Approximation Algorithms for Stochastic Geometric Optimization Problems

Jian Li

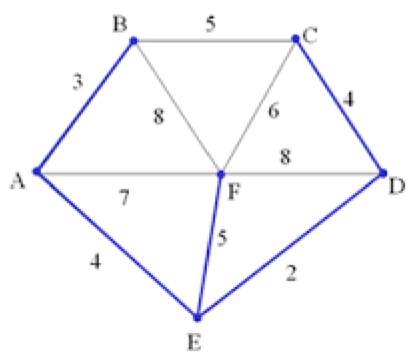
Institute of Interdisciplinary Information Sciences

Tsinghua University

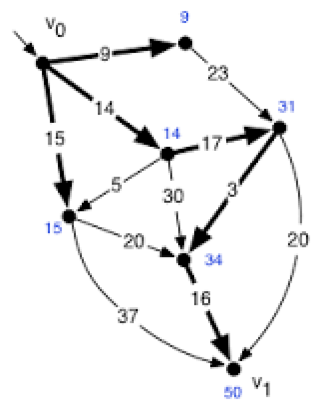
Outline

- Introduction
- Stochastic Geometry Models
- ϵ-Kernels/Coresets
- ϵ-Kernels for Stochastic Geometry
- ϵ-Expectation-Kernels
- Other Kernels/Coresets for Stochastic Geometry
- More Stochastic Combinatorial/Geometric Optimization Problems
- Conclusion

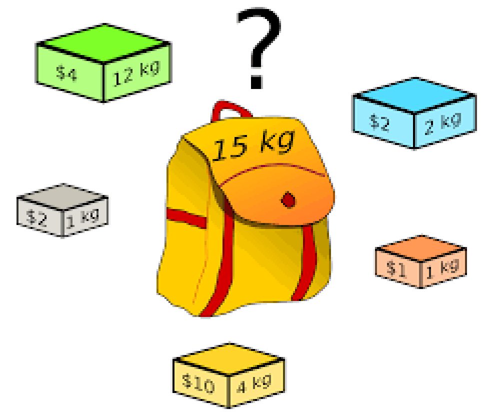
Combinatorial and Geometric Optimization problems



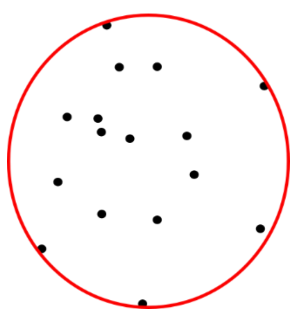
Minimum Spanning Tree



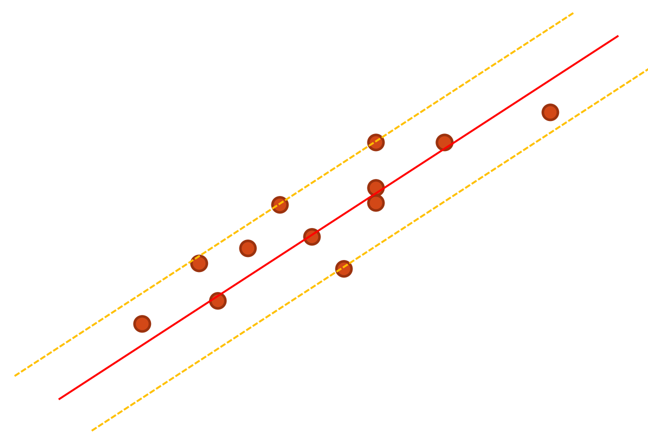
Shortest Path



Knapsack



Minimum Enclosing Ball



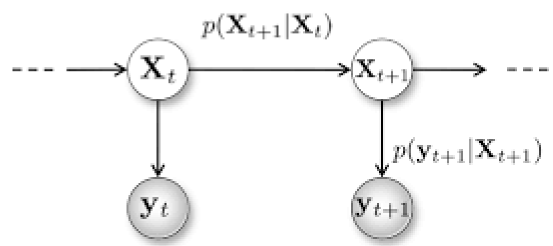
Minimum j-fat center

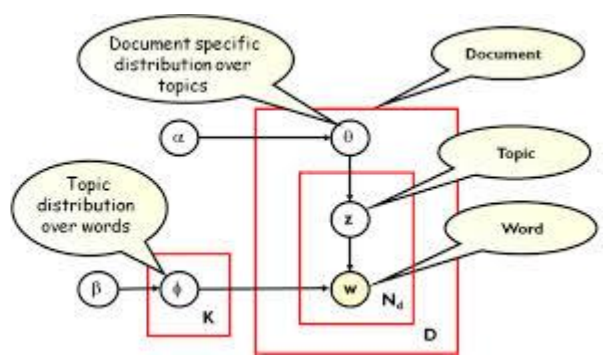
Uncertain Data and Stochastic Model

- Data Integration and Information Extraction
- Sensor Networks; Information Networks
- Probabilistic models in machine learning

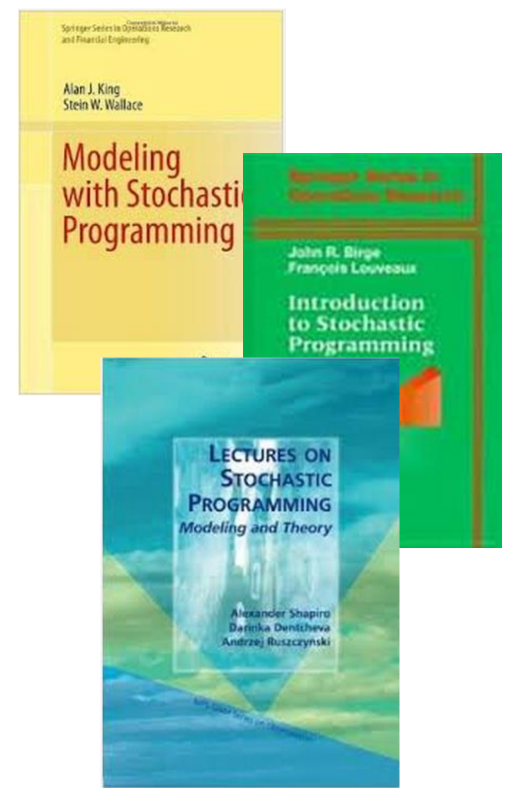
Sensor ID	Temp.
1	Gauss(40,4)
2	Gauss(50,2)
3	Gauss(20,9)
…	…

Probabilistic databases





Probabilistic Models in machine learning



Stochastic models in operation research

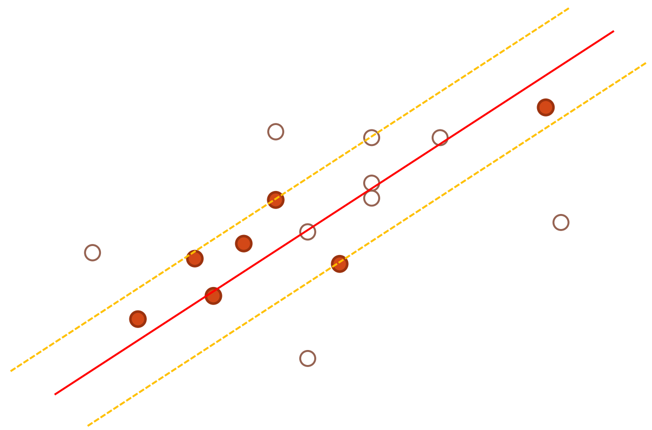
Stochastic Optimization

- Danzig in 1950s (linear programming with stochastic coefficients – stochastic programming)
- Depending on how the decision process interacts with the uncertainty, we may be able to formulate different versions of stochastic optimization problems
  - Estimation (no decision)
  - Single-stage
  - 2-stage
  - Multi-stage
  - Online (adaptive/non-adaptive))
  - Geometric Optimization problems

Stochastic Minimum j-flat Center

- Every point i exists with prob p_i
- Find a j-flat F (an affine subspace of dim j) such that

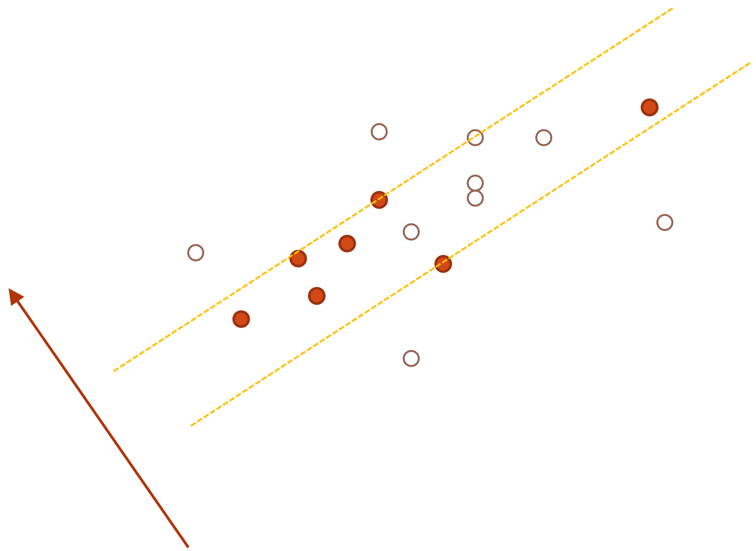
$$\mathrm{E}[\max_i d(i, F)] \text{ is miminized}$$



Stochastic Minimum Width

- Every point i exists with prob p_i
- Find a direction u such that

$$\mathrm{E}[w(Q,u)] \text{ is minimized}$$



In the deterministic setting, the minimum width problem is equivalent to min (d-1)-flat center

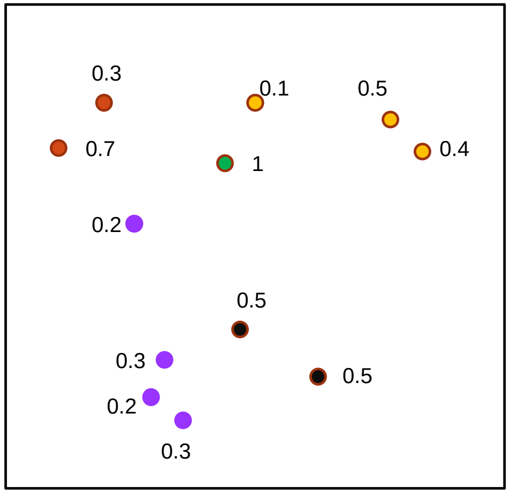
In stochastic setting, they are different.

Outline

- Introduction
- Stochastic Geometry Models
- ϵ-Kernels/Coresets
- ϵ-Kernels for Stochastic Geometry
- ϵ-Expectation-Kernels
- Other Kernels/Coresets for Stochastic Geometry
- More Stochastic Geometric Optimization Problems
- Conclusion

Stochastic Geometry Models

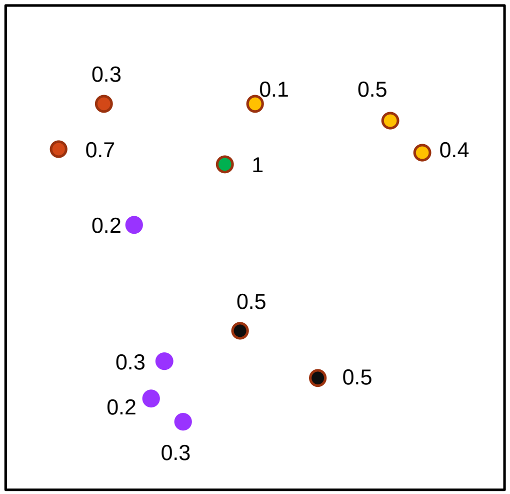
- The position of each point is random (non-i.i.d)
- All pts are independent from each other
- A popular model in wireless networks/spatial prob databases



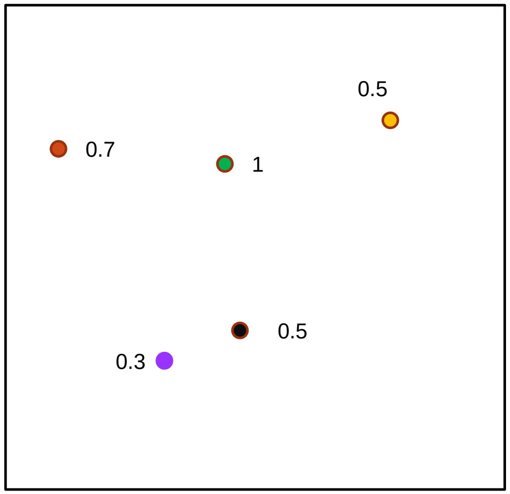
Locational uncertainty model

Stochastic Geometry Models

- The position of each point is random (non-i.i.d)
- All pts are independent from each other
- A popular model in wireless networks/spatial prob databases



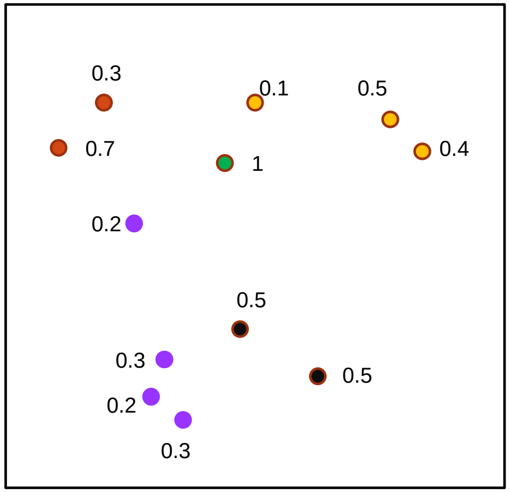
Locational uncertainty model



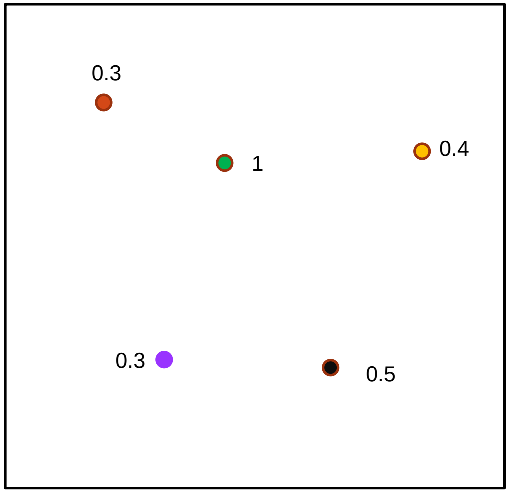
A realization (aka a possible world)
Prob=0.7*1*0.5*0.5*0.3

Stochastic Geometry Models

- The position of each point is random (non-i.i.d)
- All pts are independent from each other
- A popular model in wireless networks/spatial prob databases



Locational uncertainty model



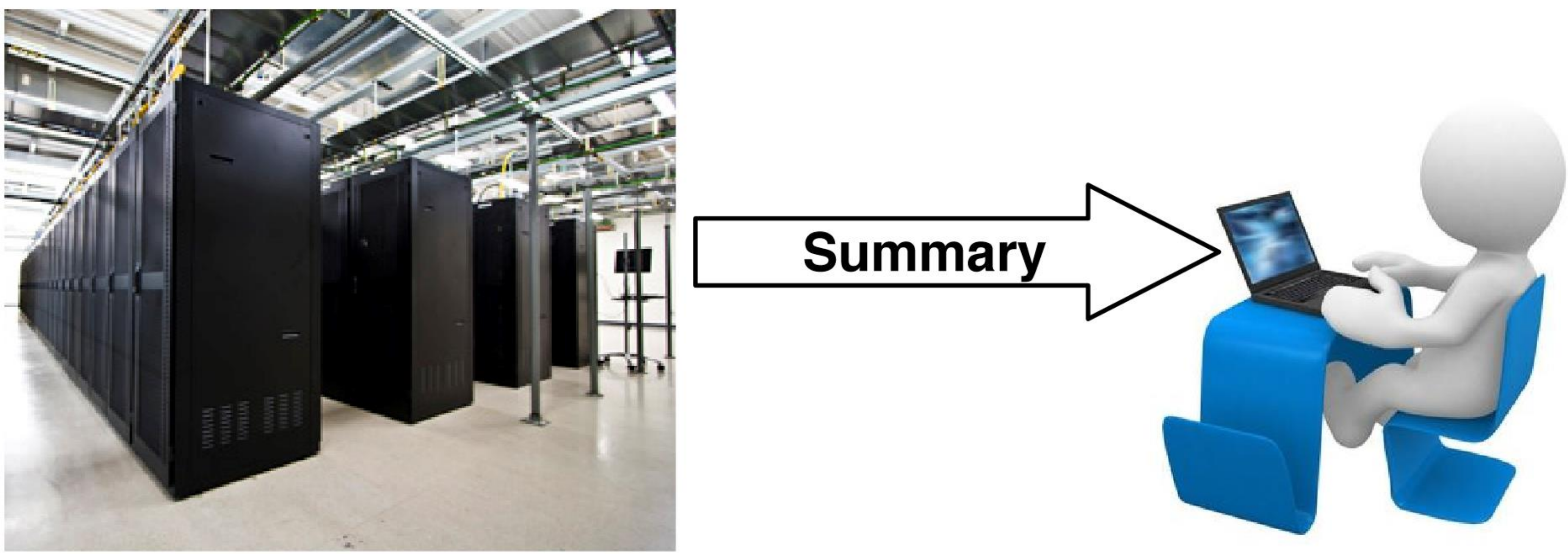
Another realization
Prob=0.3*1*0.4*0.5*0.3

Outline

- Introduction
- Stochastic Geometry Models
- ϵ-Kernels/Coresets
- ϵ-Kernels for Stochastic Geometry
- ϵ-Expectation-Kernels
- Other Kernels/Coresets for Stochastic Geometry
- More Stochastic Geometric Optimization Problems
- Conclusion

Kernel/Coreset

- Why kernel/Coreset?
- Turn BIG DATA to small data

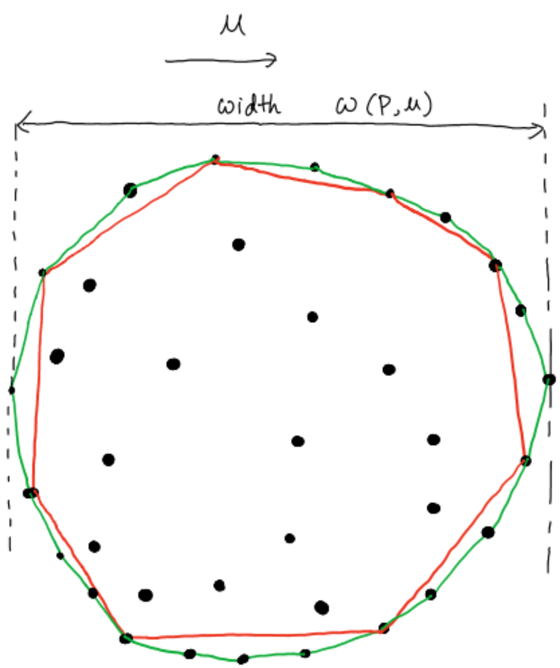


Esp-kernel

- A powerful notion in computational geometry [Agarwal et al.04]

Let $w(P,u)$ be the width of a deterministic n-point set $P \subset \mathbf{R}^d$ in a direction u. An ϵ-kernel $S \subseteq P$ s.t. for any direction u,

$$(1-\epsilon)w(P,u) \leq w(S,u) \leq w(P,u).$$



Construction of ϵ-kernels (Chan, 2006; Yu et.al., 2008):

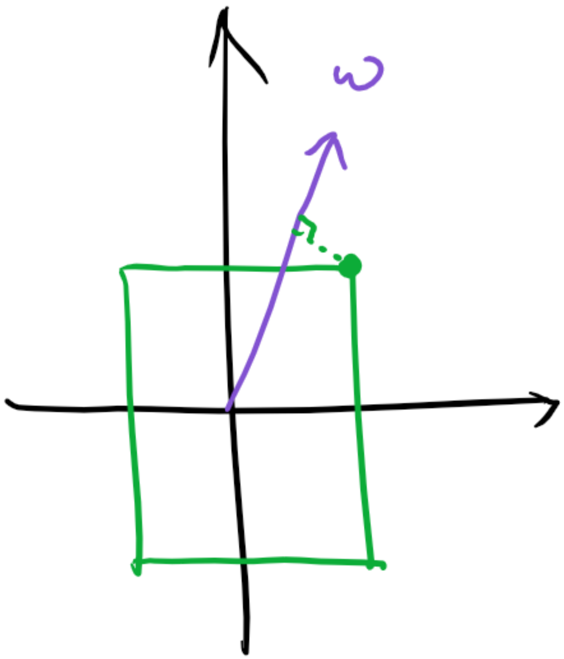
- size: $O(\epsilon^{-(d-1)/2})$,
- time: $O(n + \epsilon^{-(d-3/2)})$.

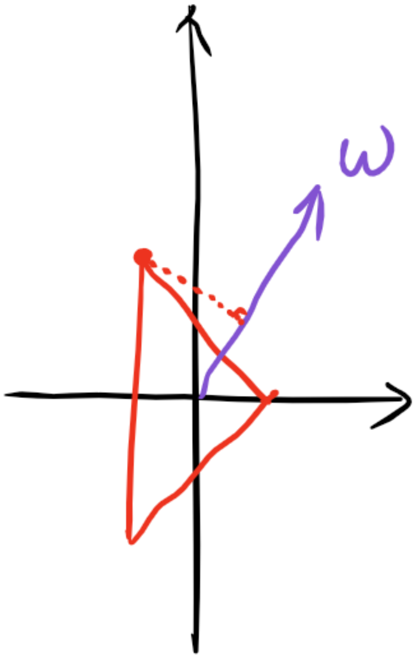
Esp-kernel

- ϵ-kernel is useful in designing efficient algorithms for many CG problems (using the linearization trick, originally used by Yao-Yao)
    1. Approximate function extent,
    2. Minimum enclosing ball,
    3. Minimum enclosing cylinder,
    4. Minimum spherical cell,
    5. Minimum cylinder cell
    6. ….
- The idea has been extended to numerous other problems: k-center, k-means, k-median, shape fitting, clustering, matrix approximation, submodular functions, connection to streaming/sketch

Support Function

- Support Function: $f(P, u) = \sup_{p \in P} \langle p, u \rangle$
- Width: $w(P, u) = f(P, u) - f(P, -u)$
- We can assume w.l.o.g. that $\|u\| = 1$





Outline

- Introduction
- Stochastic Geometry Models
- ϵ-Kernels/Coresets
- ϵ-Kernels for Stochastic Geometry
- ϵ-Expectation-Kernels
- Other Kernels/Coresets for Stochastic Geometry
- More Stochastic Combinatorial/Geometric Optimization Problems
- Conclusion

Stochastic Points

- How to extend the notion of ϵ-kernel to stochastic points??
  - The directional width is not a number any more! It is a random variable.

- Definition 1: Approximate the expectation of the directional width for all directions - ϵ-Exp-Kernel
- Definition 2: Approximate the distribution of the directional width for all directions - (ϵ, τ)-Quant-Kernel

Outline

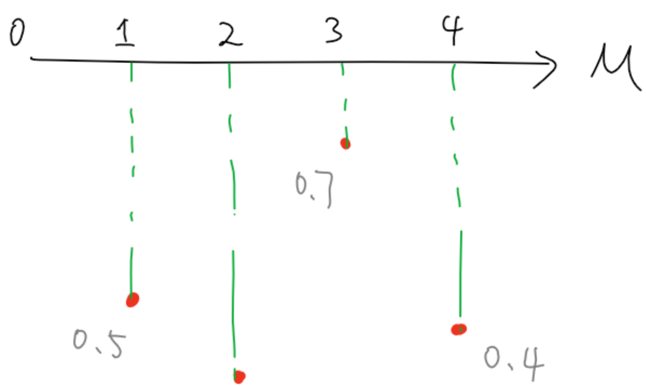
- Introduction
- Stochastic Geometry Models
- ϵ-Kernels/Coresets
- ϵ-Kernels for Stochastic Geometry
- ϵ-Expectation-Kernels
- Other Kernels/Coresets for Stochastic Geometry
- More Stochastic Combinatorial/Geometric Optimization Problems
- Conclusion

0,2

ϵ -Expectation-Kernel

- Define the expected value of the directional width

$$w(\mathcal{P}, u) = \mathbb{E}_{Q \sim \mathcal{P}}[w(Q, u)]$$



$$0.5 \times 0.4 \times 3 + \left((1-0.5) \times 0.2 \times 0.4 + 0.5 \times 0.7 \times (1-0.4)\right) \times 2 + \left(0.5 \times 0.2 \times (1-0.7) \times (1-0.4) + (1-0.5) \times 0.2 \times 0.7 \times (1-0.4) + (1-0.5)(1-0.2)\, 0.7 \times 0.4\right) \times 1$$

- ϵ-exp-kernel S: for any direction u:

$$(1-\epsilon) w(\mathcal{P}, u) \leq w(S, u) \leq w(\mathcal{P}, u).$$

In fact, we can choose S to be a constant-sized set of deterministic points

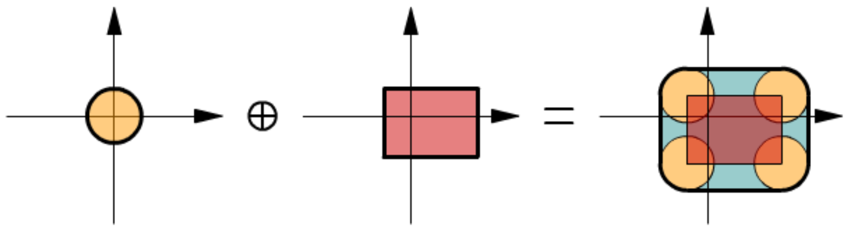
ϵ -Expectation-Kernel

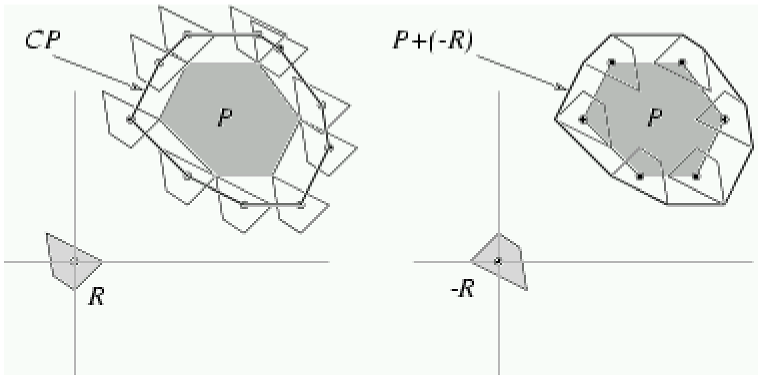
- Question 1: Does such kernel even exist?
- Question 2: How to find it efficiently?
- Question 3: What it is good for?

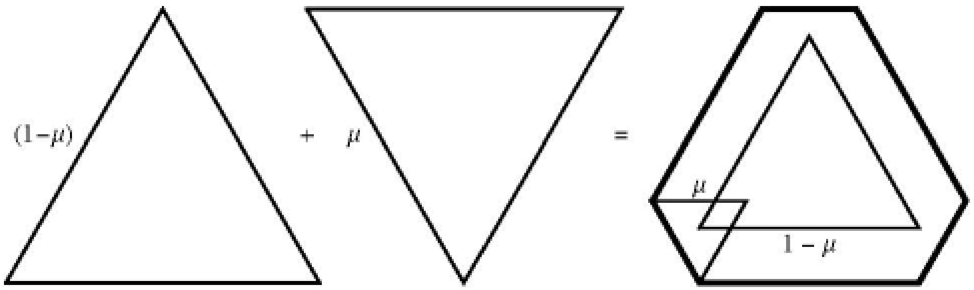
Minkovski Sum

- For sets A and B

  their Minkovski sum $A + B = \{a + b \mid a \in A, b \in B\}$



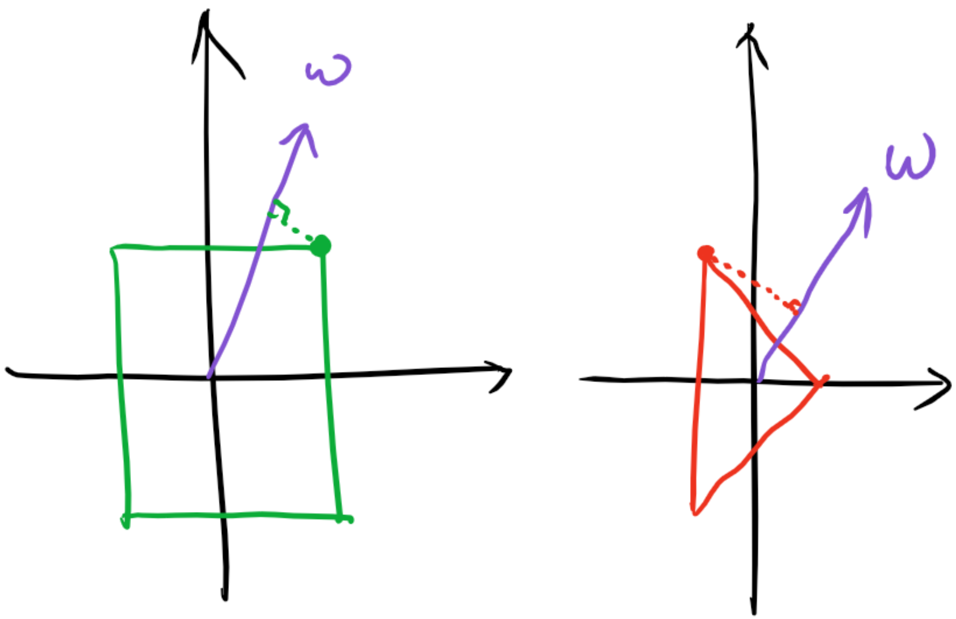




Minkovski Sum

- An important property of Minkovski Sum

$$f(P,u)+f(Q,u)=f(P+Q,u)$$



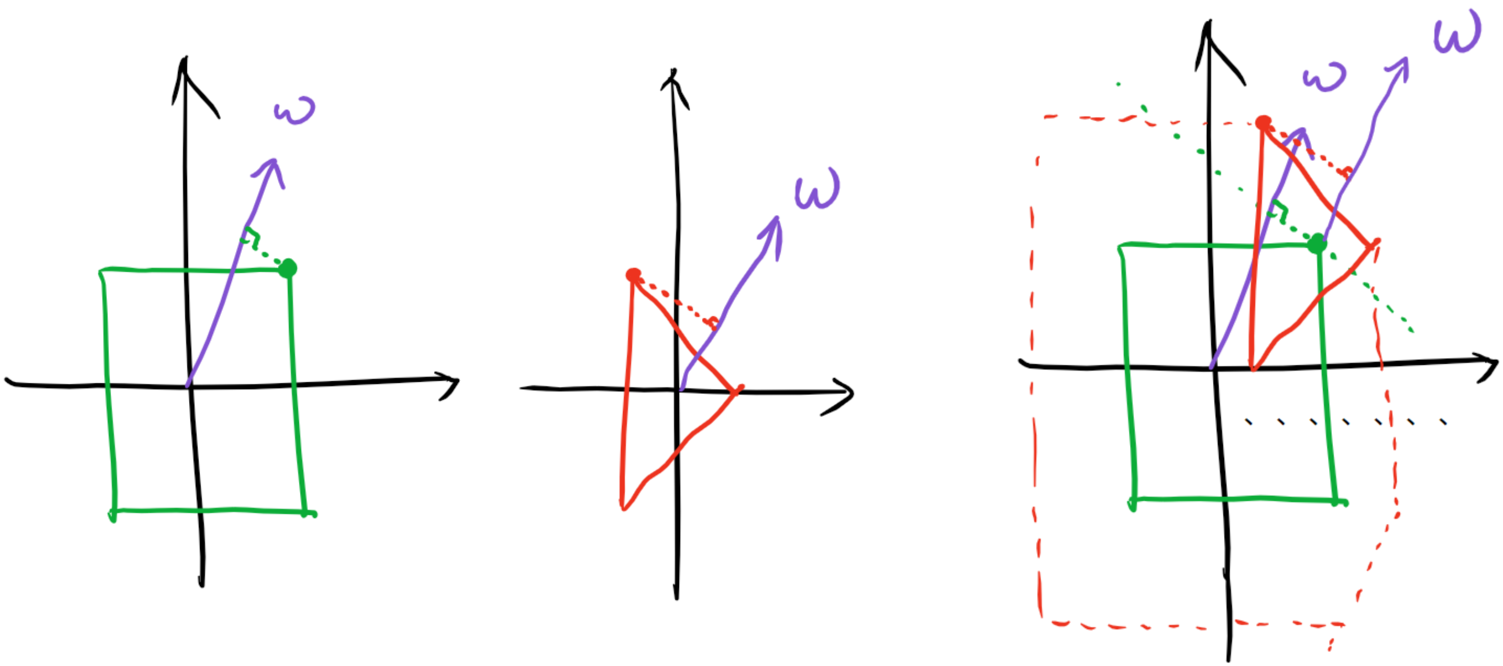
- Another easy property (α is a real number)

$$f(\alpha P,u)=\alpha f(P,u)$$

Minkovski Sum

- An important property of Minkovski Sum

$$f(P,u) + f(Q,u) = f(P+Q,u)$$



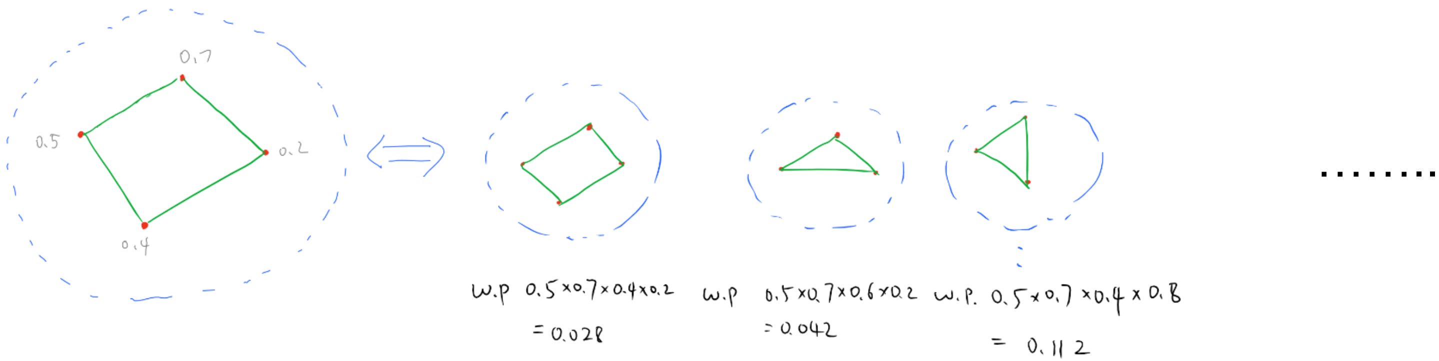
- Another easy property (α is a real number)

$$f(\alpha P,u) = \alpha f(P,u)$$

Existence of ϵ -Exp-Kernel

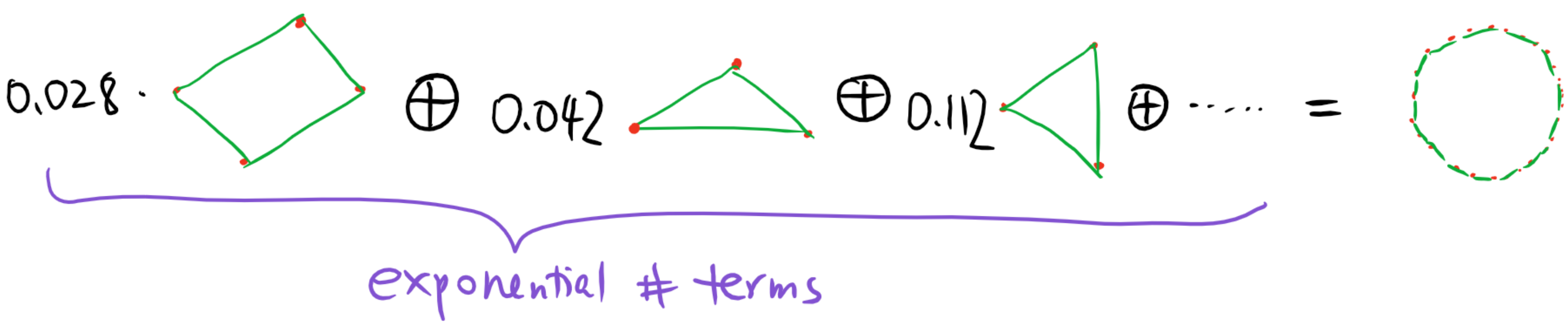
- Consider the expected value of the support function

$$E_Q[f(Q,u)] = \sum \Pr[\mathrm{Q}] f(Q,u)$$



- $E_Q[f(Q,u)] = \sum \Pr[\mathrm{Q}] f(Q,u) = f(\sum \Pr[Q] Q, u)$

Minkovski Sum



Existence of ϵ -Exp-Kernel

- We just show that

  There exists a deterministic convex shape *M* such that

$$w(M,u) = \mathrm{E}_Q[w(Q,u)]$$

- Every deterministic convex shape has an ϵ-kernel of size $\epsilon^{-(d-1)/2}$ [Agarwal et al '04]
- So, we have proved the existence!
- How to construct it efficiently?
- Let us first try to understand the deterministic convex shape *M*

  (which is the Minkovski sum of exponential convex shapes)

A Deep Understanding

- Let us first try to understand the deterministic convex shape M (which is the Minkovski sum of exponential convex shapes)
- What is the complexity of M (i.e., #vertices)?
- It seems to be exponential…..

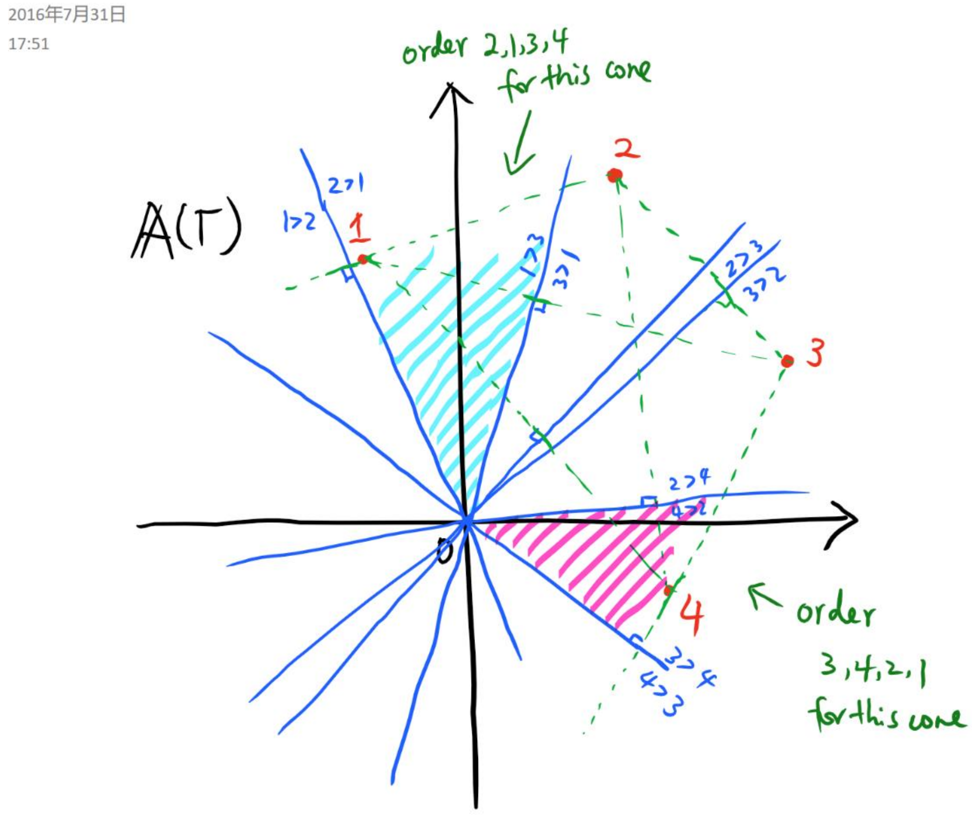
A Deep Understanding

- Let us first try to understand the deterministic convex shape *M* (which is the Minkovski sum of exponential convex shapes)
- What is the complexity of M (i.e., #vertices)?
- It seems to be exponential…..
- But we are going to prove it is polynomial !

$$O(\binom{n^2}{d-1}) = O(n^{2d-2})$$

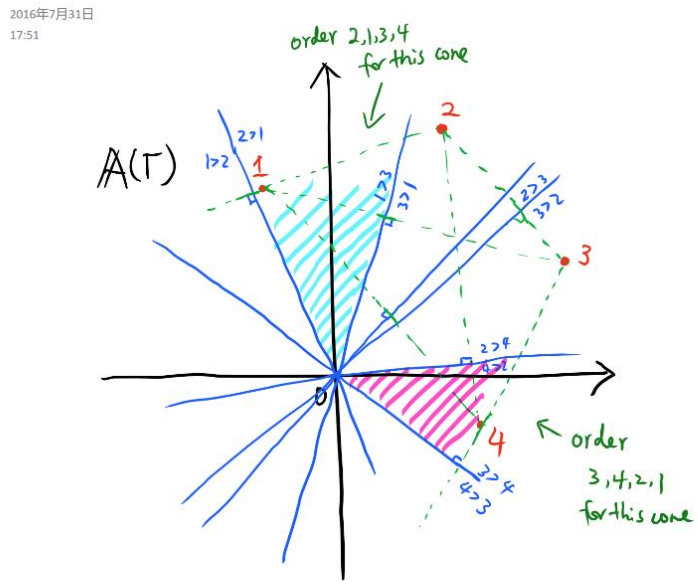
A Polynomial Size Bound

- Consider the existential uncertainty model
- Consider the arrangement $\mathbb{A}(\Gamma)$



A Polynomial Size Bound

- Consider the existential uncertainty model
- Consider the arrangement $\mathbb{A}(\Gamma)$



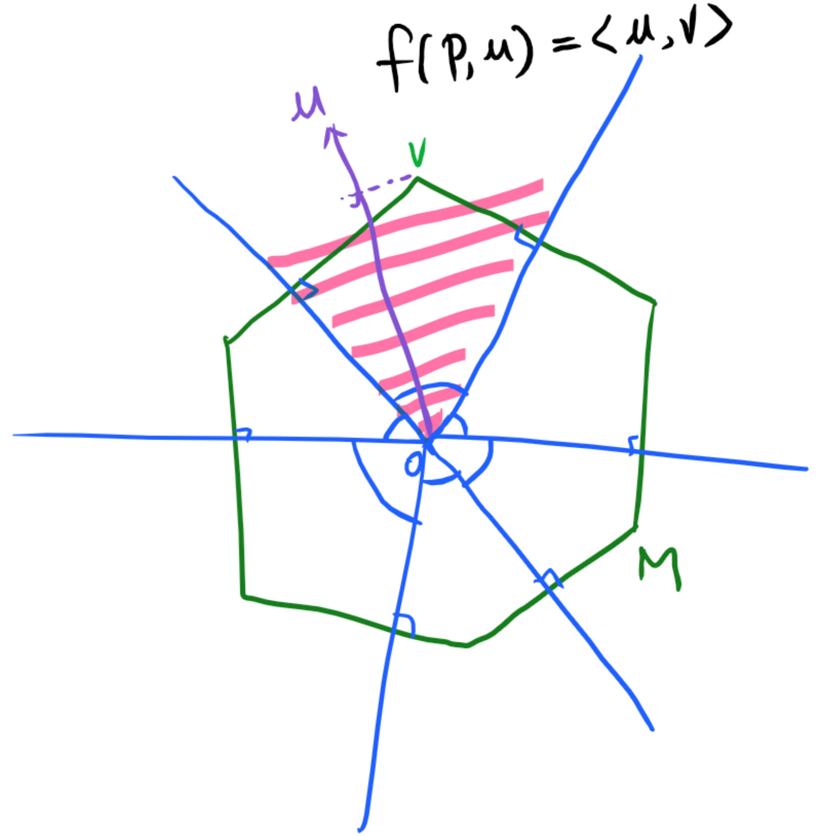
- THM: $|M| = |\mathbb{A}(\Gamma)|$

Moreover, each cone C in $\mathbb{A}(\Gamma)$ corresponds to a vertex in M as follows:

$$\nabla f(M, u) = v \ for\ all\ u \in \operatorname{int} C$$

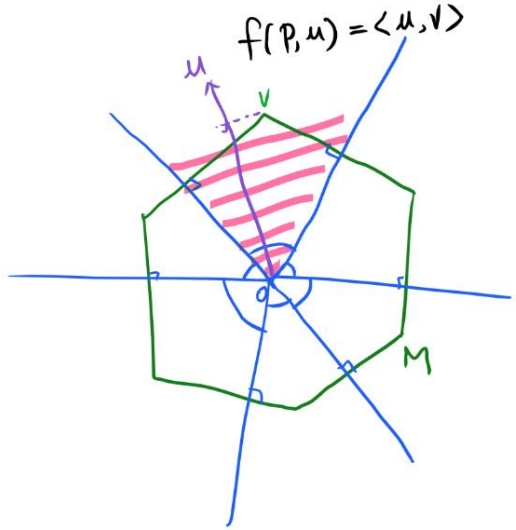
A Polynomial Size Bound - Proof

- Fact: For each convex body M, we can divide the space into $|M|$ cones, such that each cone C_v corresponds to a vertex v of M and $f(M,u) = \langle v, u \rangle$ for any $u \in C_v$.



A Polynomial Size Bound - Proof

- Fact: For each convex body M, we can divide the space into $|M|$ cones, such that each cone C_v corresponds to a vertex v of M and $f(M,u) = \langle v, u \rangle$ for any $u \in C_v$.



- Hence, for any $u \in C_v$

$$\nabla f(M,u) = \left\{\frac{\partial f(M,u)}{\partial u_j}\right\}_{j\in[d]} = \left\{\frac{\partial \langle u, v\rangle}{\partial u_j}\right\}_{j\in[d]} = \left\{\frac{\partial \sum_{j\in[d]} v_j u_j}{\partial u_j}\right\}_{j\in[d]} = v,$$

- Conclusion 1: $\nabla f(M,u)$ is a constant vector for each cone C_v

Proof - Cont

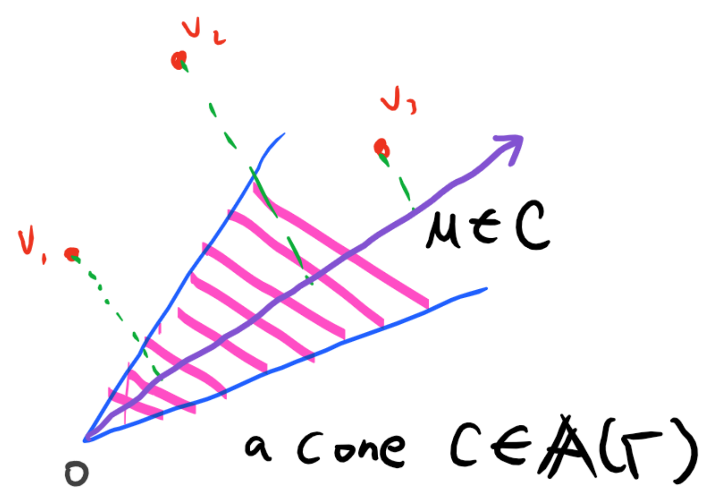
- Now, consider a cone C in $\mathbb{A}(\Gamma)$
- We show $\nabla f(M, u)$ is a constant vector for all $u \in \operatorname{int} C$.

Proof - Cont

- Now, consider a cone C in $\mathbb{A}(\Gamma)$
- We show $\nabla f(M, u)$ is a constant vector for all $u \in \operatorname{int} C$
- First, we notice that

$$f(M, u) = f(\mathcal{P}, u) = \sum_{v \in \mathcal{P}} \Pr^R(v, u)\langle v, u\rangle$$

$$\Pr^R(v, u) = \prod_{v' \succ_u v}(1 - p_{v'})p_v$$



Proof - Cont

- Now, consider a cone C in $\mathbb{A}(\Gamma)$
- We show $\nabla f(M,u)$ is a constant vector for all $u \in \operatorname{int} C$
- First, we notice that

$$f(M,u) = f(\mathcal{P},u) = \sum_{v\in\mathcal{P}} \Pr^R(v,u)\langle v,u\rangle$$

$$\Pr^R(v,u) = \prod_{v' \succ_u v}(1-p_{v'})p_v$$

- In cone C, the order doesn't change (So $\Pr^R(v,u)$ does not change. In particular, it does not depend on u)
- Hence, we can see that

$$\nabla f(M,u) = \sum_{v\in\mathcal{P}} \Pr^R(v,u)v$$

a constant independent of u

A Polynomial Size Bound

- $\nabla f(M, u)$ is a piecewise constant in $\mathbb{A}(\Gamma)$
- It is not hard to show the constant is not the same for different cones
- Hence, $|M| = |\mathbb{A}(\Gamma)|$
- $\binom{n}{2}$ hyperplanes (passing the origin) can divide the d-dim space into this many cones

$$O(\binom{n^2}{d-1}) = O(n^{2d-2})$$

- This can be made constructive: we can spend this amount of time to construct M

ϵ -Expectation-Kernel

- Question 1: Does such kernel even exist?

- Question 2: How to find it efficiently?

- Question 3: What it is good for?

A Nearly Linear Time Algorithm

- Constructing *M* is expensive (e.g., d=10)
- Can we construct the kernel without constructing *M* explicitly?

- Yes, we can.
- We can do this in $O(2^d n \log n)$ time
- A key procedure:

We are able to find the extreme vertex of *M* for a given direction in $O(n \log n)$ time.

The idea follows from our previous proof!

$$\nabla f(M,u) = \sum_{v \in \mathcal{P}} \Pr^R(v,u) v$$

ϵ -Expectation-Kernel

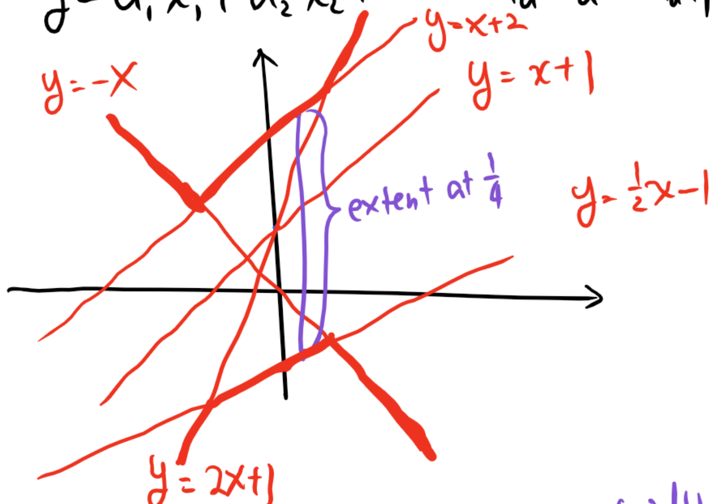
- Question 1: Does such kernel even exist?
- Question 2: How to find it efficiently?
- Question 3: What it is good for?

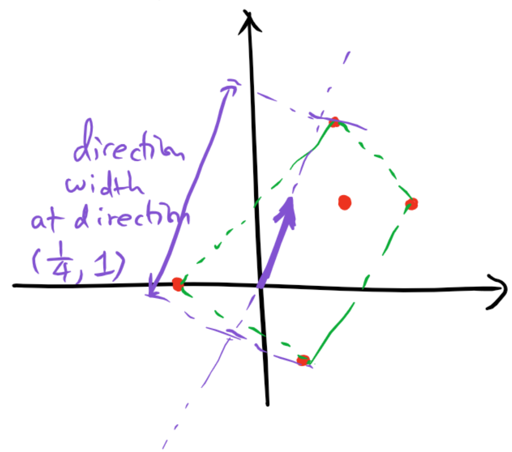
Applications

- Function extent
- Duality transform:

map linear function to point

$$y = a_1x_1 + a_2x_2 + \cdots + a_d x_d + a_{d+1} \iff (a_1, \ldots, a_{d+1})$$





$$\text{extend} = \left(\tfrac{1}{4} + 2\right) - \left(\tfrac{1}{2} \times \tfrac{1}{4} - 1\right)$$

$$\text{width} = \langle (\tfrac{1}{4}, 1), (1, 2) \rangle - \langle (\tfrac{1}{4}, 1), (\tfrac{1}{2}, -1) \rangle$$

Applications

- Each function appears with some probability
- We are interested in the expectation of the extent
- By duality, it is equivalent to the direction width problem!

- By the linearization trick, we can give PTAS for the problem minimizing the expected areas of the enclosing ball and the enclosing annulus in the plane.

Application

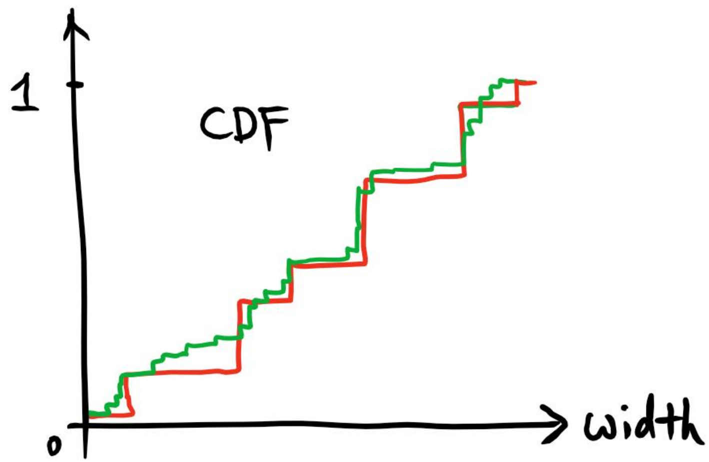
- Stochastic Moving Points
- A set of stochastic points, each moving along a polynomial trajectory
- By our function extent result, we can show that we can construct a constant number of deterministic moving points, such that the directional width approximates the expected direction width of the stochastic points, for any direction and any time!

Outline

- Introduction
- Stochastic Geometry Models
- ϵ-Kernels/Coresets
- ϵ-Kernels for Stochastic Geometry
- ϵ-Expectation-Kernels
- Other Kernels/Coresets for Stochastic Geometry
- More Stochastic Combinatorial/Geometric Optimization Problems
- Conclusion

Approximate the Distribution

- Want to approximate the distribution for every direction



- (ϵ, τ)-Quant-Kernel: For every direction u,

$$\Pr_{P\sim\mathcal{P}}\Big[\omega(P,u) \le (1-\varepsilon)x\Big] - \tau \le \Pr_{S\sim\mathcal{S}}\Big[\omega(S,u) \le x\Big] \le \Pr_{P\sim\mathcal{P}}\Big[\omega(P,u) \le (1+\varepsilon)x\Big] + \tau$$

Algorithm for Quant-Kernel

Algorithm:

- Take N samples from the stochastic model where

$$N = O\left(\tau^{-2}\varepsilon^{-(d-1)}\log(1/\varepsilon)\right)$$

- Compute the ϵ-kernel K_i for each sample Q_i
- Quant-Kernel = $\{K_1, K_2, \ldots, K_N\}$, each w.p. $1/N$

▶ **Theorem** *An* (ε,τ)-QUANT-KERNEL *of size* $\widetilde{O}\left(\tau^{-2}\varepsilon^{-3(d-1)/2}\right)$ *can be constructed in* $\widetilde{O}\left(n\tau^{-2}\varepsilon^{-(d-1)}\right)$ *time, under both existential and locational uncertainty models.*

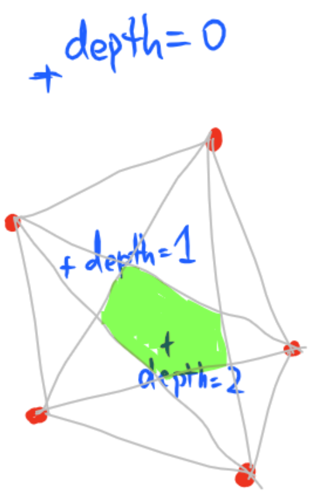
- Proof uses the celebrated VC (Vapnik-Chervonenkis) uniform convergence theory + VC-dimension for union of half spaces

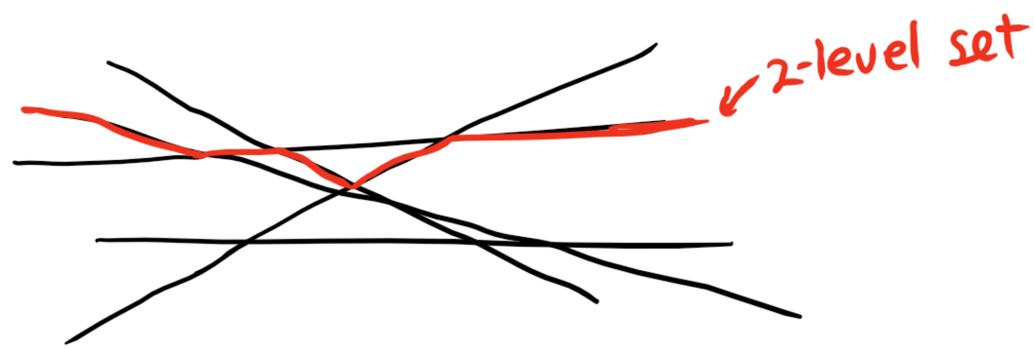
Algorithm for Quant-Kernel

- The above result can be improved for existential model:

▶ **Theorem** *$\mathcal{P}$ is a set of uncertain points in $\mathbb{R}^d$ with existential uncertainty. Let $\lambda = \sum_{v \in \mathcal{P}}(-\ln(1-p_v))$. There exists an (ε, τ)-*QUANT-KERNEL* for $\mathcal{P}$, which consists of a set of independent uncertain points of cardinality $\min\{\widetilde{O}(\tau^{-2}\max\{\lambda^2, \lambda^4\}), \widetilde{O}(\varepsilon^{-(d-1)}\tau^{-2})\}$. The algorithm for constructing such a coreset runs in $\widetilde{O}(n \log^{O(d)} n)$ time.*

- A more complicated construction and analysis
- Interesting connections to Tukey Depth and k-Level set





Other Kernel/Coresets

- Approximate Fractional Power [H**L**PW,ESA16]

$$T_r(P,u) = \max_{v \in P} \langle u, v \rangle^{1/r} - \min_{v \in P} \langle u, v \rangle^{1/r}$$

- Fractional power kernel *S*:

$$(1-\varepsilon)\mathbb{E}_{P\sim\mathcal{P}}[T_r(P,u)] \le \mathbb{E}_{P\sim\mathcal{S}}[T_r(P,u)] \le (1+\varepsilon)\mathbb{E}_{P\sim\mathcal{P}}[T_r(P,u)].$$

- Minimum Enclosing Balls [MSF, SCG'14]
- Minimum j-flat center [H**L**,SODA'17]
- Minimum k-center [H**L**,SODA'17]

Outline

Many More Problems

- Geometric optimization
  - Nearest neighbor queries
  - Range queries
  - Hyperplane Separation (SVM)
  - Coresets
  - Shape fitting (minimum enclosing ball, minimum j-flat center, minimum k-center etc.)
  - …….

Conclusion

- Bayesian mechanism design (essentially stochastic optimization problems)
- Learning+Optimization
  - We don't have to first learn the distributions first, and then solve the stochastic optimization problem. We can do it together and use less samples!
- A fascinating topic with interesting connections to many subareas in TCS (counting, coresets, geometry, VC theory, bandits, online algorithms, mechanism design,….) and probability theory/statistics
- A lot more interesting problems to be studied
- Many open problems
- A Survey: Jian Li, Yu Liu. Approximation Algorithms for Stochastic Combinatorial Optimization Problems. 2016

Thanks

lijian83@mail.tsinghua.edu.cn