How Good Are Sparse Cutting Planes?

Santanu S. Dey Marco Molinaro Qianyi Wang

ISyE - Georgia Tech

Thanks to Marco Molinaro for the figures

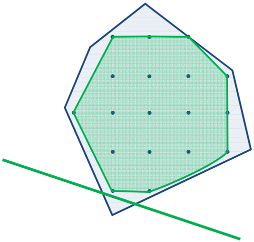
Outline

1
Introduction and Motivation

Cutting-planes

Cutting Plane

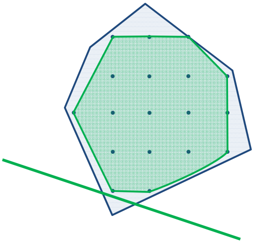
1. Cutting-planes in a linear inequality that is valid for all integer feasible points, but may not be valid for the linear programming relaxation.



Cutting-planes

Cutting Plane

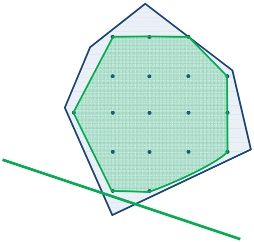
1. Cutting-planes in a linear inequality that is valid for all integer feasible points, but may not be valid for the linear programming relaxation.
2. The convex hull of integer feasible solutions is a rational polyhedron, so a finite number of cutting-planes gives the convex hull.



Cutting-planes

Cutting Plane

1. Cutting-planes in a linear inequality that is valid for all integer feasible points, but may not be valid for the linear programming relaxation.
2. The convex hull of integer feasible solutions is a rational polyhedron, so a finite number of cutting-planes gives the convex hull.
3. Huge amount of research in Integer Programming on problem-specific and general purpose cutting-planes.
4. Justified by (or more importantly?) being extremely useful in practice to solve IPs.



Implementation of cutting planes is a different matter altogether

A double edged sword?

More commercial/successful IP solvers have very sophisticated methods of cutting-planes selection and use.

Implementation of cutting planes is a different matter altogether

A double edged sword?

More commercial/successful IP solvers have very sophisticated methods of cutting-planes selection and use.

1. "Dept of cut" and variants ("Rotated distance", "Distance with bounds",)
2. "Parallelism" ("Objective function parallelism", "Cutting plane parallelism")
3. "Numerical stability"
4. "Cutting-plane sparsity"

'Cut pool management system', 'Cutting-plane filter'

Implementation of cutting planes is a different matter altogether

A double edged sword?

More commercial/successful IP solvers have very sophisticated methods of cutting-planes selection and use.

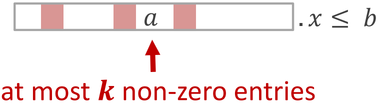
1. "Dept of cut" and variants ("Rotated distance", "Distance with bounds",)
2. "Parallelism" ("Objective function parallelism", "Cutting plane parallelism")
3. "Numerical stability"
4. "Cutting-plane sparsity"

'Cut pool management system', 'Cutting-plane filter'

Cutting plane Sparsity

In practice

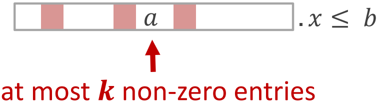
1. Prefer to use sparse cutting planes.



Cutting plane Sparsity

In practice

1. Prefer to use sparse cutting planes.



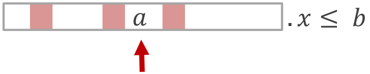
2. Why use sparse cutting-planes:

Pros: It is easy to solve LPs in B&B tree with sparse cutting-planes.

Cutting plane Sparsity

In practice

1. Prefer to use sparse cutting planes.

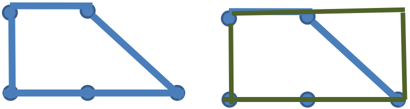


at most k non-zero entries

2. Why use sparse cutting-planes:

Pros: It is easy to solve LPs in B&B tree with sparse cutting-planes.

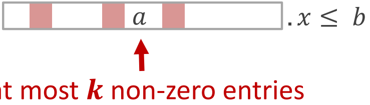
Cons: Sparse cutting planes do not give the convex hull.



Cutting plane Sparsity

In practice

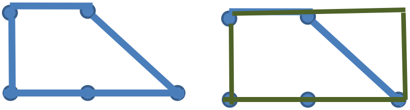
1. Prefer to use sparse cutting planes.



2. Why use sparse cutting-planes:

Pros: It is easy to solve LPs in B&B tree with sparse cutting-planes.

Cons: Sparse cutting planes do not give the convex hull.



How good are sparse cutting-planes?

2
Setting up the problem

Geometric Problem

Considered abstractly

1. We have a polytope $P \subseteq \mathbb{R}^n$, which represents the integer hull.

Geometric Problem

Considered abstractly

1. We have a polytope $P \subseteq \mathbb{R}^n$, which represents the integer hull.
2. P^k : The set of points that are valid for all k-sparse inequalities.

Geometric Problem

Considered abstractly

1. Instead of considering a general polytope, we have a polytope $\cancel{P \subseteq \mathbb{R}^n}$ $P \subseteq [0\ 1]^n$, which represents the integer hull.
2. P^k : The set of points that are valid for all k-sparse inequalities.

Geometric Problem

Considered abstractly

1. Instead of considering a general polytope, we have a polytope $\cancel{P \subseteq \mathbb{R}^n}$ $P \subseteq [0\ 1]^n$, which represents the integer hull.
2. P^k : The set of points that are valid for all k-sparse inequalities.

$\boxed{\text{How well does } P^k \text{ approximate } P?}$

Geometric Problem

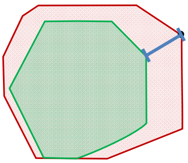
Considered abstractly

1. Instead of considering a general polytope, we have a polytope $\cancel{P \subseteq \mathbb{R}^n}$ $P \subseteq [0\ 1]^n$, which represents the integer hull.
2. P^k : The set of points that are valid for all k-sparse inequalities.

$$\boxed{\text{How well does } P^k \text{ approximate } P?}$$

We wanted a measure that is well-defined in all cases:

$$d(P, P^k) = \max_{x \in P^k} \{\text{distance}(x, P)\}$$



Comments:

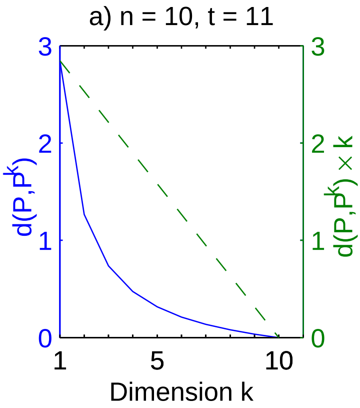
1. At most $\sqrt{n}$.
2. Upper bound on dept of cut measure.

3
Some Examples

Example 1: "Single phase"

$$P := \{x \in [0\ 1]^n | \sum_{i=1}^n x_i \le 1\}$$

$d(P, P^k) \approx \frac{\sqrt{n}}{k}$ This is Good!

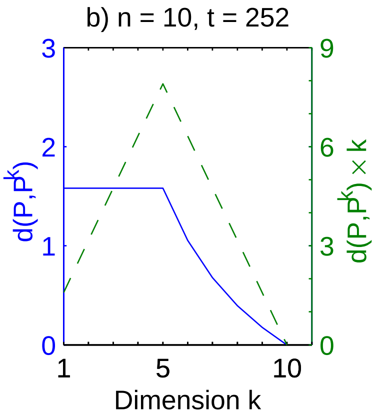


Example 1: "Two phase"

$$P := \{x \in [0\ 1]^n | \sum_{i=1}^n x_i \leq \frac{n}{2}\}$$

Phase 1: $d(P, P^k) \approx \sqrt{n}/2$ This is Bad!

Phase 2: $d(P, P^k) \approx \frac{n\sqrt{n}}{2k}$



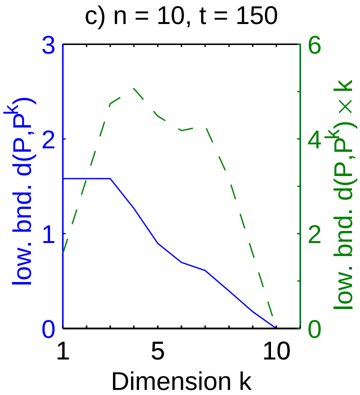
Example 1: "Three phases"

$P :=$ convex hull of random 0/1 points

Phase 1: $d(P, P^k) \approx \sqrt{n}/2$

Phase 2: $d(P, P^k) \propto \frac{1}{\sqrt{k}}$

Phase 3: $d(P, P^k) \propto \frac{1}{k}$

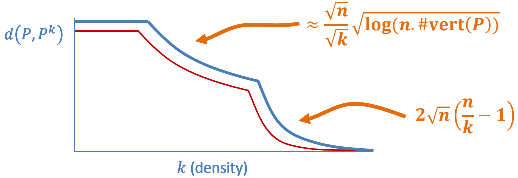


4

Outline of main results

Main results - Informally

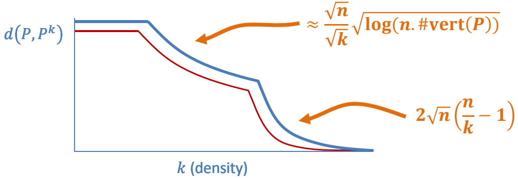
1. Upper bounds on $d(P, P^k)$ depending on n, k, and number of vertices.



Consequences

Main results - Informally

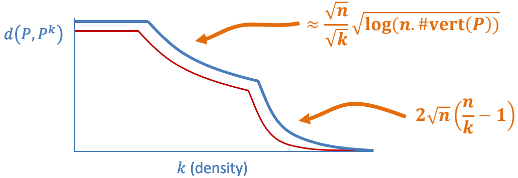
1. Upper bounds on $d(P, P^k)$ depending on n, k, and number of vertices.
2. Matching lower bound: a random 0/1 polytope, with prob $\frac{1}{4}$



Consequences

Main results - Informally

1. Upper bounds on $d(P, P^k)$ depending on n, k, and number of vertices.
2. Matching lower bound: a random 0/1 polytope, with prob $\frac{1}{4}$

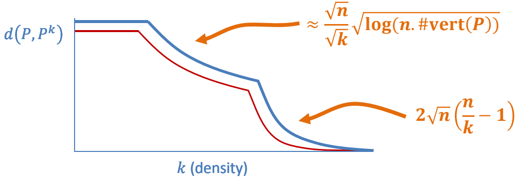


Consequences

1. Polynomial number of vertices as a function of dimension with $\sim \frac{1}{2}$ sparsity, implies $d(P, P^k)$ is very small ($\approx \sqrt{\log n}$), i.e. sparse cutting planes are good.

Main results - Informally

1. Upper bounds on $d(P, P^k)$ depending on n, k, and number of vertices.
2. Matching lower bound: a random 0/1 polytope, with prob $\frac{1}{4}$

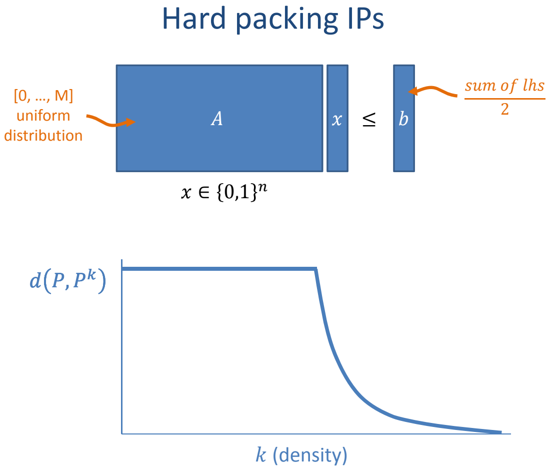


Consequences

1. Polynomial number of vertices as a function of dimension with $\sim \frac{1}{2}$ sparsity, implies $d(P, P^k)$ is very small ($\approx \sqrt{\log n}$), i.e. sparse cutting planes are good.
2. As the number of vertices grow, the location of vertices becomes more important....

Main results - Informally

3. Random packing instances: Sparse cutting-planes are bad.



Main results - Informally

3. Random packing instances: Sparse cutting-planes are bad.

4 k-sparse cutting-planes for extended formulation + projection to original space is significantly stronger than k-sparse cutting-planes for original polytope.

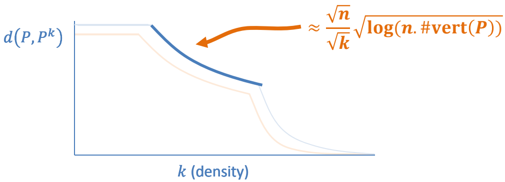
4
Results and outline of proofs

Upper Bound

Theorem 1 [Upper Bound on $d(P, P^k)$]

Let $n \geq 2$. Let $P \subseteq [0,1]^n$ be the convex hull of points $\{p^1, \ldots, p^t\}$. Then

1. $d(P, P^k) \leq 4 \max \left\{ \frac{n^{1/4}}{\sqrt{k}} \sqrt{8 \max_{i \in [t]} \|p^i\|} \sqrt{\log 4tn}, \frac{8\sqrt{n}}{3k} \log 4tn \right\}$
2. $d(P, P^k) \leq 2\sqrt{n} \left(\frac{n}{k} - 1 \right)$. $\square$



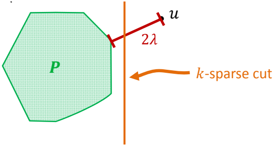
Proof Outline

To show: $d(P, P^k) \leq 2\lambda \triangleq \frac{\sqrt{n}}{\sqrt{k}}\sqrt{\log(n \times \text{number of vertices})}$.

Proof Outline

To show: $d(P, P^k) \leq 2\lambda \triangleq \frac{\sqrt{n}}{\sqrt{k}}\sqrt{\log(n \times \text{number of vertices})}$.

1. Let $u \in \mathbb{R}^n$ such that distance between u and P greater than 2λ.



2. How to obtain a k-sparse inequality to separate u?

Proof Outline

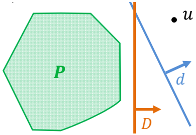
To show: $d(P, P^k) \leq 2\lambda \triangleq \frac{\sqrt{n}}{\sqrt{k}} \sqrt{\log(n \times \text{number of vertices})}$.

1. Let $u \in \mathbb{R}^n$ such that distance between u and P greater than 2λ.
2. How to obtain a k-sparse inequality to separate u?
3. Start with valid inequality $d^\top x \leq b$ of P; d is unit norm such that
   3.1 $2\lambda d = (u - v)$; v is the closest point to u in P.

Proof Outline

To show: $d(P, P^k) \leq 2\lambda \triangleq \frac{\sqrt{n}}{\sqrt{k}}\sqrt{\log(n \times \text{number of vertices})}$.

1. Let $u \in \mathbb{R}^n$ such that distance between u and P greater than 2λ.
2. How to obtain a k-sparse inequality to separate u?
3. Start with valid inequality $d^\top x \leq b$ of P; d is unit norm such that
   3.1 $2\lambda d = (u - v)$; v is the closest point to u in P.



4. Let $\alpha = k/(2\sqrt{n})$. Randomly "sparsify" d to a vector D:

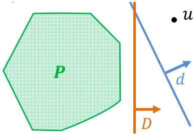
$\alpha\lvert d_i\rvert \leq 1$	$\alpha\lvert d_i\rvert > 1$
$D_i = \begin{cases} \frac{sign(d_i)}{\alpha} & \text{with probability} \alpha\lvert d_i\rvert \\ 0 & \text{with probability} 1 - \alpha\lvert d_i\rvert \end{cases}$	$D_i = d_i$

Proof Outline

To show: $d(P, P^k) \leq 2\lambda \triangleq \frac{\sqrt{n}}{\sqrt{k}}\sqrt{\log(n \times \text{number of vertices})}$.

1. Let $u \in \mathbb{R}^n$ such that distance between u and P greater than 2λ.

2. How to obtain a k-sparse inequality to separate u?
3. Start with valid inequality $d^\top x \leq b$ of P; d is unit norm such that
   3.1 $2\lambda d = (u - v)$; v is the closest point to u in P.



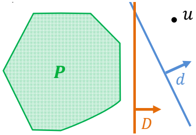
4. Let $\alpha = k/(2\sqrt{n})$. Randomly "sparsify" d to a vector D:

$\alpha\lvert d_i\rvert \leq 1$	$\alpha\lvert d_i\rvert > 1$
$D_i = \begin{cases} \frac{sign(d_i)}{\alpha} & \text{with probability} \alpha\lvert d_i\rvert \\ 0 & \text{with probability} 1-\alpha\lvert d_i\rvert \end{cases}$	$D_i = d_i$

Proof Outline

To show: $d(P, P^k) \leq 2\lambda \triangleq \frac{\sqrt{n}}{\sqrt{k}}\sqrt{\log(n \times \text{number of vertices})}$.

1. Let $u \in \mathbb{R}^n$ such that distance between u and P greater than 2λ.
2. How to obtain a k-sparse inequality to separate u?
3. Start with valid inequality $d^\top x \leq b$ of P; d is unit norm such that
   3.1 $2\lambda d = (u - v)$; v is the closest point to u in P.



4. Let $\alpha = k/(2\sqrt{n})$. Randomly "sparsify" d to a vector D:

$\alpha\lvert d_i\rvert \leq 1$	$\alpha\lvert d_i\rvert > 1$
$D_i = \begin{cases} \frac{sign(d_i)}{\alpha} & \text{with probability} \alpha\lvert d_i\rvert \\ 0 & \text{with probability} 1 - \alpha\lvert d_i\rvert \end{cases}$	$D_i = d_i$

5. With probability > 0, D is k-sparse, $D^\top x \leq (D^\top v + \lambda)$ is valid for P, and $D^\top u > (D^\top v + \lambda)$.

Upper Bound: "Sparsity of random vectors"

Claim: With high probability D is k-sparse.

Proof

1. $E[\#$ of non-zero components of $D] = \sum_{i=1}^{n} \alpha d_i \leq k/2$.

Upper Bound: "Sparsity of random vectors"

Claim: With high probability D is k-sparse.

Proof

1. $E[\#$ of non-zero components of $D] = \sum_{i=1}^{n} \alpha d_i \leq k/2$.
2. $Std[\#$ of non-zero components of $D] \leq \sqrt{k/2}$

Upper Bound: "Sparsity of random vectors"

Claim: With high probability D is k-sparse.

Proof

1. $E[\#$ of non-zero components of $D] = \sum_{i=1}^{n} \alpha d_i \leq k/2$.
2. $Std[\#$ of non-zero components of $D] \leq \sqrt{k/2}$
3. $\#$ of non-zero components of $D \lesssim k/2 + \sqrt{k/2}$

Upper Bound: "Sparsity of random vectors"

Claim: With high probability D is k-sparse.

Proof

1. $E[\#$ of non-zero components of $D] = \sum_{i=1}^{n} \alpha d_i \leq k/2$.
2. $Std[\#$ of non-zero components of $D] \leq \sqrt{k/2}$
3. $\#$ of non-zero components of $D \lesssim k/2 + \sqrt{k/2}$

$$P(\#\text{ of non-zero components of } D \leq k) \underbrace{\geq}_{\text{Bernstein's Inequality}} 1 - \frac{1}{4n}.$$

Upper Bound: "Validity of random inequality"

Observation

Let $w \in \mathbb{R}^n$. Then $E(D^\top w) = d^\top w$, $\text{Var}(D^\top w) = \frac{1}{\alpha} \sum_{i=1}^n w_i^2 |d_i|$.

Upper Bound: "Validity of random inequality"

Observation

Let $w \in \mathbb{R}^n$. Then $E(D^\top w) = d^\top w$, $\text{Var}(D^\top w) = \frac{1}{\alpha} \sum_{i=1}^n w_i^2 |d_i|$.

Claim: With high probability $D^\top x \leq (D^\top v + \lambda)$ is a valid inequality for P.

Proof

1. Fix a vertex p of P.

Upper Bound: "Validity of random inequality"

Observation

Let $w \in \mathbb{R}^n$. Then $E(D^\top w) = d^\top w$, $\mathrm{Var}(D^\top w) = \frac{1}{\alpha} \sum_{i=1}^n w_i^2 |d_i|$.

Claim: With high probability $D^\top x \leq (D^\top v + \lambda)$ is a valid inequality for P.

Proof

1. Fix a vertex p of P.
2. $E[D^\top (p - v)] = d^\top (p - v) \leq 0$.
3. $Std[D^\top (p - v)] \leq \frac{\lambda}{\sqrt{\log(n \times \text{ number of vertices})}}$

Upper Bound: "Validity of random inequality"

Observation

Let $w \in \mathbb{R}^n$. Then $E(D^\top w) = d^\top w$, $\mathrm{Var}(D^\top w) = \frac{1}{\alpha}\sum_{i=1}^n w_i^2 |d_i|$.

Claim: With high probability $D^\top x \leq (D^\top v + \lambda)$ is a valid inequality for P.

Proof

1. Fix a vertex p of P.
2. $E[D^\top (p - v)] = d^\top (p - v) \leq 0$.
3. $Std[D^\top (p - v)] \leq \frac{\lambda}{\sqrt{\log(n \times \text{ number of vertices})}}$
4. $P\left(D^\top p > D^\top v + \lambda\right) \leq P\left(D^\top (p - v) > E[D^\top (p - v)] + \lambda\right) \leq \frac{1}{4n \times \text{ number of vertices}}$.

Upper Bound: "Validity of random inequality"

Observation

Let $w \in \mathbb{R}^n$. Then $E(D^\top w) = d^\top w$, $\mathrm{Var}(D^\top w) = \frac{1}{\alpha}\sum_{i=1}^n w_i^2 |d_i|$.

Claim: With high probability $D^\top x \leq (D^\top v + \lambda)$ is a valid inequality for P.

Proof

1. Fix a vertex p of P.
2. $E[D^\top (p - v)] = d^\top (p - v) \leq 0$.
3. $Std[D^\top (p - v)] \leq \frac{\lambda}{\sqrt{\log(n \times \text{ number of vertices})}}$
4. $P\left(D^\top p > D^\top v + \lambda\right) \leq P\left(D^\top (p - v) > E[D^\top (p - v)] + \lambda\right) \leq \frac{1}{4n \times \text{ number of vertices}}$.
5. Probability that atleast one vertex does not satisfy $D^\top x \leq D^\top (v + \lambda)$ $\leq \left(\frac{1}{4n \times \text{ number of vertices}} \times \text{ number of vertices}\right)$. (Union Bound)

Upper Bound: "Validity of random inequality"

Observation

Let $w \in \mathbb{R}^n$. Then $E(D^\top w) = d^\top w$, $\mathrm{Var}(D^\top w) = \frac{1}{\alpha}\sum_{i=1}^n w_i^2 |d_i|$.

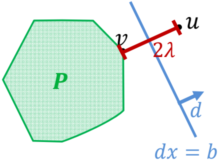
Claim: With high probability $D^\top x \leq (D^\top v + \lambda)$ is a valid inequality for P.

Proof

1. Fix a vertex p of P.
2. $E[D^\top(p - v)] = d^\top(p - v) \leq 0$.
3. $Std[D^\top(p - v)] \leq \frac{\lambda}{\sqrt{\log(n \times \text{ number of vertices})}}$
4. $P\Big(D^\top p > D^\top v + \lambda\Big) \leq P\Big(D^\top(p - v) > E[D^\top(p - v)] + \lambda\Big) \leq \frac{1}{4n \times \text{ number of vertices}}$.
5. Probability that atleast one vertex does not satisfy $D^\top x \leq D^\top(v + \lambda)$ $\leq \left(\frac{1}{4n \times \text{ number of vertices}} \times \text{ number of vertices}\right)$. (Union Bound)

$$P\Big(D^\top x \leq (D^\top v + \lambda) \text{ is a valid inequality}\Big) \geq 1 - \frac{1}{4n}.$$

Upper Bound: "Separation of u"



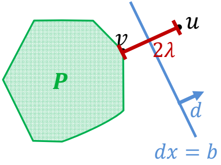
$$
\begin{aligned}
P\Big(D^\top u > (D^\top v + \lambda)\Big) &= P\Big(D^\top (u - v) > \lambda)\Big) \\
&= P\Big(D^\top (2\lambda d) > \lambda)\Big) \\
&= P\Big(D^\top d > 1/2)\Big)
\end{aligned}
$$

Claim: With probability at least $\frac{1}{2n-1}$, $D^\top d > 1/2$.

Proof

1. $E[D^\top d] = 1.$
2. $D^\top d \leq n$

Upper Bound: "Separation of u"



$$
\begin{aligned}
P\Big(D^\top u > (D^\top v + \lambda)\Big) &= P\Big(D^\top (u - v) > \lambda)\Big) \\
&= P\Big(D^\top (2\lambda d) > \lambda)\Big) \\
&= P\Big(D^\top d > 1/2)\Big)
\end{aligned}
$$

Claim: With probability at least $\frac{1}{2n-1}$, $D^\top d > 1/2$.

Proof

1. $E[D^\top d] = 1$.
2. $D^\top d \leq n$
3. With probability at least $\frac{1}{2n-1}$, $D^\top d > 1/2$. (Using Markov's inequality on the random variable $n - D^\top d$)

Upper Bound: Putting it all together

1. With probability at most $\frac{1}{4n}$, the inequality is not k-sparse.
2. With probability at most $\frac{1}{4n}$, the inequality is not valid.
3. With probability at most $1 - \frac{1}{2n-1}$, the inequality is does not separate u.

Upper Bound: Putting it all together

1. With probability at most $\frac{1}{4n}$, the inequality is not k-sparse.
2. With probability at most $\frac{1}{4n}$, the inequality is not valid.
3. With probability at most 1 - $\frac{1}{2n-1}$, the inequality is does not separate u.

With probablity at most $1 - \frac{1}{(2n-1)(2n)}$, the inequality is either not k-sparse, not valid or does not separate u.

$\Rightarrow$ With probablity at least $\frac{1}{(2n-1)(2n)} > 0$, the inequality is k-sparse, valid inequality and separates u.

Lower Bound

Theorem 2 [‘Matching’ Lower Bound on $d(P, P^k)$]

Let $k, t, n \in \mathbb{Z}_+$ satisfy

1. $64 \leq k \leq n$
2. $(0.5k^2 \log n + 2k + 1)^2 \leq t \leq e^n$.

Let $X^1, X^2, \ldots, X^t$ be independent uniformly random points in $\{0,1\}^n$ and let $P = \text{conv}(X^1, X^2, \ldots, X^t)$. Then with probability at least $1/4$ we have that

$$d(P, P^k) \geq \min\left\{\frac{\sqrt{n}}{\sqrt{k}}\frac{\sqrt{\log t}}{110\sqrt{\log n}}, \frac{\sqrt{n}}{8}\right\}\left(\frac{1}{2} - \frac{1}{k^{3/2}}\right) - 3\sqrt{\log t}. \qquad \square$$

Lower Bound

Theorem 2 ['Matching' Lower Bound on $d(P, P^k)$]

Let $k, t, n \in \mathbb{Z}_+$ satisfy

1. $64 \leq k \leq n$
2. $(0.5k^2 \log n + 2k + 1)^2 \leq t \leq e^n$.

Let $X^1, X^2, \ldots, X^t$ be independent uniformly random points in $\{0,1\}^n$ and let $P = \text{conv}(X^1, X^2, \ldots, X^t)$. Then with probability at least $1/4$ we have that

$$d(P, P^k) \geq \min\left\{ \frac{\sqrt{n}}{\sqrt{k}} \frac{\sqrt{\log t}}{110\sqrt{\log n}}, \frac{\sqrt{n}}{8} \right\} \left(\frac{1}{2} - \frac{1}{k^{3/2}} \right) - 3\sqrt{\log t}. \qquad \square$$

Upper bound $\approx \frac{\sqrt{n}}{\sqrt{k}} \sqrt{\log(n \times t)}$.

Lower Bound: Proof Outline

We show that there are points in P^k far from P.

1. With high probability the inequality $\sum_{i=1}^{n} x_i \leq \frac{n}{2} + 3\sqrt{n \log t}$ is valid for P.

Lower Bound: Proof Outline

We show that there are points in P^k far from P.

1. With high probability the inequality $\sum_{i=1}^{n} x_i \leq \frac{n}{2} + 3\sqrt{n \log t}$ is valid for P.
2. With high propobablity if the inequality $a^\top x \leq b$ is valid for P^k then $b \geq \left(1 + \frac{\sqrt{\log t}}{110\sqrt{k}\sqrt{\log n}}\left(1 - \frac{1}{k^2}\right)\right)\frac{\sum_{i=1}^{n} a_i}{2} - \frac{||a||_\infty}{2k^2}$

Lower Bound: Proof Outline

We show that there are points in P^k far from P.

1. With high probability the inequality $\sum_{i=1}^{n} x_i \leq \frac{n}{2} + 3\sqrt{n \log t}$ is valid for P.
2. With high propobablity if the inequality $a^\top x \leq b$ is valid for P^k then $b \geq \left(1 + \frac{\sqrt{\log t}}{110\sqrt{k}\sqrt{\log n}}\left(1 - \frac{1}{k^2}\right)\right) \frac{\sum_{i=1}^{n} a_i}{2} - \frac{||a||_\infty}{2k^2}$

3. The point $\sim \left(\frac{1}{2} + \frac{\sqrt{\log t}}{\sqrt{k}}\right)e$ belongs to P^k

Lower Bound: Proof Outline

We show that there are points in P^k far from P.

1. With high probability the inequality $\sum_{i=1}^{n} x_i \leq \frac{n}{2} + 3\sqrt{n \log t}$ is valid for P.
2. With high propobablity if the inequality $a^\top x \leq b$ is valid for P^k then $b \geq \left(1 + \frac{\sqrt{\log t}}{110\sqrt{k}\sqrt{\log n}}\left(1 - \frac{1}{k^2}\right)\right)\frac{\sum_{i=1}^{n} a_i}{2} - \frac{||a||_\infty}{2k^2}$
3. The point $\sim (\frac{1}{2} + \frac{\sqrt{\log t}}{\sqrt{k}})e$ belongs to P^k
4. The point $\sim (\frac{1}{2})e$ is closest point to above point in P. $\Rightarrow$

$$d(P, P^k) \approx \frac{\sqrt{n}}{\sqrt{k}}\sqrt{\log t}$$

Lower Bound: Proof Outline

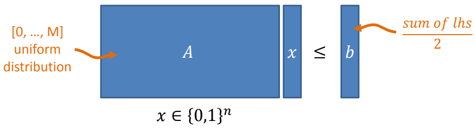
We show that there are points in P^k far from P.

1. With high probability the inequality $\sum_{i=1}^{n} x_i \leq \frac{n}{2} + 3\sqrt{n \log t}$ is valid for P.
2. With high propobablity if the inequality $a^\top x \leq b$ is valid for P^k then $b \geq \left(1 + \frac{\sqrt{\log t}}{110\sqrt{k}\sqrt{\log n}}\left(1 - \frac{1}{k^2}\right)\right)\frac{\sum_{i=1}^{n} a_i}{2} - \frac{||a||_\infty}{2k^2}$
    - 2.1 New anti-concentration inequality (approx): $P\left(ax \geq E[ax] + \gamma\frac{|a|_1}{\sqrt{n}}\right) \geq e^{-\gamma^2}$
    - 2.2 If $c^\top x \leq f$ is a facet-defining inequality of a $0/1$ polytope in R^k, then $||c||_\infty \leq k^{k/2}$.
3. The point $\sim (\frac{1}{2} + \frac{\sqrt{\log t}}{\sqrt{k}})e$ belongs to P^k
4. The point $\sim (\frac{1}{2})e$ is closest point to above point in P. $\Rightarrow$

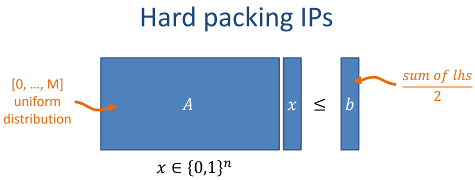
$$d(P, P^k) \approx \frac{\sqrt{n}}{\sqrt{k}}\sqrt{\log t}$$

Hard Packing Instances

Hard packing IPs



Hard Packing Instances



Theorem 3

Consider $n, m, M \in N$ such that $n \geq 50$ and $8 \log 8n \leq m \leq n$. Then with probability at least $1/2$,
$d(P, P^k) \geq \frac{\sqrt{n}}{2}\left(\frac{2}{\max\{\alpha,1\}}(1-\epsilon)^2 - (1+\epsilon')\right)$, where $c = k/n$ and

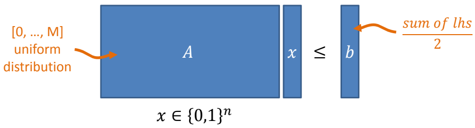
$$\frac{1}{\alpha} = \frac{M}{2(M+1)}\left[\frac{n - 2\sqrt{n\log 8m}}{c((2-c)n+1) + 2\sqrt{10cnm}}\right], \quad \epsilon = \frac{24\sqrt{\log 4n^2 m}}{\sqrt{n}},$$

$$\epsilon' = \frac{3\sqrt{\log 8n}}{\sqrt{m} - 2\sqrt{\log 8n}}.$$

□

Hard Packing Instances

Hard packing IPs



Theorem 3

Consider $n, m, M \in N$ such that $n \geq 50$ and $8 \log 8n \leq m \leq n$. Then with probability at least $1/2$,
$d(P, P^k) \geq \frac{\sqrt{n}}{2}\left(\frac{2}{\max\{\alpha, 1\}}(1-\epsilon)^2 - (1+\epsilon')\right)$, where $c = k/n$ and

$$\frac{1}{\alpha} = \frac{M}{2(M+1)}\left[\frac{n - 2\sqrt{n \log 8m}}{c((2-c)n+1) + 2\sqrt{10cnm}}\right], \quad \epsilon = \frac{24\sqrt{\log 4n^2 m}}{\sqrt{n}},$$

$$\epsilon' = \frac{3\sqrt{\log 8n}}{\sqrt{m} - 2\sqrt{\log 8n}}. \qquad \square$$

$\sim$ With probably $1/2$, $d(P, P^k) \gtrsim \sqrt{n}\left(\frac{n}{k} - 1\right)$ for $k \geq n/2$.

Sparse cutting-planes and extended formulations

Let $\text{proj}_x : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ denote the projection operator onto the first n coordinates.

Proposition

Consider a polyhedron $P \subseteq \mathbb{R}^n$ and an extended formulation $Q \subseteq \mathbb{R}^n \times \mathbb{R}^m$ for it. Then $\text{proj}_x(Q^k) \subseteq (\text{proj}_x(Q))^k = P^k$.

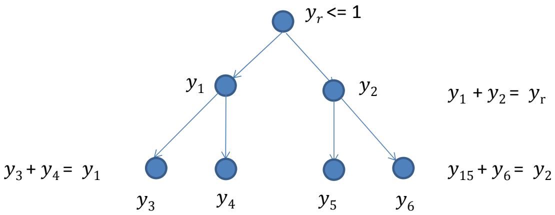
Proposition

Consider $n \in N$ and assume it is a power of 2. Then there is a polytope $P \subseteq \mathbb{R}^n$ such that:

1. *$d(P, P^k) = \sqrt{n/2}$ for all $k \leq n/2$.*
2. *There is an extended formulation $Q \subseteq \mathbb{R}^n \times \mathbb{R}^{2n-1}$ of P such that $\text{proj}_x(Q^3) = P$.*

Extended Formulation

Original Polytope:
$P = \{x \in [0\ 1]^n \mid \sum_{i=1}^n x_i \leq n/2\}$



Extended Formulation:

$$
\begin{aligned}
& y_r \leq 1 \\
& y_v = y_{\text{left}(v)} + y_{\text{right}(v)}, \quad \forall v \in \text{int}(T) \\
& y_v = \frac{2}{n} x_{i(v)}, \quad \forall v \in T \setminus \text{int}(T) \\
& y_v \geq 0, \quad \forall v \in T \\
& x_i \in [0, 1], \quad \forall i \in [n].
\end{aligned} \tag{1}
$$

Thank You!