower bound imply that randomization provably helps.

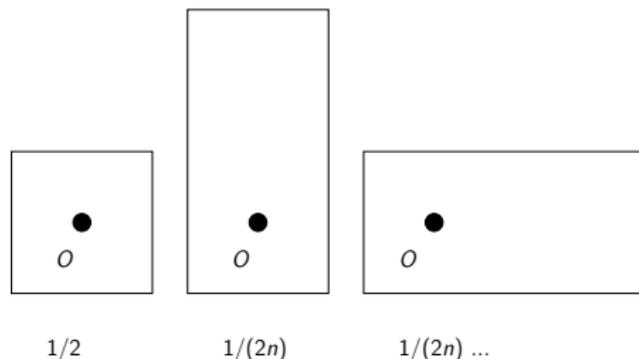
- ▶ The complexity of a randomized approximation of the volume depends on roundedness of the input. Our lower bound holds even for well-rounded convex bodies.

# Yao's lemma

- ▶ Yao's lemma:  
The probability of failure of a randomized algorithm on the worst input is at least the probability of failure of the best deterministic algorithm against some distribution.

# Easy lower bound for randomized algorithms

- ▶ Lower bound:  $\Omega(n)$ .
  - ▶ Distribution: Axis-aligned cube or brick.
  - ▶ Result: less than  $n - 1$  queries  $\implies$  fails to approximate volume up to a constant factor with probability at least  $1/2n$



## First idea for improved lower bound

- ▶ Instead of axis-aligned, consider all rotations of cube and brick,  $\sim n^2$  “degrees of freedom”.
- ▶ But given that our proof is in the flavor of information theory, we prefer the space of transformations to be a subset of Euclidean space with the usual measure. Instead of rotations, we allow all linear transformations of a cube, suitably restricted to satisfy roundness condition.  
     $\implies$  more variability in volume.

## Our results: lower bounds, volume, determinant

- ▶ Theorem (volume): Any randomized algorithm needs  $\Omega(n^2 / \log n)$  queries to approximate the volume up to a constant factor with probability  $1 - 1/\text{poly}(n)$  (even well-rounded).
- ▶ Actually, it is hard for parallelepipeds of the form: for  $A \in \mathbb{R}^{n \times n}$

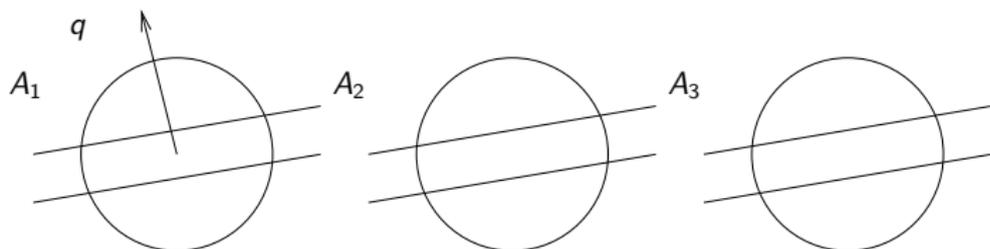
$$\{x \in \mathbb{R}^n : (\forall i) |A_i \cdot x| \leq 1\}$$

Oracle: given  $x \in \mathbb{R}^n$ , decide whether  $(\forall i) |A_i \cdot x| \leq 1$ .

Theorem (determinant): Any randomized algorithm needs  $\Omega(n^2 / \log n)$  queries to approximate  $|\det A|$  up to a constant factor with probability  $1 - 1/\text{poly}(n)$ .

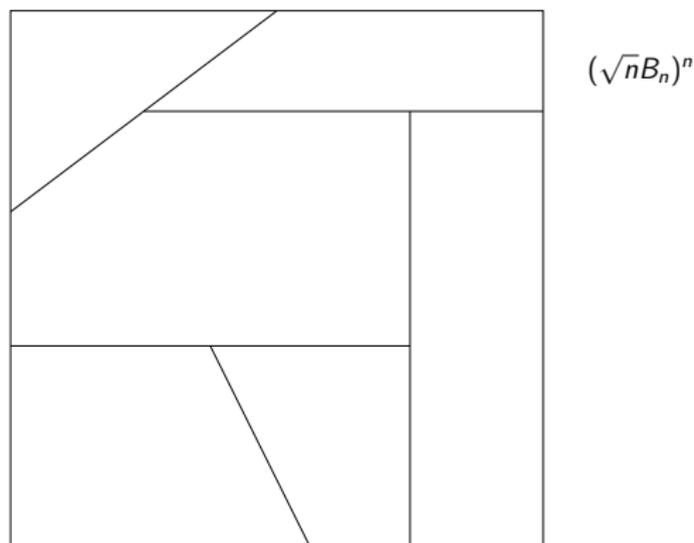
# Proof of the volume lower bound

- ▶ Yao's lemma: enough to prove a lower bound for deterministic algorithms against a distribution.
- ▶ Distribution: Parallelepipeds " $\|Ax\|_\infty \leq 1$ " given by matrices  $A$  with each row uniformly and independently from  $\sqrt{n}B_n$ .
- ▶ A deterministic algorithm can be seen as a decision tree  $\implies$  partition of the input space  $((\sqrt{n}B_n)^n)$ .



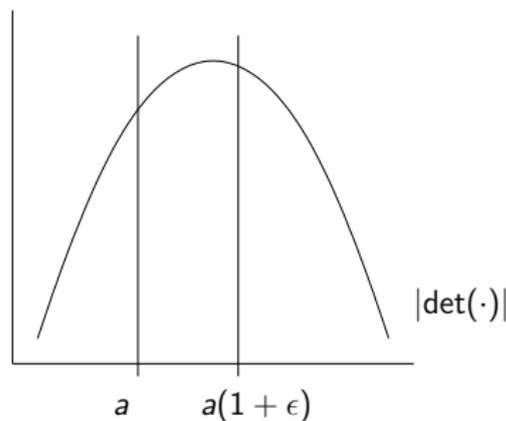
# Proof of the volume lower bound

- ▶ After making all its queries, what the algorithm knows is exactly on which part of the partition of  $(\sqrt{n}B_n)^n$  we are



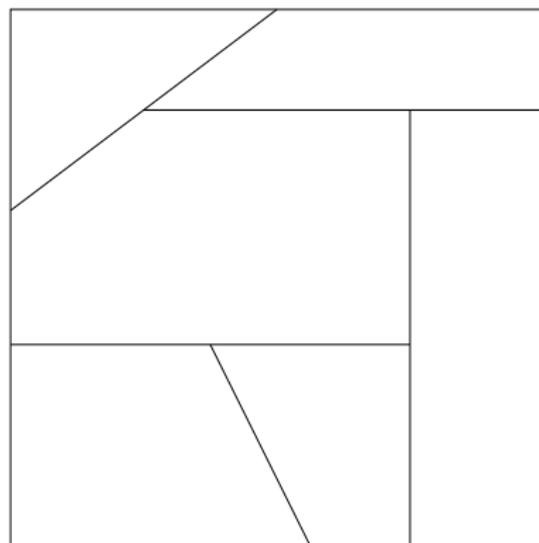
# Proof of the volume lower bound

- ▶ Failure of the algorithm  $\Leftrightarrow$  dispersion of  $|\det(\cdot)|$  on most parts.
- ▶ Constant dispersion: there is at least  $1/\text{poly}(n)$  mass outside of any interval of constant multiplicative length (i.e. of the form  $[a, a(1 + \epsilon)]$ ).



# Proof of the volume lower bound

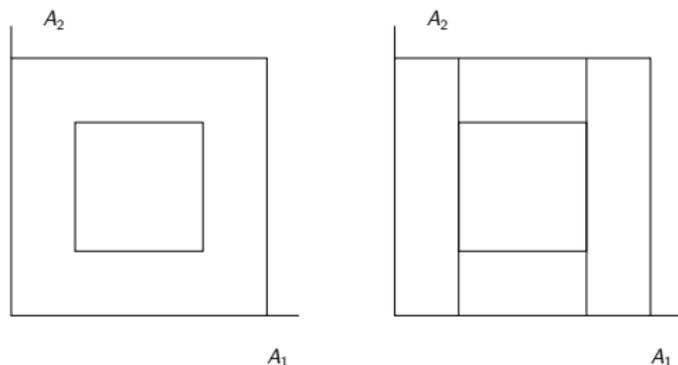
- ▶ Leaf  $\leftrightarrow$  part. # parts = # leafs  $\leq 2^{\text{depth}}$ .



$$(\sqrt{n}B_n)^n$$

# Proof of the volume lower bound

- ▶ Parts can be assumed to be “product sets along rows”.

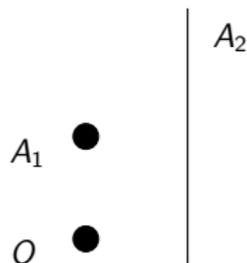


Implemented by modified oracle.



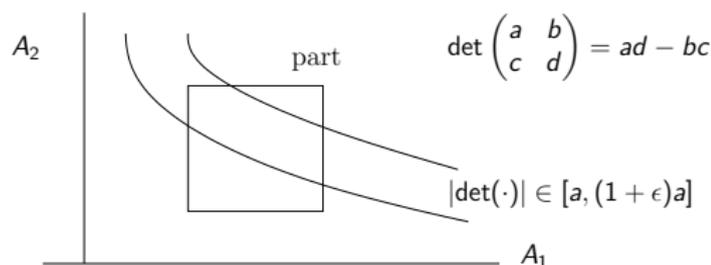
# Proof of the volume lower bound

- ▶ Warm up: What if a part is a “product set along rows” and the determinant on that part has a fixed value?  
Then, regions for each row are
  - ▶ a point,
  - ▶ a line parallel to the span of the previous point,
  - ▶ a plane parallel to the span of the previous regions ...



# Proof of the volume lower bound

- ▶ Intuition: shape of parts is very different from level sets of  $|\det(\cdot)|$ , as parts can be assumed to be “product sets along rows”. If  $|\det(\cdot)|$  is not dispersed in a part, then the part is small  $\implies$  many parts  $\implies$  tree of large depth ( $n^2 / \log n$ ).



# Main lemma for volume lower bound

- ▶ (dispersion of the determinant)  
For any partition of  $B_n^n$  into  $\leq 2^{n^2}$  parts that are “product sets along rows”, for half of the parts in measure we have

$$\Pr(|\det X| \notin [u, u(1+c)]) \geq \frac{1}{2^7 n^6}$$

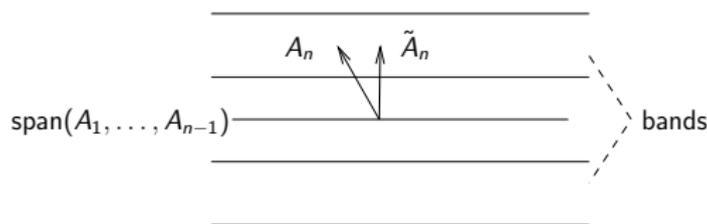
for any  $u$ , for  $X$  a random point (matrix) in the part, and for  $c$  a universal constant.

## Proof of lemma

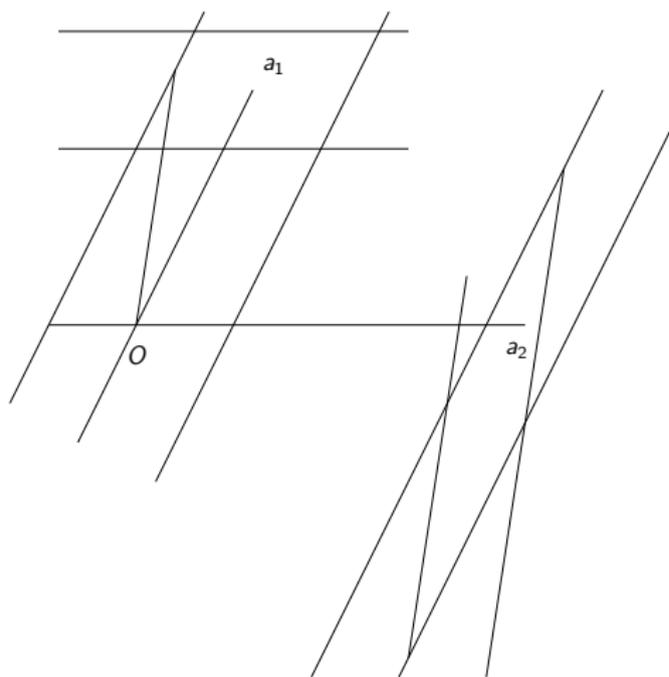
- ▶ Proof idea: Assume (for a contradiction) that a part  $P = \prod_i P_i$  is such that, for most matrices in it,  $|\det(\cdot)|$  fits in a small interval. Write  $|\det(A)| = \prod_i \|\tilde{A}_i\|$ , where  $\tilde{A}_i$  is the projection of  $A_i$  to the space orthogonal to  $A_1, \dots, A_{i-1}$ . Then

$$\|\tilde{A}_n\| \in [a, a(1 + \epsilon)] \prod_{i=1}^{n-1} \|\tilde{A}_i\|^{-1}.$$

Thus, for most choices of  $A_1, \dots, A_{n-1}$ , most of  $A_n \in P_n$  is forced to be contained in a “band”. A refinement of this argument gives a band on  $P_i$  for each pair  $i, j$ .



## Proof of lemma: pairs of bands



# Proof of lemma

- ▶  $\text{vol}(\sqrt{n}B_n)^n \geq 2^{n^2} + 2^{n^2}$  parts.
    - $\implies$  half of the parts (in measure) have volume at least  $1/2$  each.
  - ▶  $|\det(\cdot)|$  concentrated + “product set of matrices”
    - $\implies$  bands
    - $\implies$  volume of corresponding part less than  $1/2$ .
- Contradiction.

# Conclusion

We proved:

- ▶ an  $n^2 / \log n$  lower bound for the query complexity of approximating the volume of a convex body up to a constant factor by a randomized algorithm.

# Conclusion

Open problems:

- ▶ Complexity of approximating the volume: matching upper and lower bounds. Same with a separation oracle.