Extended Formulations and Information Complexity

Ankur Moitra
Massachusetts Institute of Technology

Dagstuhl, March 2014

The Permutahedron

The Permutahedron

Let $\vec{t} = [1, 2, 3, \ldots n]$, $P = \text{conv}\left\{\pi(\vec{t}) \,\middle|\, \pi \text{ is permutation}\right\}$

The Permutahedron

Let $\vec{t} = [1, 2, 3, \ldots n]$, $P = \text{conv}\{\pi(\vec{t}) \mid \pi \text{ is permutation}\}$

How many facets of P have?

The Permutahedron

Let $\vec{t} = [1, 2, 3, \ldots n]$, $P = \text{conv}\{\pi(\vec{t}) \mid \pi \text{ is permutation}\}$

How many facets of P have?

exponentially many!

The Permutahedron

Let $\vec{t} = [1, 2, 3, \ldots n]$, $P = \text{conv}\{\pi(\vec{t}) \mid \pi \text{ is permutation}\}$

How many facets of P have?

exponentially many!

e.g. $S \subset [n]$, $\Sigma_{i \text{ in } S}\, x_i \geq 1 + 2 + \ldots + |S| = |S|(|S|+1)/2$

The Permutahedron

Let $\vec{t} = [1, 2, 3, \ldots n]$, $P = \text{conv}\{\pi(\vec{t}) \mid \pi \text{ is permutation}\}$

How many facets of P have?

exponentially many!

e.g. $S \subset [n]$, $\Sigma_{i \text{ in } S}\, x_i \geq 1 + 2 + \ldots + |S| = |S|(|S|+1)/2$

Let $Q = \{A \mid A \text{ is doubly-stochastic}\}$

The Permutahedron

Let $\vec{t} = [1, 2, 3, \dots n]$, $P = \text{conv}\{\pi(\vec{t}) \mid \pi \text{ is permutation}\}$

How many facets of P have?

exponentially many!

e.g. $S \subset [n]$, $\Sigma_{i \text{ in } S}\, x_i \geq 1 + 2 + \dots + |S| = |S|(|S|+1)/2$

Let $Q = \{A \mid A \text{ is doubly-stochastic}\}$

Then P is the projection of Q: $P = \{A\vec{t} \mid A \text{ in } Q\}$

The Permutahedron

Let $\vec{t} = [1, 2, 3, \ldots n]$, $P = \text{conv}\{\pi(\vec{t}) \mid \pi \text{ is permutation}\}$

How many facets of P have?

exponentially many!

e.g. $S \subset [n]$, $\Sigma_{i \text{ in } S}\, x_i \geq 1 + 2 + \ldots + |S| = |S|(|S|+1)/2$

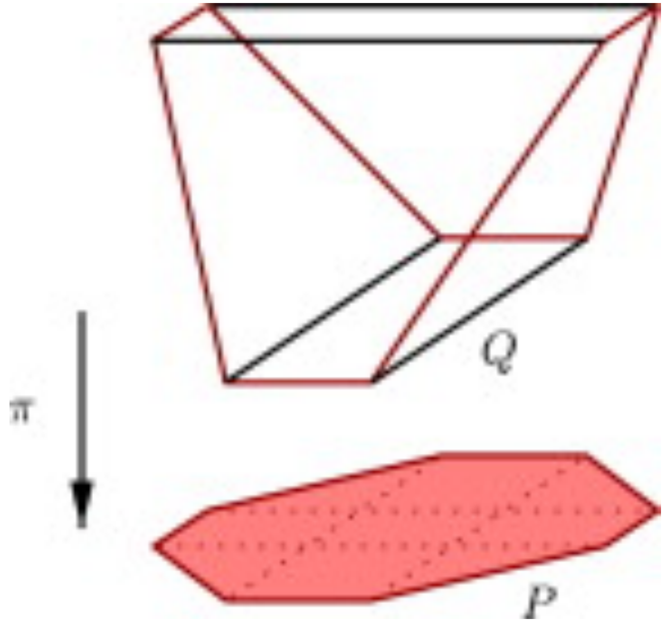
Let $Q = \{A \mid A \text{ is doubly-stochastic}\}$

Then P is the projection of Q: $P = \{A\vec{t} \mid A \text{ in } Q\}$

Yet Q has only $O(n^2)$ facets

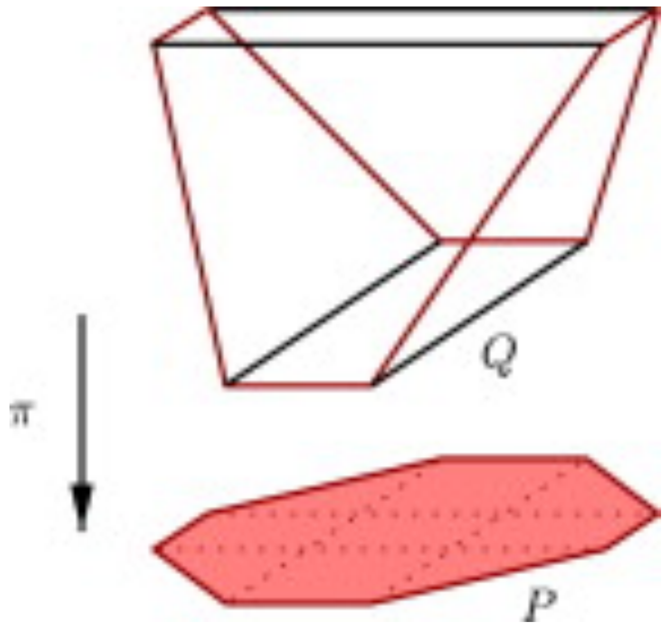
Extended Formulations

The **extension complexity (xc)** of a polytope P is the minimum number of facets of Q so that P = proj(Q)



Extended Formulations

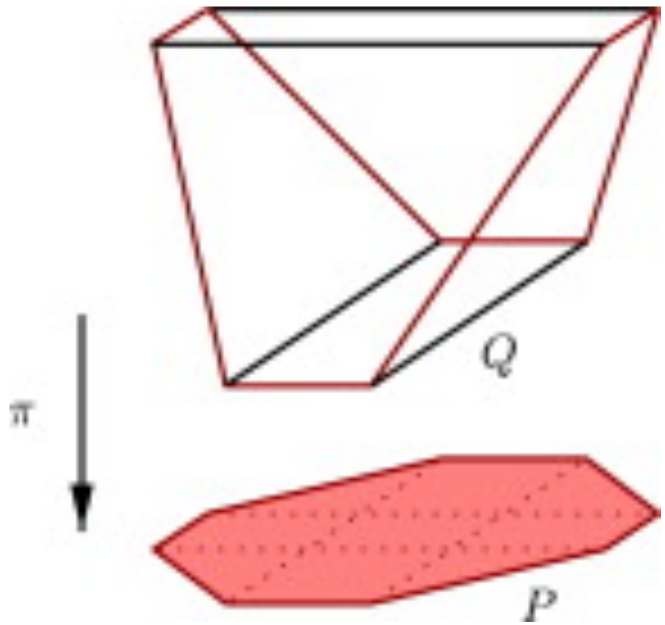
The **extension complexity (xc)** of a polytope P is the minimum number of facets of Q so that P = proj(Q)



e.g. xc(P) = $\Theta(n \log n)$
for permutahedron

Extended Formulations

The **extension complexity (xc)** of a polytope P is the minimum number of facets of Q so that P = proj(Q)

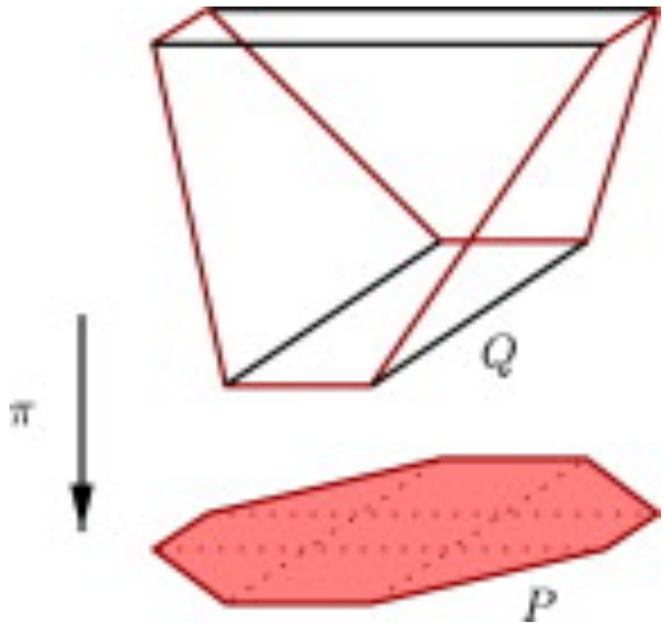


e.g. xc(P) = Θ(n logn) for permutahedron

xc(P) = Θ(logn) for a regular n-gon, but Ω(√n) for its perturbation

Extended Formulations

The **extension complexity (xc)** of a polytope P is the minimum number of facets of Q so that P = proj(Q)



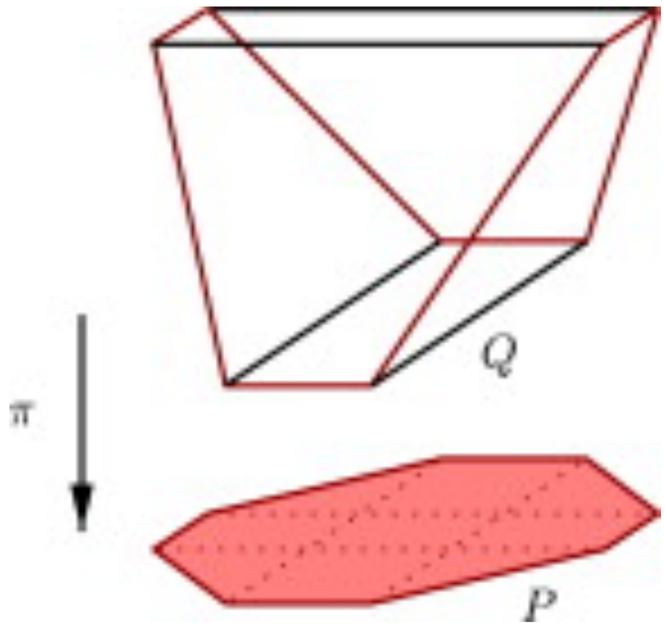
e.g. $xc(P) = \Theta(n \log n)$ for permutahedron

$xc(P) = \Theta(\log n)$ for a regular n-gon, but $\Omega(\sqrt{n})$ for its perturbation

In general, $P = \{x \mid \exists y, (x,y) \text{ in } Q\}$

Extended Formulations

The **extension complexity (xc)** of a polytope P is the minimum number of facets of Q so that P = proj(Q)



e.g. xc(P) = Θ(n logn) for permutahedron

xc(P) = Θ(logn) for a regular n-gon, but $\Omega(\sqrt{n})$ for its perturbation

In general, P = {x | $\exists$ y, (x,y) in Q}

…analogy with **quantifiers** in Boolean formulae

Applications of EFs

In general, P = $\{x \mid \exists y, (x,y) \text{ in } Q\}$

Applications of EFs

In general, $P = \{x \mid \exists y, (x,y) \text{ in } Q\}$

Through EFs, we can reduce # facets exponentially!

Applications of EFs

In general, $P = \{x \mid \exists y, (x,y) \text{ in } Q\}$

Through EFs, we can reduce # facets exponentially!

Hence, we can run standard LP solvers instead of the ellipsoid algorithm

Applications of EFs

In general, $P = \{x \mid \exists y, (x,y) \text{ in } Q\}$

Through EFs, we can reduce # facets exponentially!

Hence, we can run standard LP solvers instead of the ellipsoid algorithm

EFs often give, or are based on new combinatorial insights

Applications of EFs

In general, $P = \{x \mid \exists y, (x,y) \text{ in } Q\}$

Through EFs, we can reduce # facets exponentially!

Hence, we can run standard LP solvers instead of the ellipsoid algorithm

EFs often give, or are based on new combinatorial insights

e.g. Birkhoff-von Neumann Thm and permutahedron

Applications of EFs

In general, P = $\{x \mid \exists y, (x,y) \text{ in } Q\}$

Through EFs, we can reduce # facets exponentially!

Hence, we can run standard LP solvers instead of the ellipsoid algorithm

EFs often give, or are based on new combinatorial insights

e.g. Birkhoff-von Neumann Thm and permutahedron

e.g. prove there is low-cost object, through its polytope

Explicit, Hard Polytopes?

Explicit, Hard Polytopes?

Definition: TSP polytope:

$$P = \text{conv}\{\mathbf{1}_F \,|\, F \text{ is the set of edges on a tour of } K_n\}$$

Explicit, Hard Polytopes?

Definition: TSP polytope:

$$P = \text{conv}\{\mathbf{1}_F \mid F \text{ is the set of edges on a tour of } K_n\}$$

(If we could optimize over this polytope, then P = NP)

Explicit, Hard Polytopes?

Definition: TSP polytope:

$$P = \text{conv}\{\mathbf{1}_F \mid F \text{ is the set of edges on a tour of } K_n\}$$

(If we could optimize over this polytope, then P = NP)

Can we prove **unconditionally** there is no small EF?

Explicit, Hard Polytopes?

Definition: TSP polytope:

$$P = \text{conv}\{\mathbf{1}_F \mid F \text{ is the set of edges on a tour of } K_n\}$$

(If we could optimize over this polytope, then P = NP)

Can we prove **unconditionally** there is no small EF?

Caveat: this is unrelated to proving complexity l.b.s

Explicit, Hard Polytopes?

Definition: TSP polytope:

$$P = \text{conv}\{\mathbf{1}_F \mid F \text{ is the set of edges on a tour of } K_n\}$$

(If we could optimize over this polytope, then P = NP)

Can we prove **unconditionally** there is no small EF?

Caveat: this is unrelated to proving complexity l.b.s

[Yannakakis '90]: Yes, through the **nonnegative rank**

An Abridged History

Theorem [Yannakakis '90]: Any symmetric EF for TSP or matching has size $2^{\Omega(n)}$

An Abridged History

Theorem [Yannakakis '90]: Any symmetric EF for TSP or matching has size $2^{\Omega(n)}$

…but asymmetric EFs can be more powerful

An Abridged History

Theorem [Yannakakis '90]: Any symmetric EF for TSP or matching has size $2^{\Omega(n)}$

…but asymmetric EFs can be more powerful

⋮ ⋮

An Abridged History

Theorem [Yannakakis '90]: Any symmetric EF for TSP or matching has size $2^{\Omega(n)}$

...but asymmetric EFs can be more powerful

⋮ ⋮

Theorem [Fiorini et al '12]: Any EF for TSP has size $2^{\Omega(\sqrt{n})}$ (based on a $2^{\Omega(n)}$ lower bd for clique)

Approach: connections to non-deterministic CC

An Abridged History II

Theorem [Braun et al '12]: Any EF that approximates clique within $n^{1/2-eps}$ has size $exp(n^{eps})$

Approach: Razborov's rectangle corruption lemma

An Abridged History II

Theorem [Braun et al ’12]: Any EF that approximates clique within $n^{1/2-eps}$ has size $exp(n^{eps})$

Approach: Razborov’s rectangle corruption lemma

Theorem [Braverman, Moitra ’13]: Any EF that approximates clique within n^{1-eps} has size $exp(n^{eps})$

Approach: information complexity

An Abridged History II

Theorem [Braun et al '12]: Any EF that approximates clique within $n^{1/2-eps}$ has size $\exp(n^{eps})$

Approach: Razborov's rectangle corruption lemma

Theorem [Braverman, Moitra '13]: Any EF that approximates clique within n^{1-eps} has size $\exp(n^{eps})$

Approach: information complexity

see also **[Braun, Pokutta '13]:** reformulation using common information, applications to avg. case

An Abridged History III

Theorem [Chan et al '12]: Any EF that approximates MAXCUT within 2-eps has size $n^{\Omega(\log n / \log\log n)}$

Approach: reduction to Sherali-Adams

An Abridged History III

Theorem [Chan et al '12]: Any EF that approximates MAXCUT within 2-eps has size $n^{\Omega(\log n/\log\log n)}$

Approach: reduction to Sherali-Adams

Theorem [Rothvoss '13]: Any EF for perfect matching has size $2^{\Omega(n)}$ (same for TSP)

Approach: hyperplane separation lower bound

Outline

Part I: Tools for Extended Formulations

- Yannakakis's Factorization Theorem
- The Rectangle Bound
- A Sampling Argument

Part II: Applications

- Correlation Polytope
- Approximating the Correlation Polytope
- Matching Polytope

Outline

Part I: Tools for Extended Formulations

- **Yannakakis's Factorization Theorem**
- The Rectangle Bound
- A Sampling Argument

Part II: Applications

- Correlation Polytope
- Approximating the Correlation Polytope
- Matching Polytope

The Factorization Theorem

The Factorization Theorem

How can we prove lower bounds on EFs?

The Factorization Theorem

How can we prove lower bounds on EFs?

[Yannakakis '90]:

Geometric Parameter ⟷ Algebraic Parameter

The Factorization Theorem

How can we prove lower bounds on EFs?

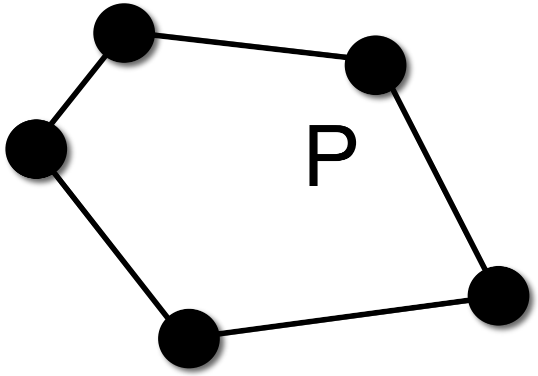
[Yannakakis ’90]:

Geometric Parameter ⟷ Algebraic Parameter

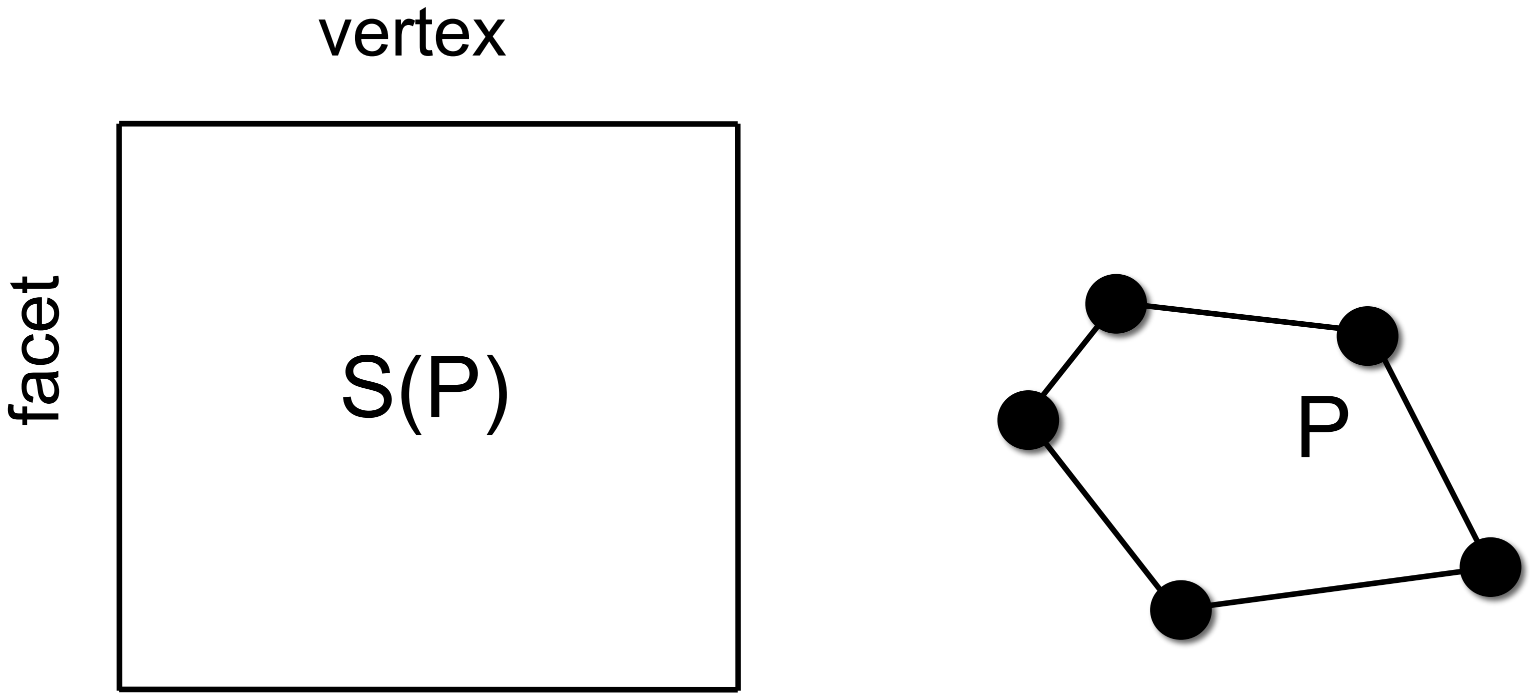
Definition of the **slack matrix**…

The Slack Matrix

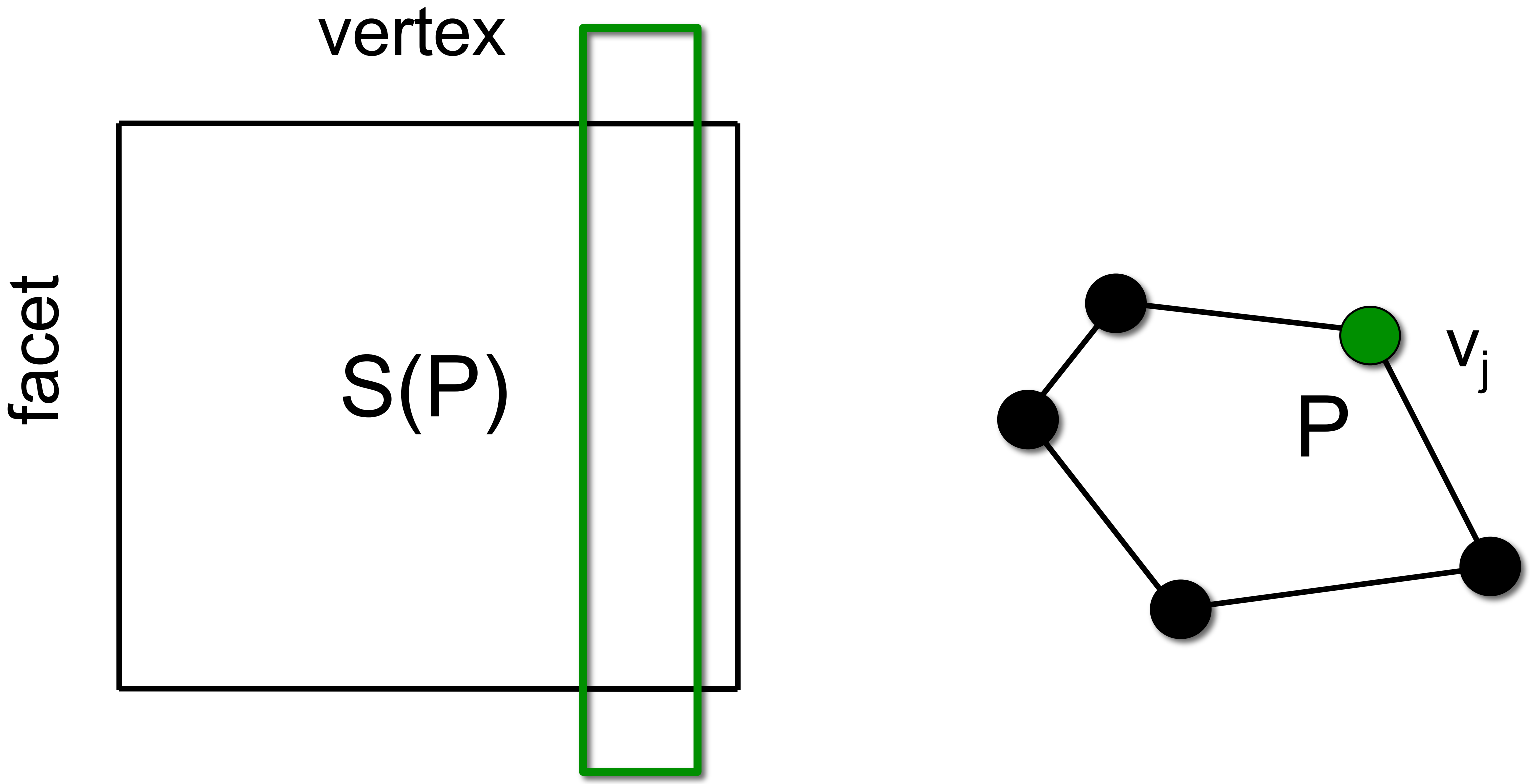
The Slack Matrix



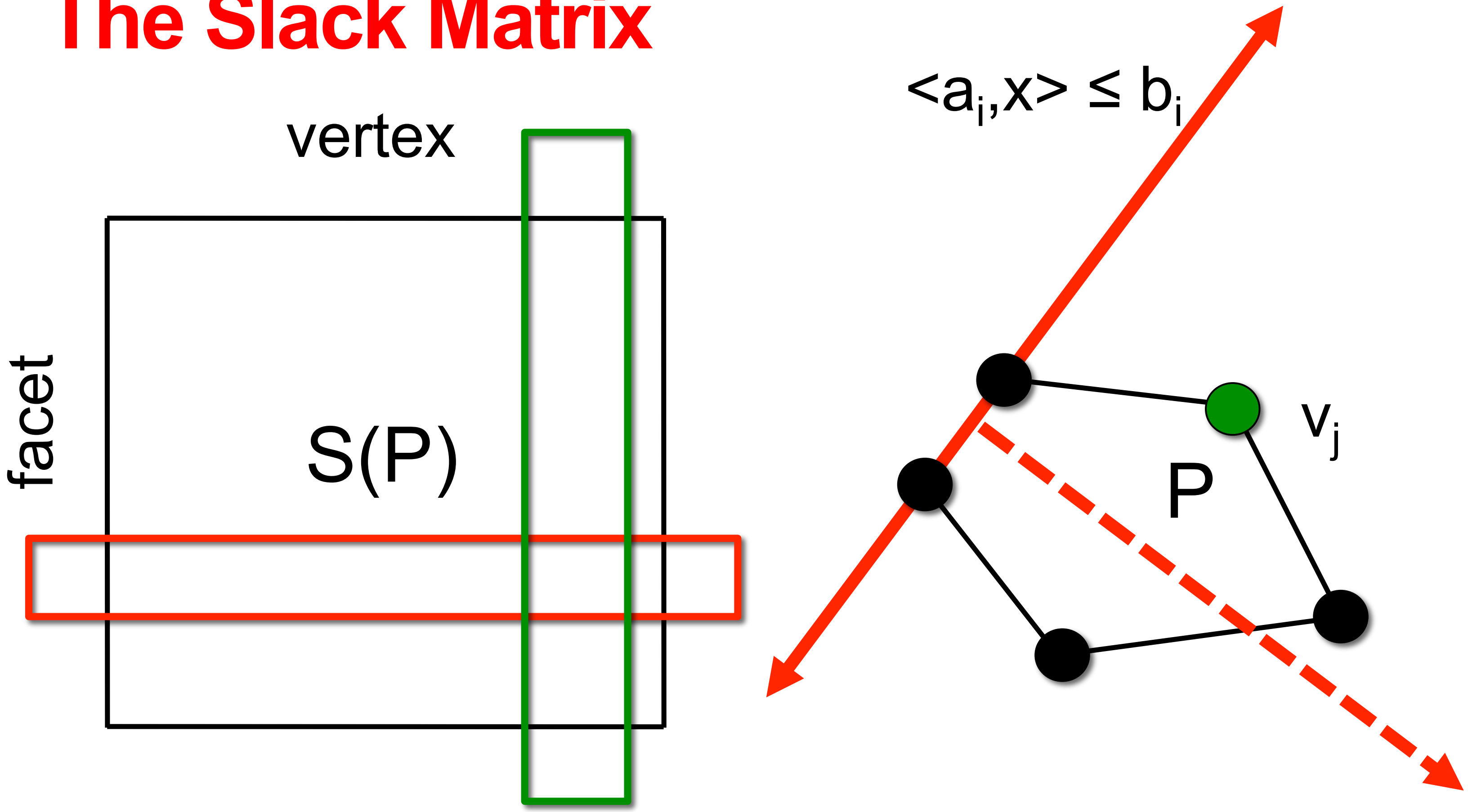
The Slack Matrix



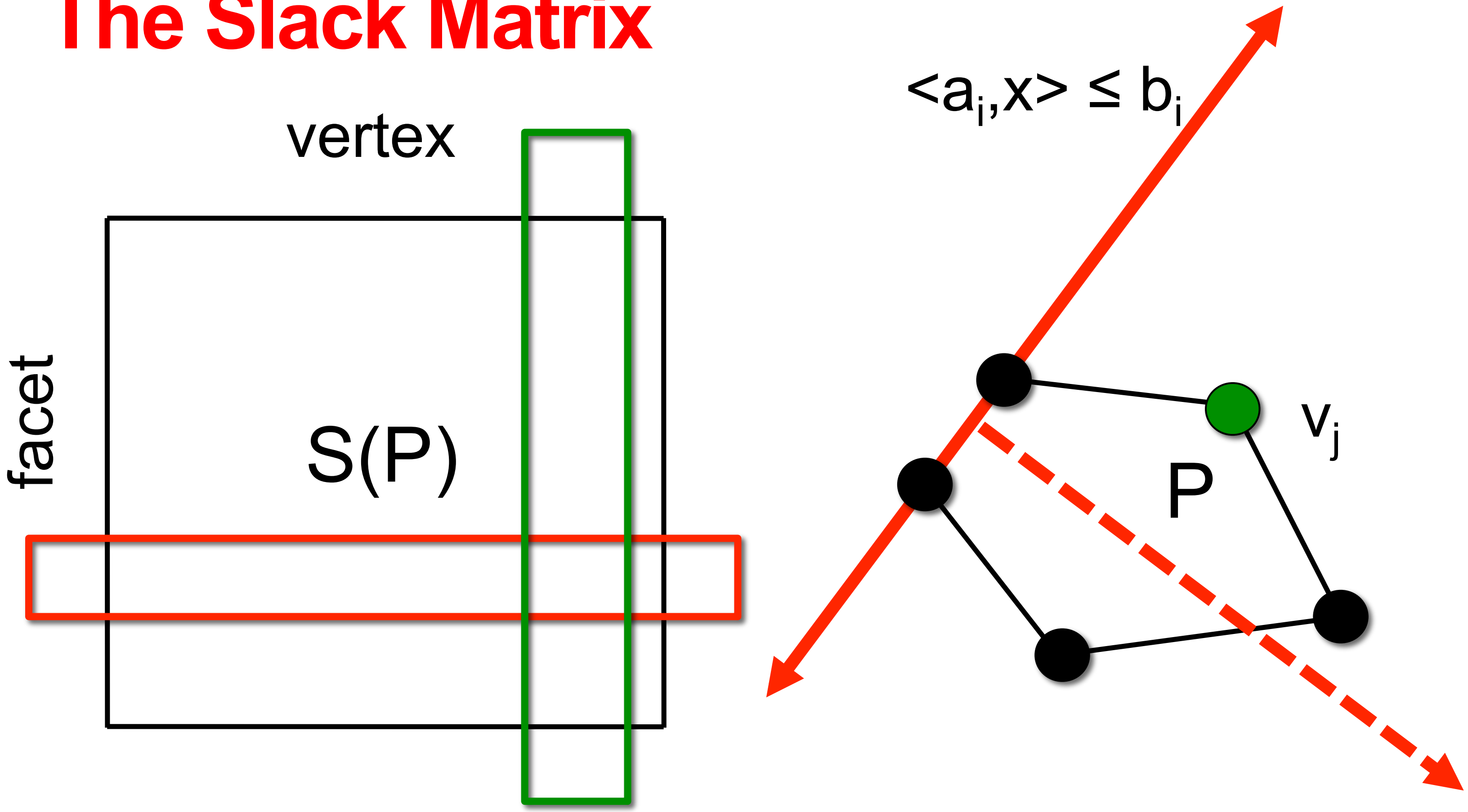
The Slack Matrix



The Slack Matrix

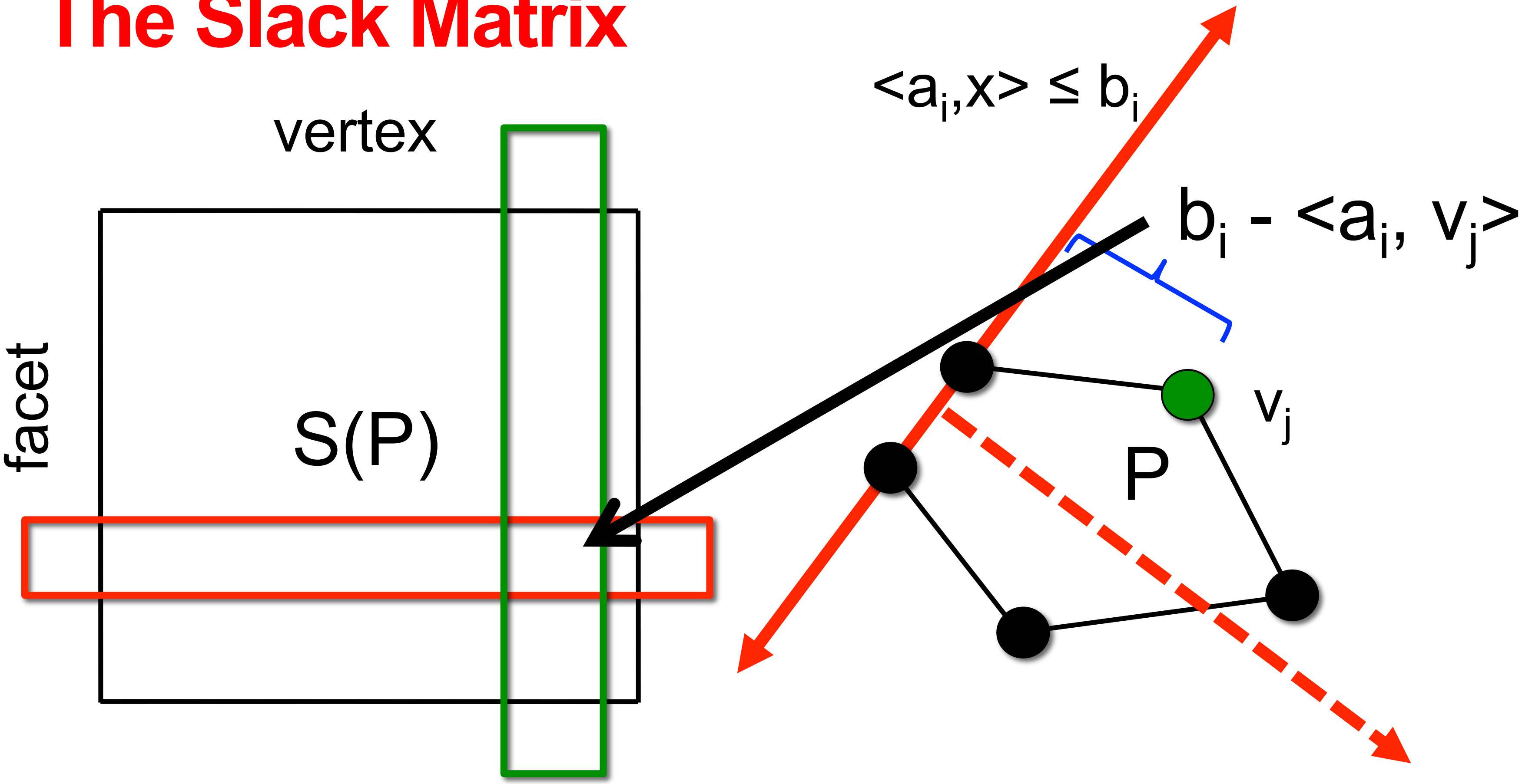


The Slack Matrix



The entry in row i, column j is how *slack* the j^{th} vertex is on the i^{th} constraint

The Slack Matrix



The entry in row i, column j is how *slack* the j^{th} vertex is on the i^{th} constraint

The Factorization Theorem

How can we prove lower bounds on EFs?

[Yannakakis ’90]:

Geometric Parameter	⟷	Algebraic Parameter

Definition of the **slack matrix**…

The Factorization Theorem

How can we prove lower bounds on EFs?

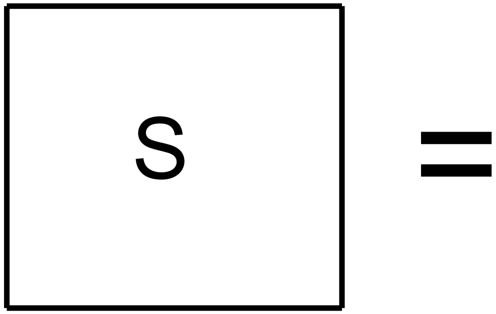
[Yannakakis '90]:

Geometric Parameter ⟺ Algebraic Parameter

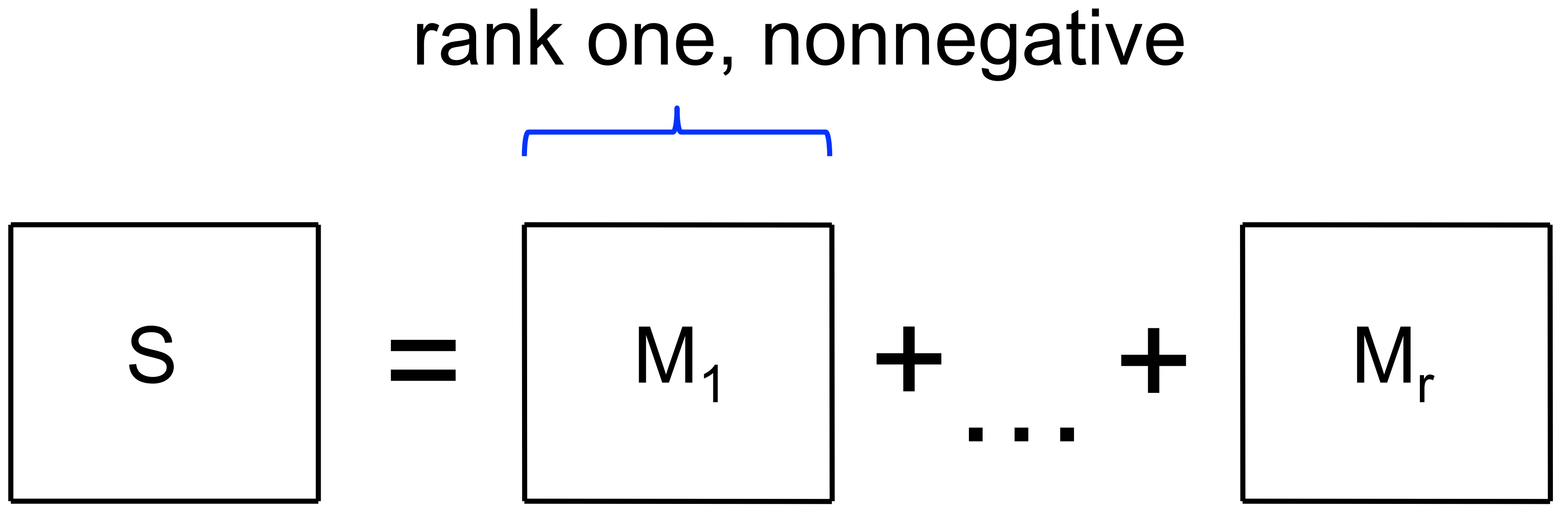
Definition of the **slack matrix**…

Definition of the **nonnegative rank**…

Nonnegative Rank



Nonnegative Rank



Nonnegative Rank

rank one, nonnegative

$$S = M_1 + \ldots + M_r$$

Definition: rank$^+$(S) is the smallest r s.t. S can be written as the sum of r rank one, nonneg. matrices

Nonnegative Rank

rank one, nonnegative

$$S = \underbrace{M_1}_{} + \dots + M_r$$

Definition: rank$^+$(S) is the smallest r s.t. S can be written as the sum of r rank one, nonneg. matrices

Note: rank$^+$(S) $\geq$ rank(S), but can be much larger too!

The Factorization Theorem

How can we prove lower bounds on EFs?

[Yannakakis '90]:

Geometric Parameter ⟷ Algebraic Parameter

The Factorization Theorem

How can we prove lower bounds on EFs?

[Yannakakis '90]: $\mathbf{xc(P) = rank^+(S(P))}$

Geometric Parameter ⟺ Algebraic Parameter

The Factorization Theorem

How can we prove lower bounds on EFs?

[Yannakakis '90]: $\mathbf{xc(P) = rank^+(S(P))}$

Geometric Parameter ⟺ Algebraic Parameter

Intuition: the factorization gives a change of variables that preserves the slack matrix!

The Factorization Theorem

How can we prove lower bounds on EFs?

[Yannakakis '90]: $\mathbf{xc(P) = rank^+(S(P))}$

Geometric Parameter ⟷ Algebraic Parameter

Intuition: the factorization gives a change of variables that preserves the slack matrix!

Next we will give a method to lower bound $rank^+$ via **information complexity**…

Outline

Part I: Tools for Extended Formulations

- Yannakakis's Factorization Theorem
- The Rectangle Bound
- A Sampling Argument

Part II: Applications

- Correlation Polytope
- Approximating the Correlation Polytope
- Matching Polytope

Outline

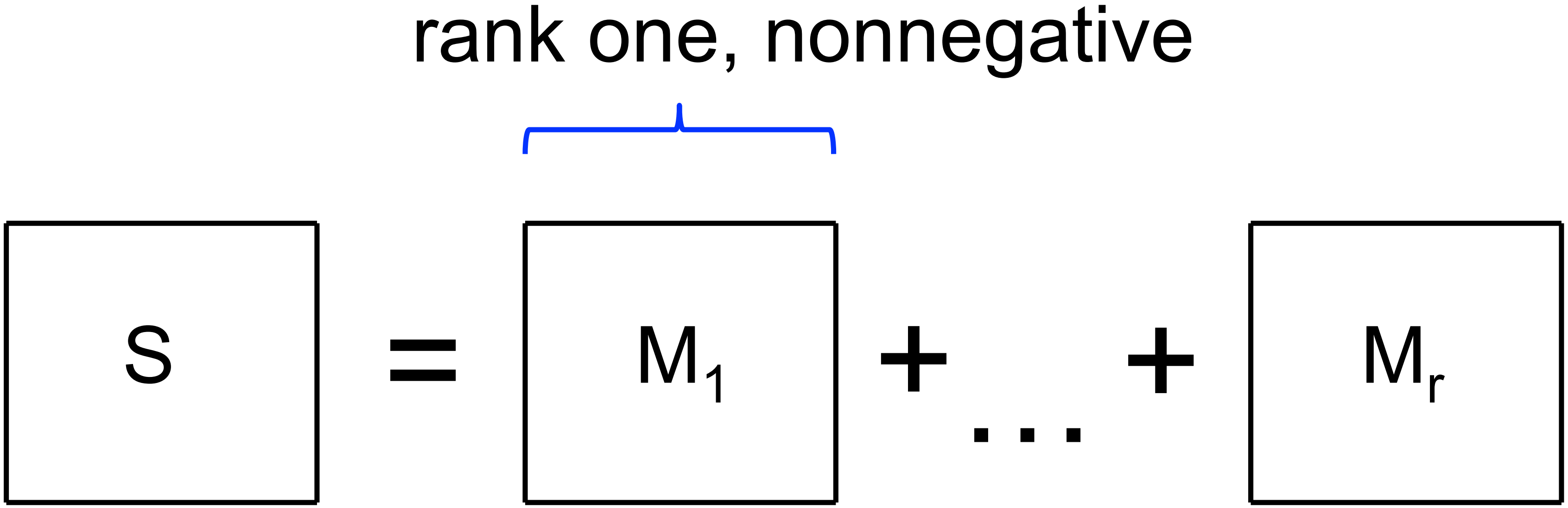
Part I: Tools for Extended Formulations

- Yannakakis's Factorization Theorem
- **The Rectangle Bound**
- A Sampling Argument

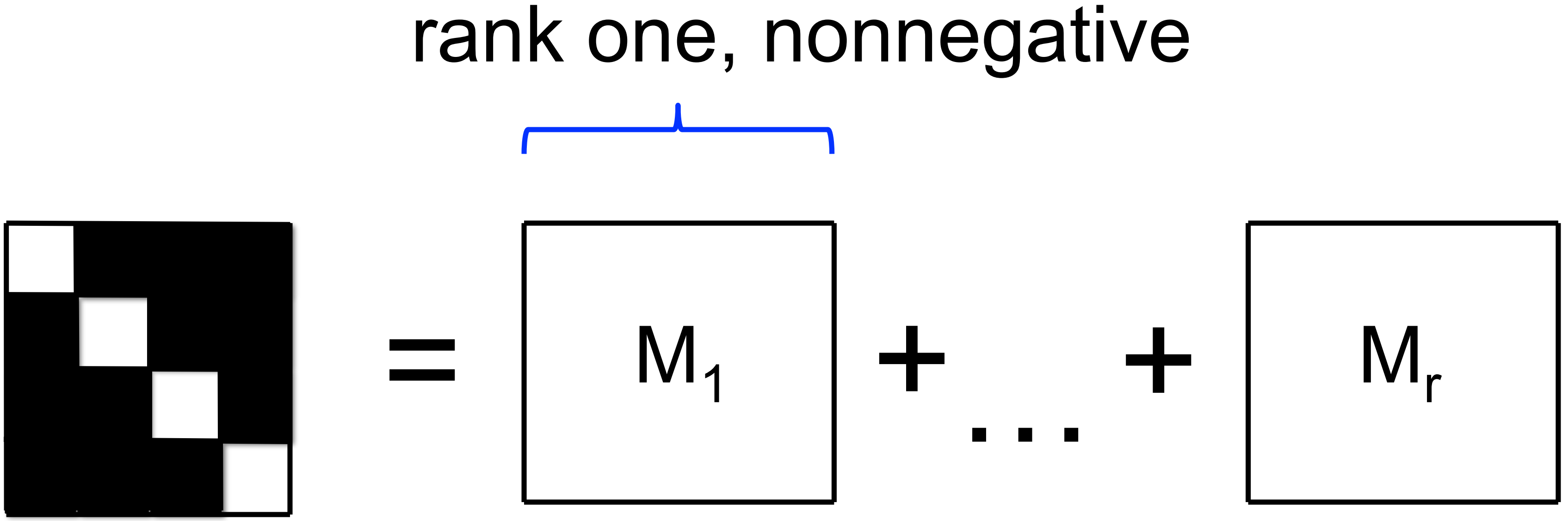
Part II: Applications

- Correlation Polytope
- Approximating the Correlation Polytope
- Matching Polytope

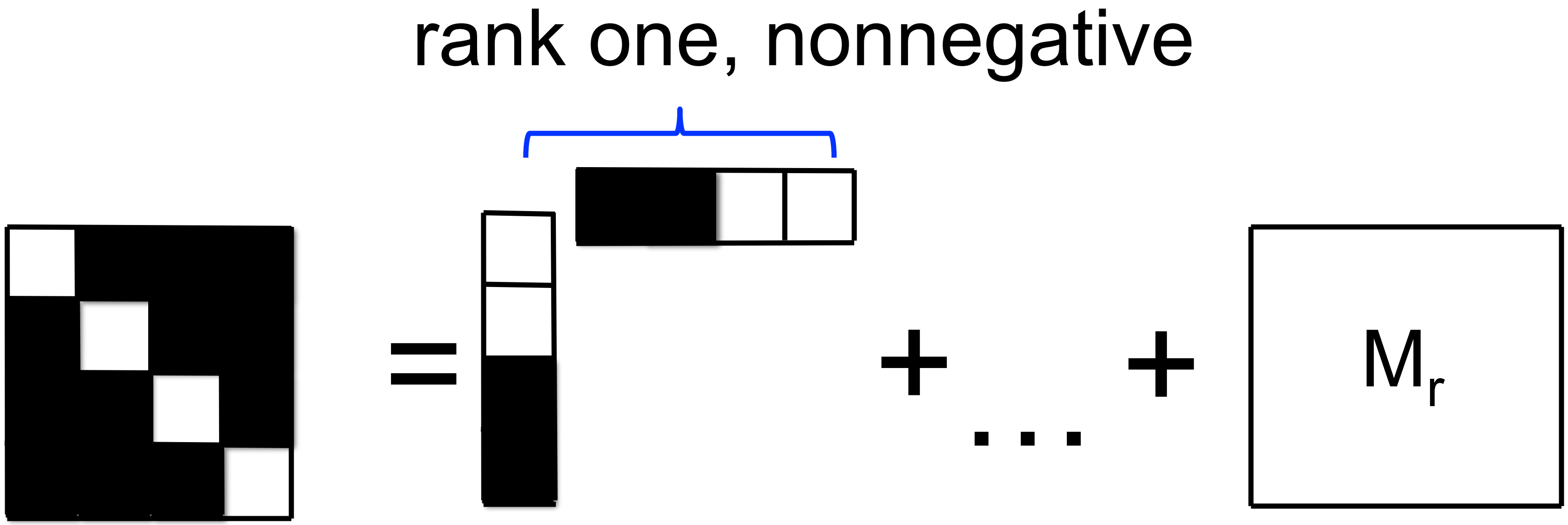
The Rectangle Bound



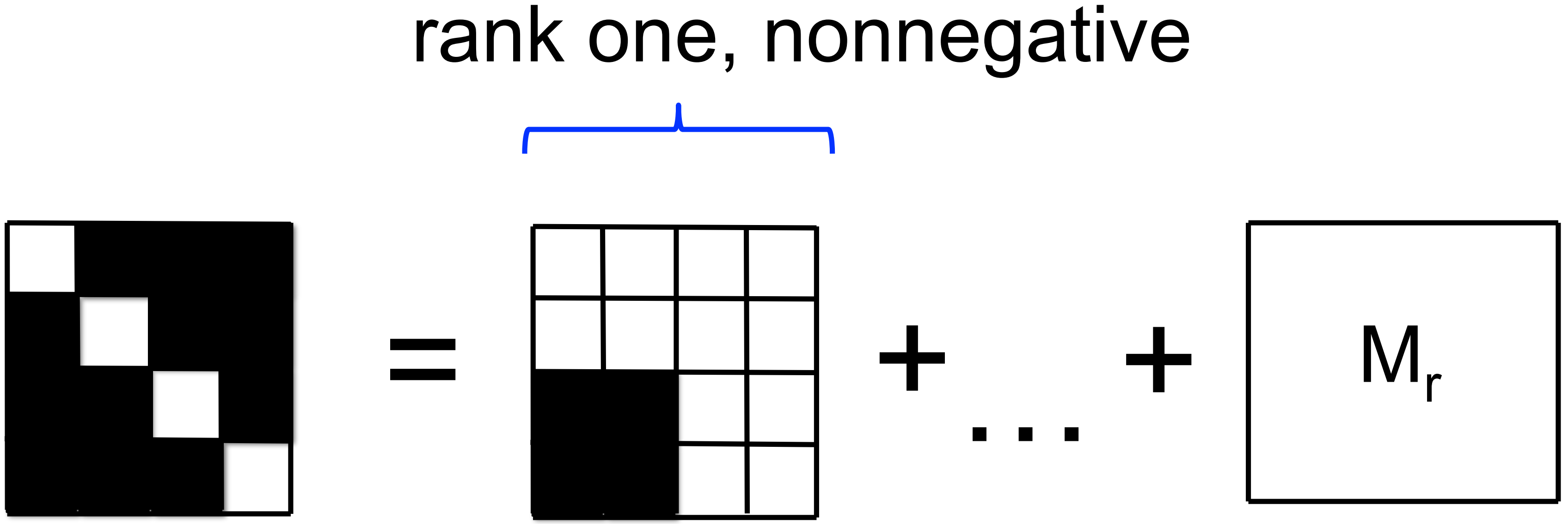
The Rectangle Bound



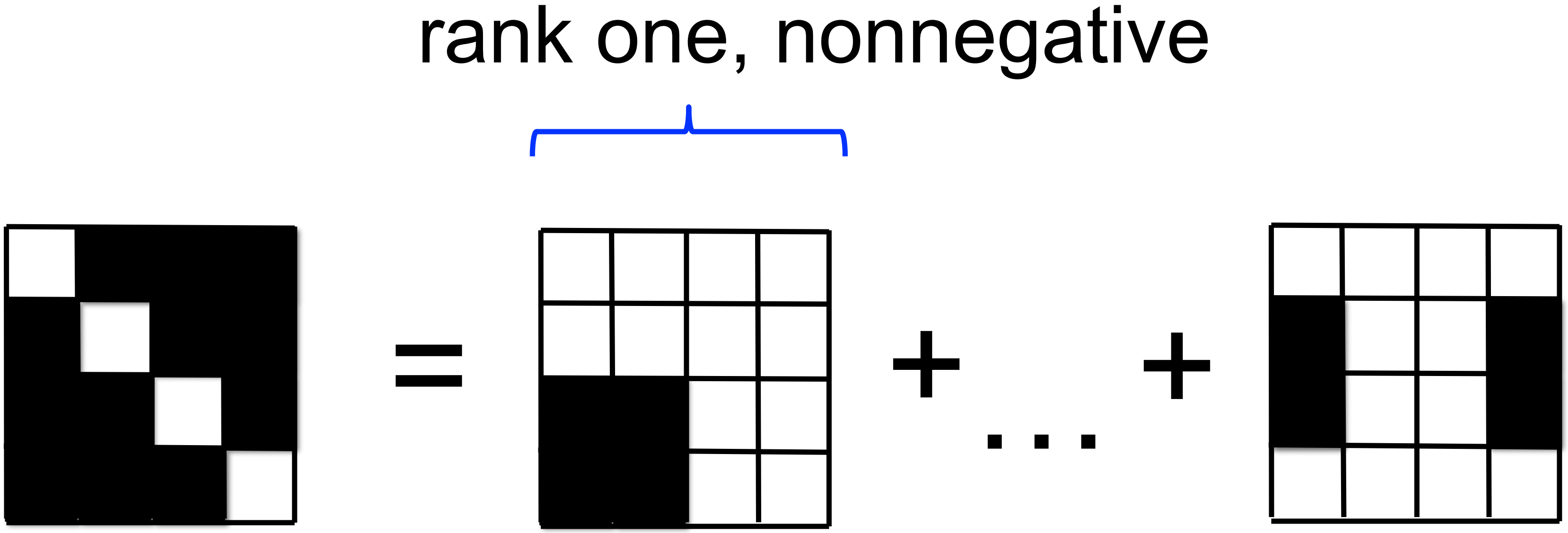
The Rectangle Bound



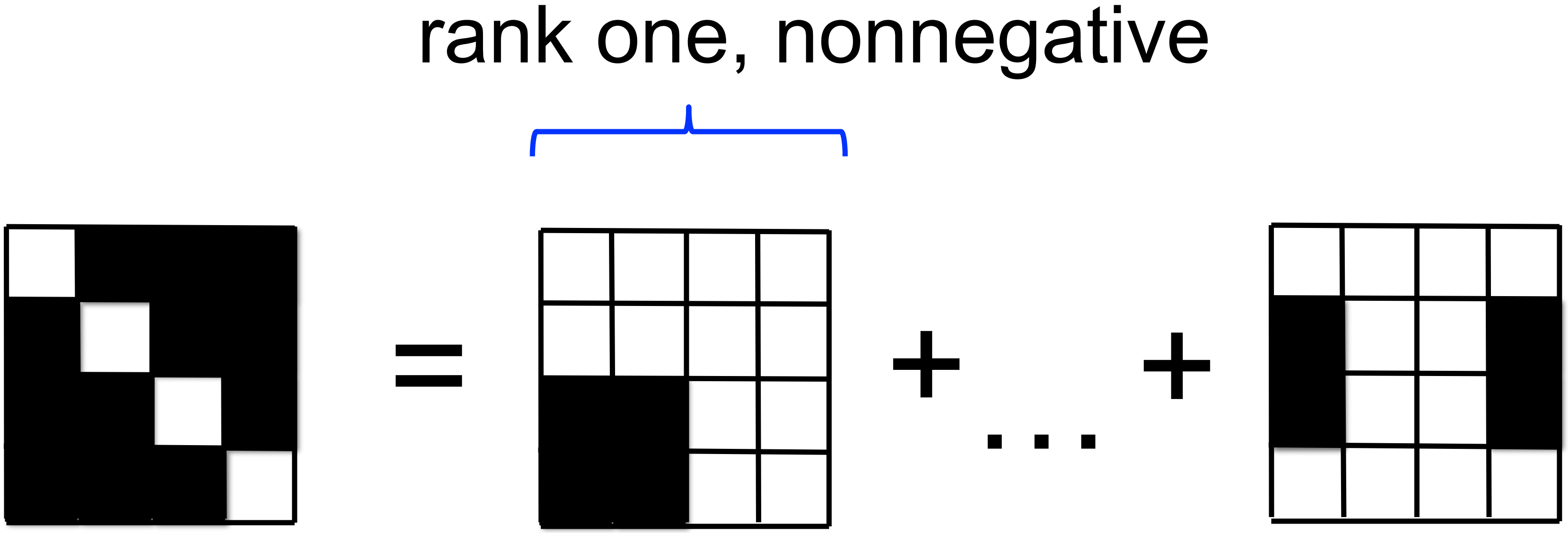
The Rectangle Bound



The Rectangle Bound

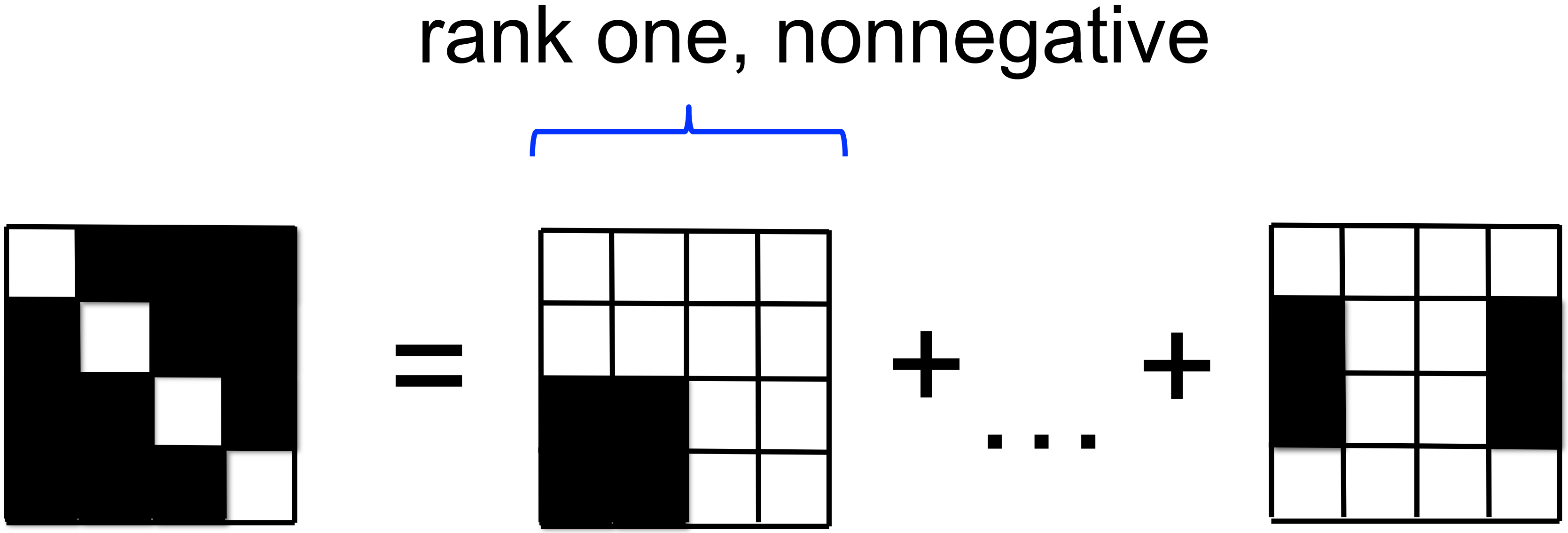


The Rectangle Bound



The support of each M_i is a combinatorial rectangle

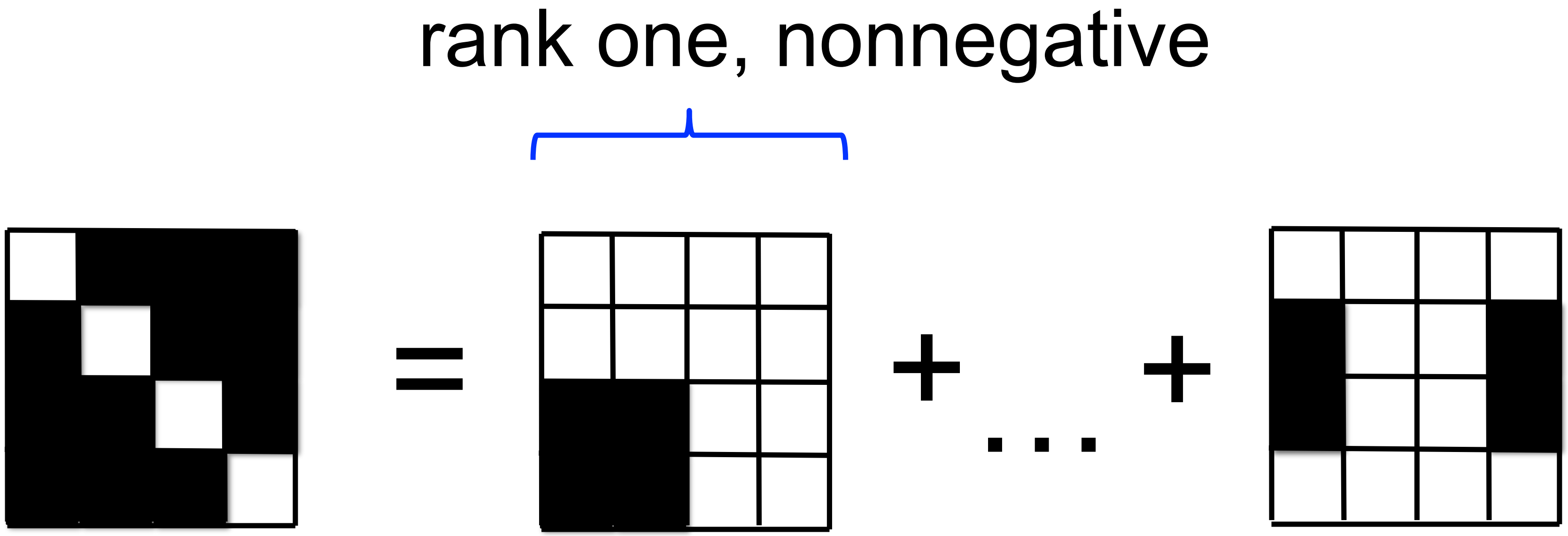
The Rectangle Bound



The support of each M_i is a combinatorial rectangle

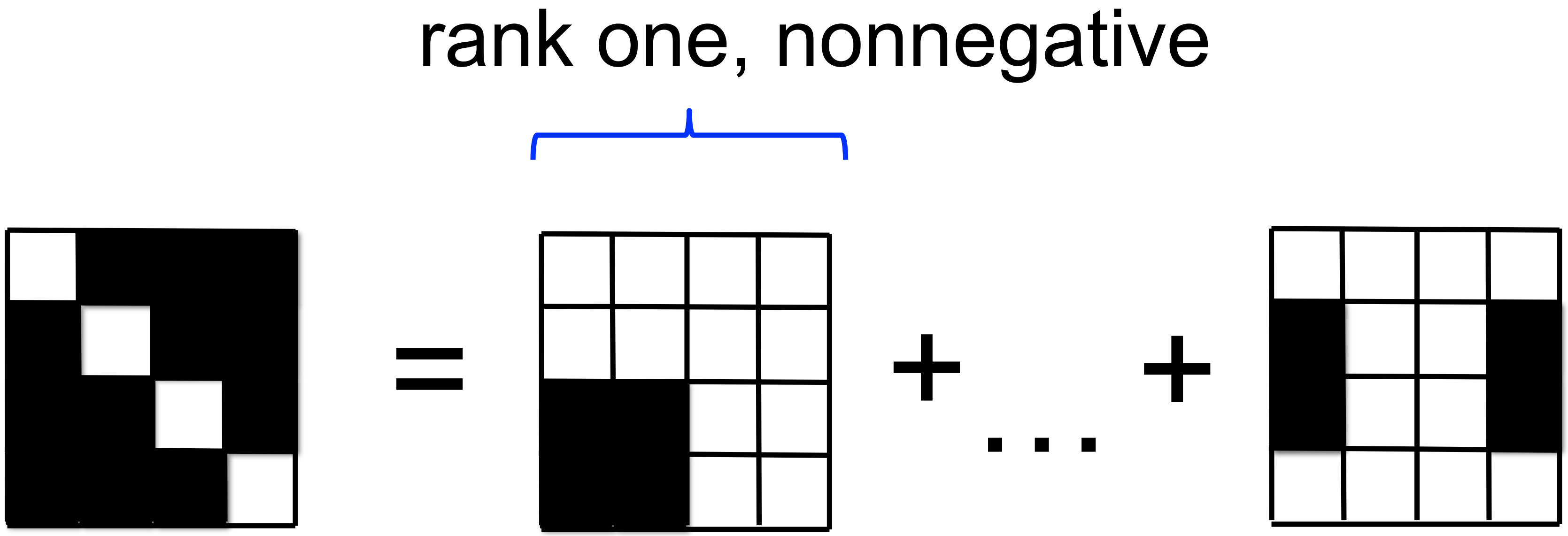
$rank^+(S)$ is at least # rectangles needed to cover supp of S

The Rectangle Bound



$rank^{+}(S)$ is at least # rectangles needed to cover supp of S

The Rectangle Bound



Non-deterministic Comm. Complexity

rank⁺(S) is at least # rectangles needed to cover supp of S

Outline

Part I: Tools for Extended Formulations

- Yannakakis's Factorization Theorem
- The Rectangle Bound
- A Sampling Argument

Part II: Applications

- Correlation Polytope
- Approximating the Correlation Polytope
- Matching Polytope

Outline

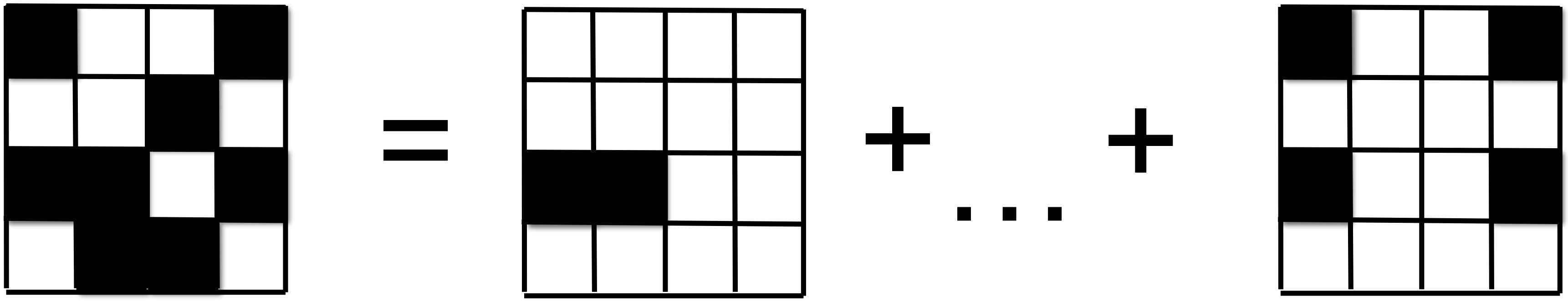
Part I: Tools for Extended Formulations

- Yannakakis's Factorization Theorem
- The Rectangle Bound
- **A Sampling Argument**

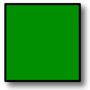
Part II: Applications

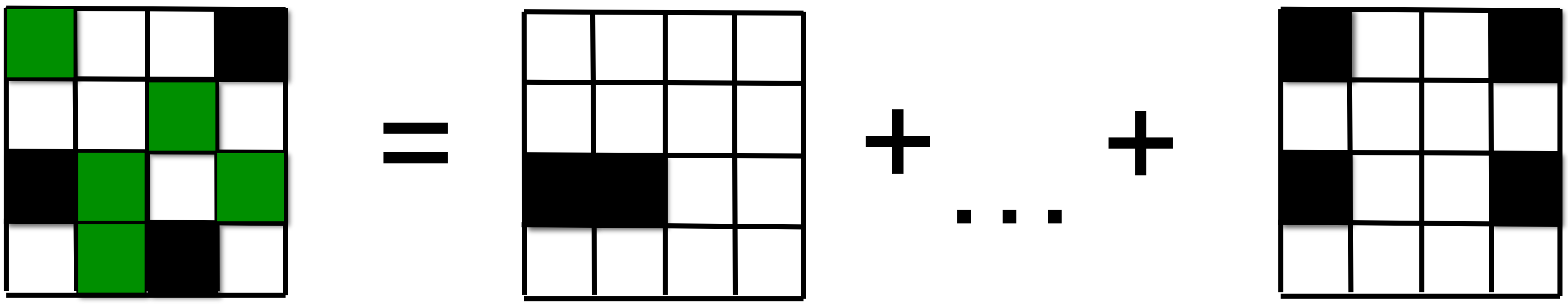
- Correlation Polytope
- Approximating the Correlation Polytope
- Matching Polytope

A Sampling Argument

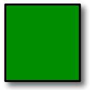


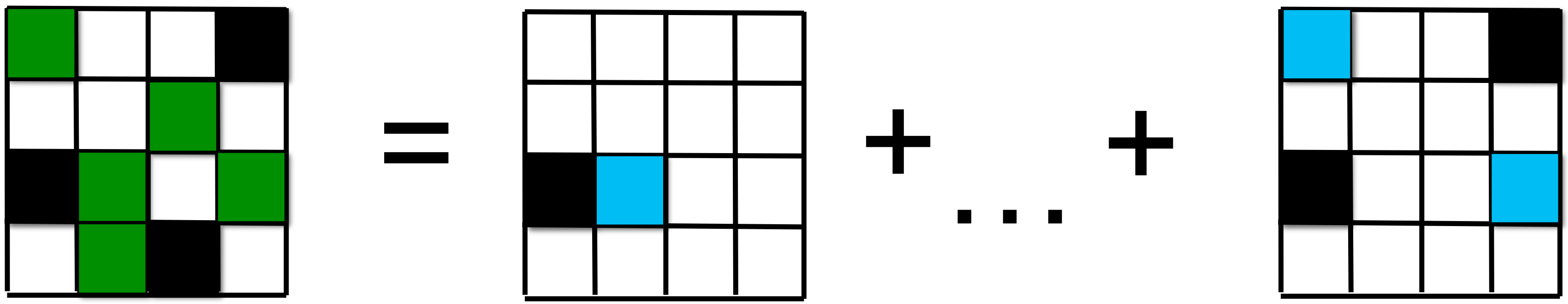
A Sampling Argument

T = {}, set of entries in S with same value

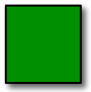


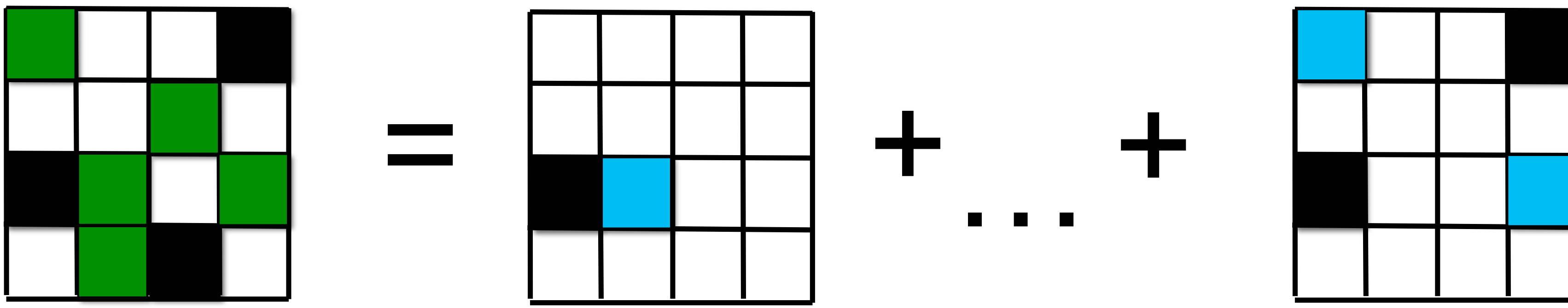
A Sampling Argument

T = {}, set of entries in S with same value



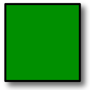
A Sampling Argument

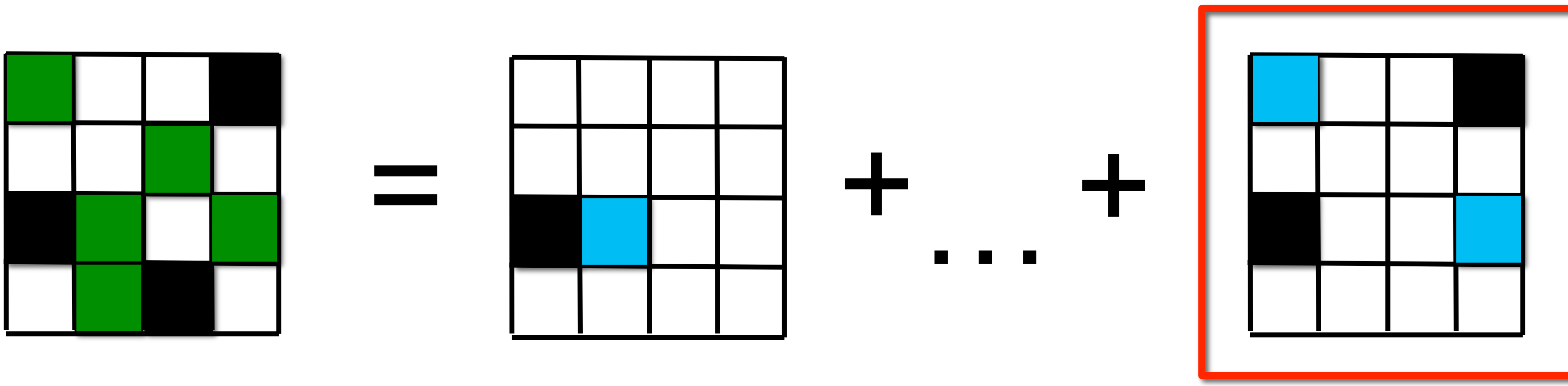
T = {}, set of entries in S with same value



Choose M_i proportional to total value on T

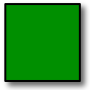
A Sampling Argument

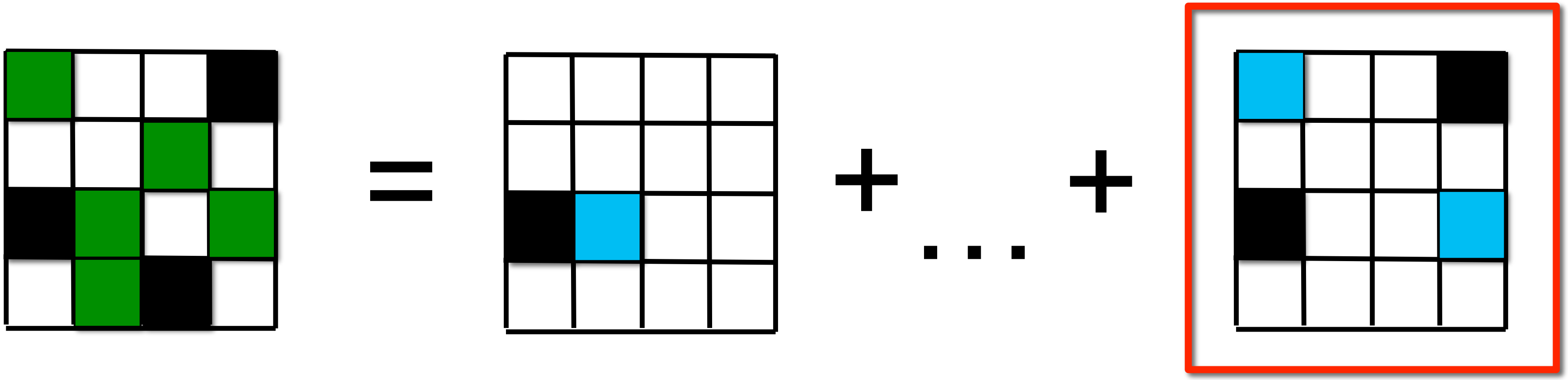
T = {}, set of entries in S with same value



Choose M_i proportional to total value on T

A Sampling Argument

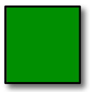
T = {}, set of entries in S with same value

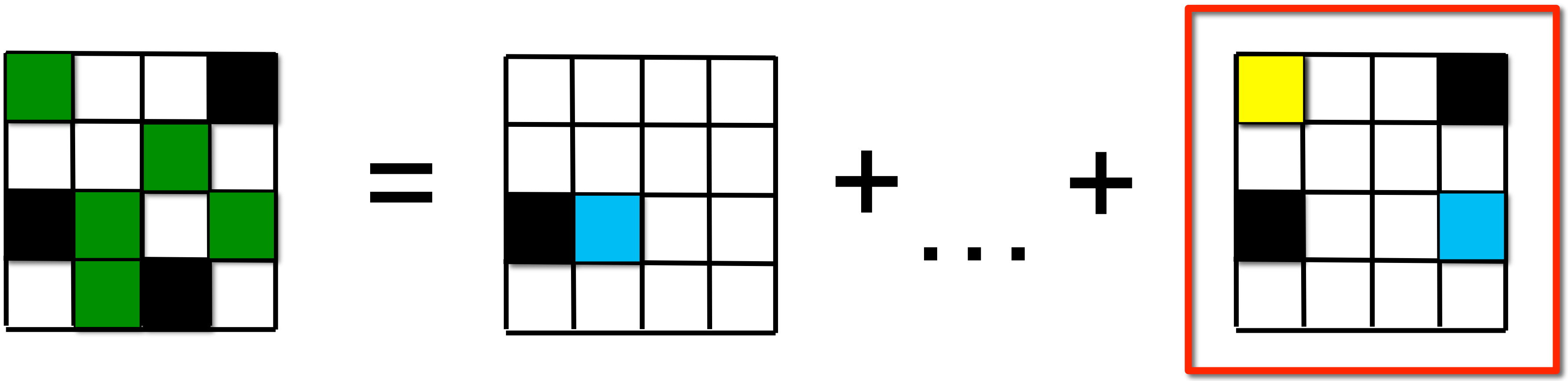


Choose M_i proportional to total value on T

Choose (a,b) in T proportional to relative value in M_i

A Sampling Argument

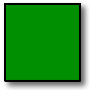
T = {}, set of entries in S with same value

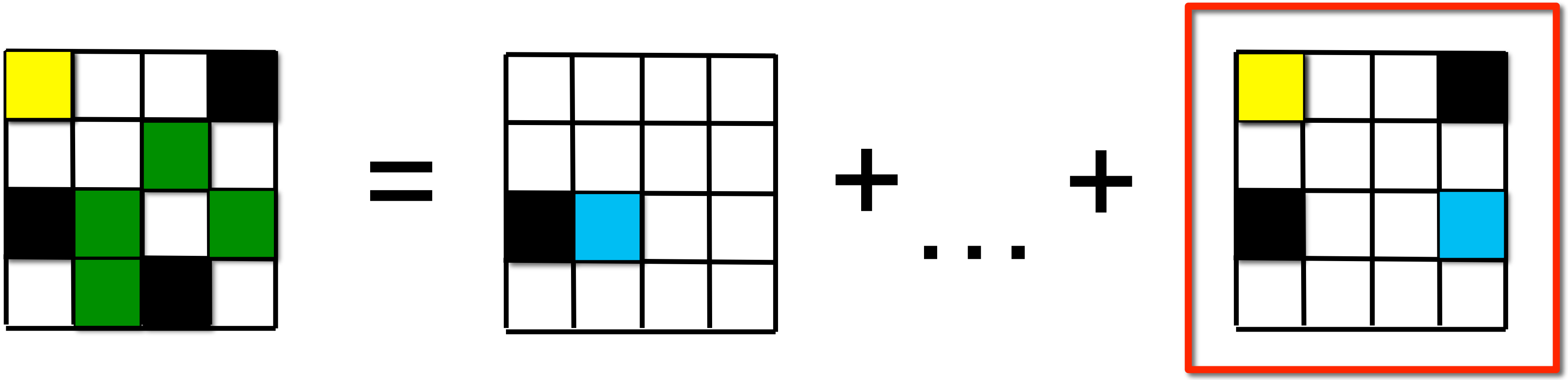


Choose M_i proportional to total value on T

Choose (a,b) in T proportional to relative value in M_i

A Sampling Argument

T = {}, set of entries in S with same value

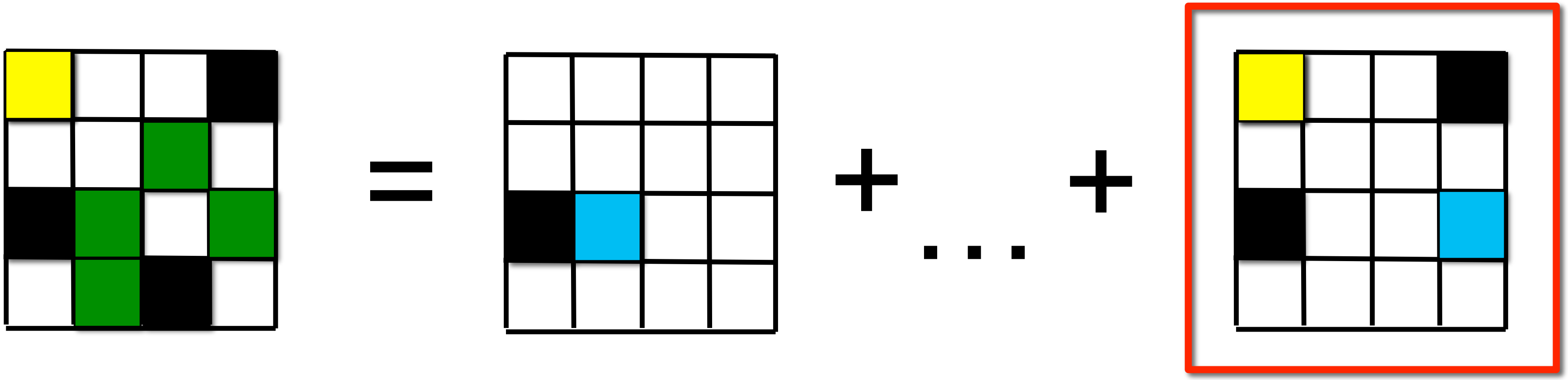


Choose M_i proportional to total value on T

Choose (a,b) in T proportional to relative value in M_i

A Sampling Argument

T = {■}, set of entries in S with same value

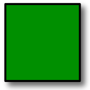


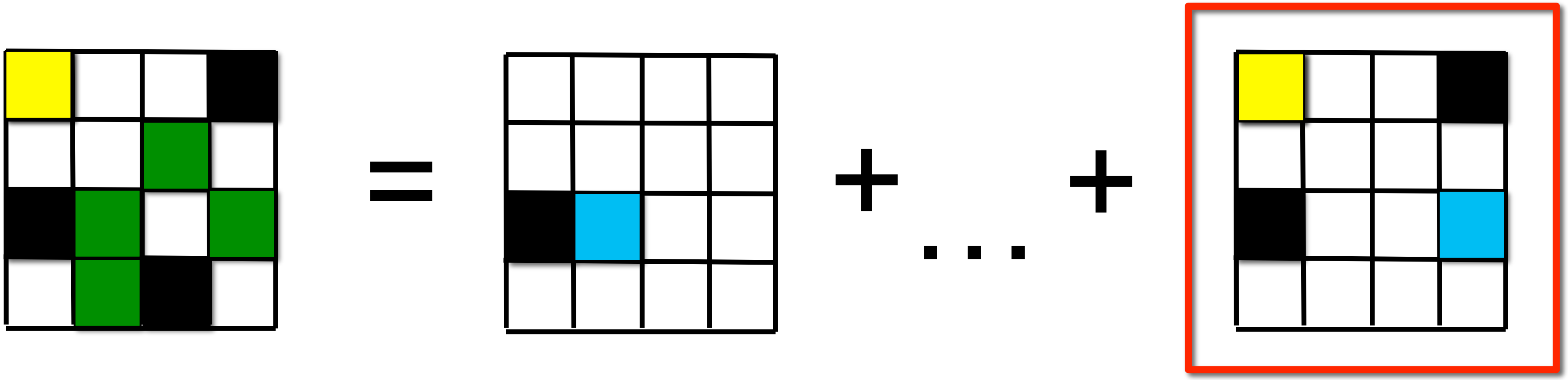
Choose M_i proportional to total value on T

Choose (a,b) in T proportional to relative value in M_i

This outputs a uniformly random sample from T

A Sampling Argument

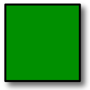
T = {  }, set of entries in S with same value

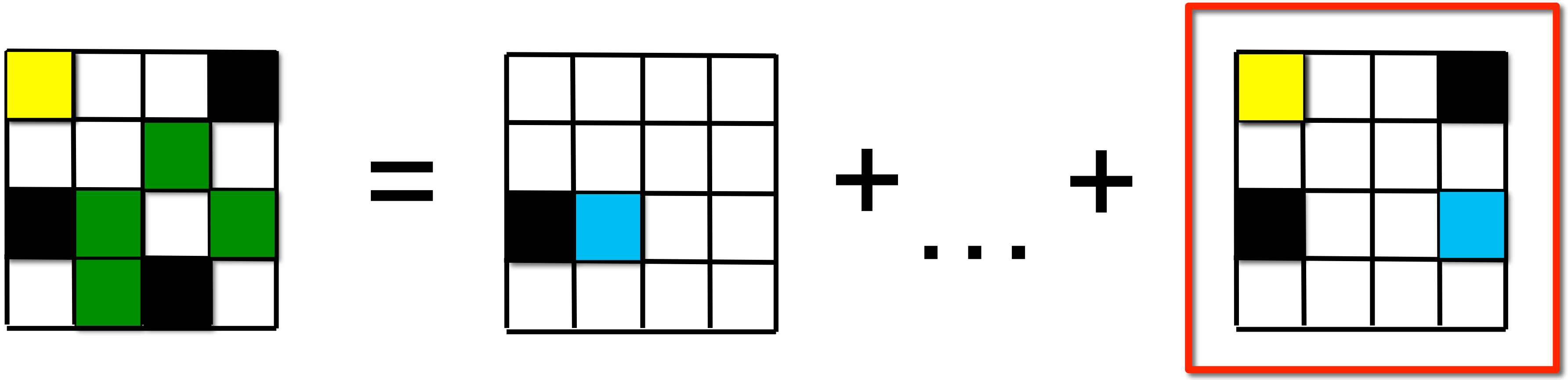


Choose M_i proportional to total value on T

Choose (a,b) in T proportional to relative value in M_i

A Sampling Argument

T = {}, set of entries in S with same value



Choose M_i proportional to total value on T

Choose (a,b) in T proportional to relative value in M_i

If r is too small, this procedure uses too little entropy!

Outline

Part I: Tools for Extended Formulations

- Yannakakis's Factorization Theorem
- The Rectangle Bound
- A Sampling Argument

Part II: Applications

- Correlation Polytope
- Approximating the Correlation Polytope
- Matching Polytope

Outline

Part I: Tools for Extended Formulations

- Yannakakis's Factorization Theorem
- The Rectangle Bound
- A Sampling Argument

Part II: Applications

- **Correlation Polytope**
- Approximating the Correlation Polytope
- Matching Polytope

The Construction of [Fiorini et al]

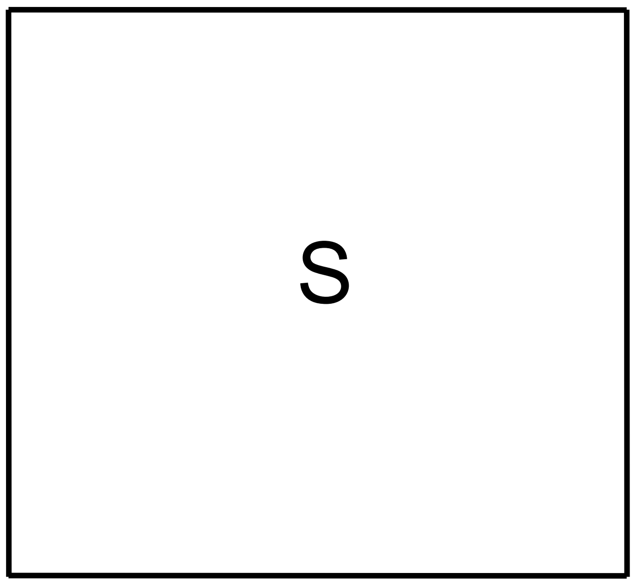
correlation polytope: $P_{corr} = conv\{aa^T | a \text{ in } \{0,1\}^n \}$

The Construction of [Fiorini et al]

correlation polytope: $P_{corr} = \text{conv}\{aa^T | a \text{ in } \{0,1\}^n \}$

vertices:

constraints:



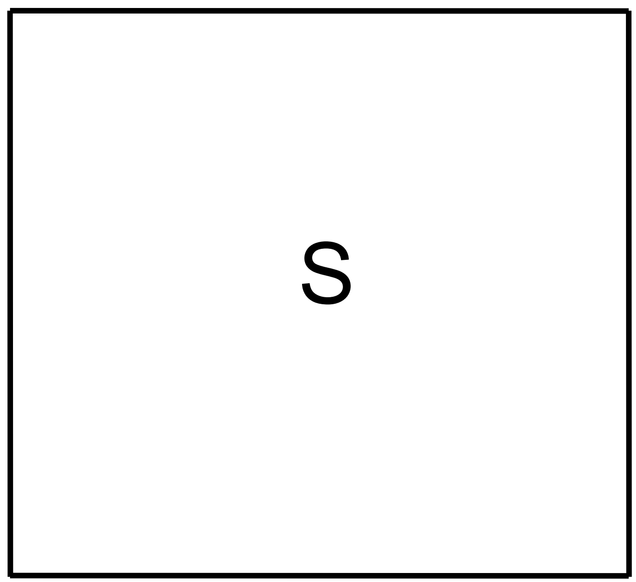
The Construction of [Fiorini et al]

correlation polytope: $P_{corr} = \text{conv}\{aa^T | a \text{ in } \{0,1\}^n\}$

vertices: a in $\{0,1\}^n$

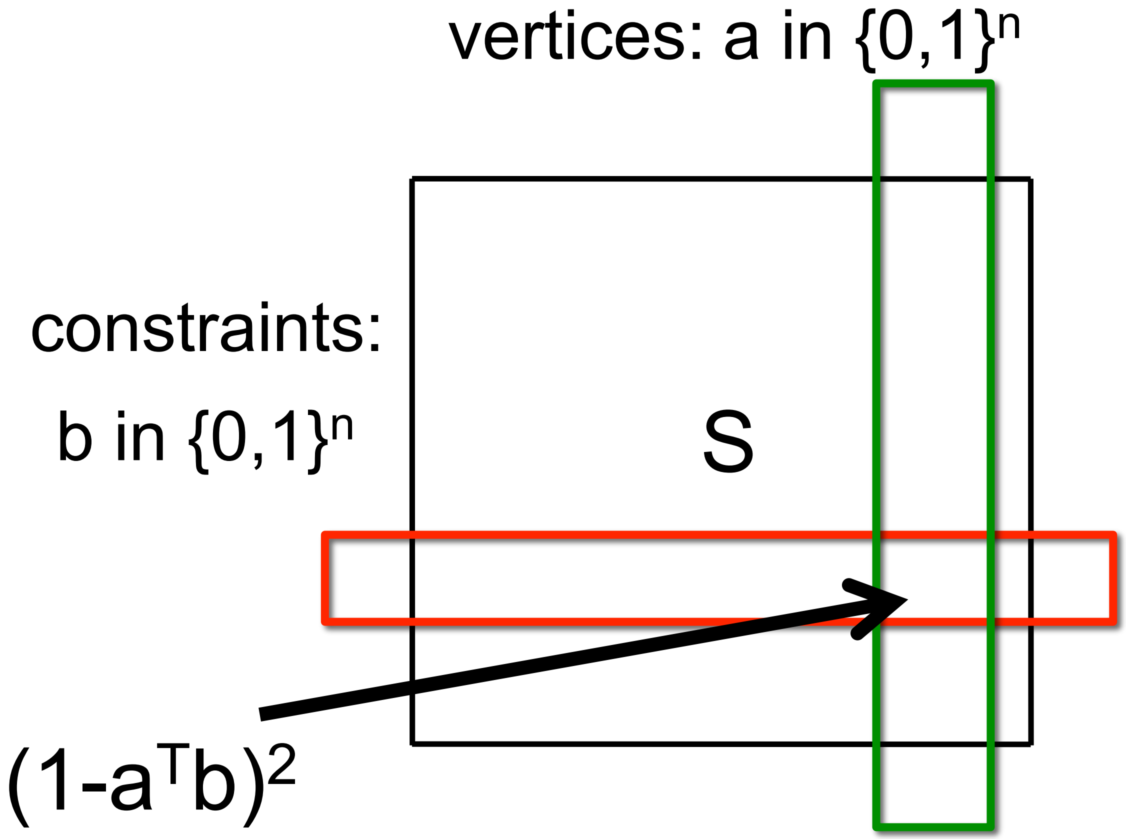
constraints:

b in $\{0,1\}^n$



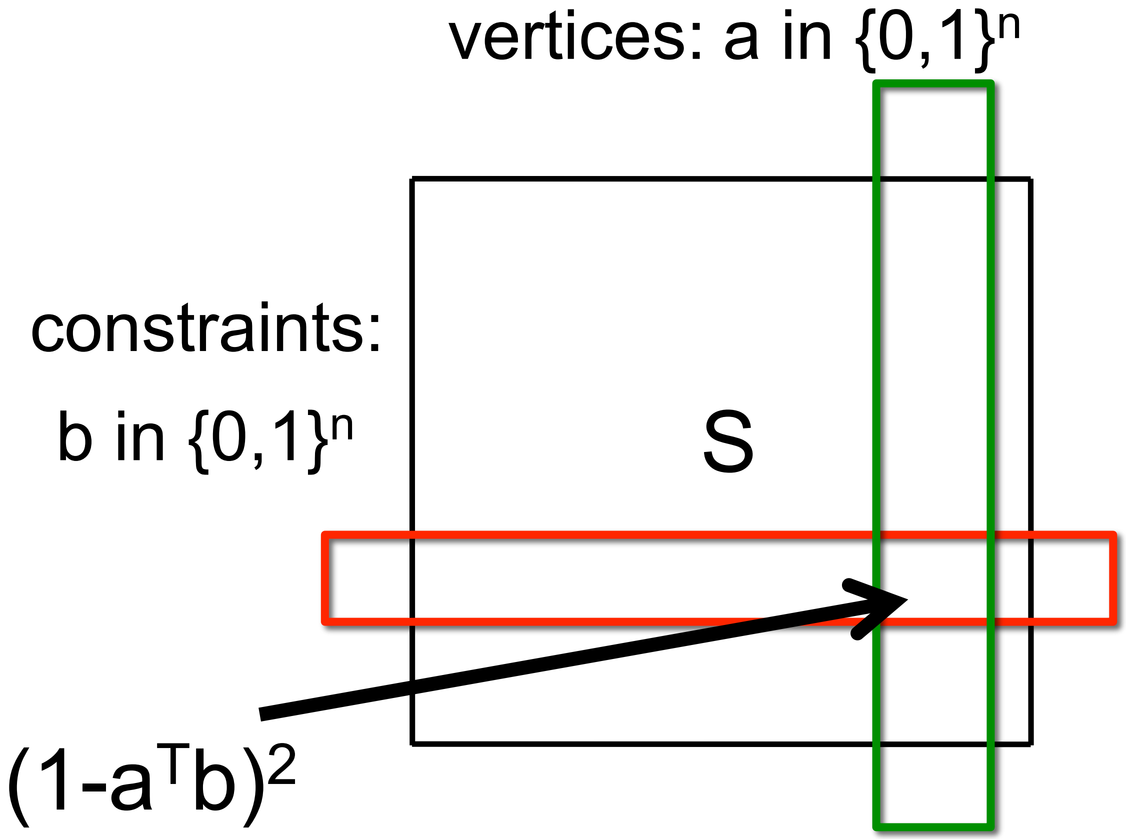
The Construction of [Fiorini et al]

correlation polytope: $P_{corr} = \text{conv}\{aa^T | a \text{ in } \{0,1\}^n\}$



The Construction of [Fiorini et al]

correlation polytope: $P_{corr} = \text{conv}\{aa^T | a \text{ in } \{0,1\}^n\}$



UNIQUE DISJ.
Output 'YES' if a and b as sets are disjoint, and 'NO' if a and b have one index in common

The Construction of [Fiorini et al]

correlation polytope: $P_{corr} = \text{conv}\{aa^T | a \text{ in } \{0,1\}^n\}$

The Construction of [Fiorini et al]

correlation polytope: $P_{corr} = \text{conv}\{aa^T | a \text{ in } \{0,1\}^n\}$

Why is that (a sub-matrix of) the slack matrix?

The Construction of [Fiorini et al]

correlation polytope: $P_{corr} = \text{conv}\{aa^T | a \text{ in } \{0,1\}^n\}$

Why is that (a sub-matrix of) the slack matrix?

$$(1-a^Tb)^2 = 1 - 2a^Tb + (a^Tb)^2$$

The Construction of [Fiorini et al]

correlation polytope: $P_{corr} = \text{conv}\{aa^T | a \text{ in } \{0,1\}^n\}$

Why is that (a sub-matrix of) the slack matrix?

$$(1-a^Tb)^2 = 1 - 2a^Tb + (a^Tb)^2$$

$$= 1 - 2\langle \text{diag}(b), aa^T \rangle + \langle bb^T, aa^T \rangle$$

The Construction of [Fiorini et al]

correlation polytope: $P_{corr} = \text{conv}\{aa^T | a \text{ in } \{0,1\}^n\}$

Why is that (a sub-matrix of) the slack matrix?

$$(1-a^Tb)^2 = 1 - 2a^Tb + (a^Tb)^2$$

$$= 1 - 2\langle \text{diag}(b), aa^T \rangle + \langle bb^T, aa^T \rangle$$

$$\Rightarrow \quad 1 \geq \langle 2\text{diag}(b) - bb^T, aa^T \rangle$$

The Construction of [Fiorini et al]

correlation polytope: $P_{corr} = \text{conv}\{aa^T | a \text{ in } \{0,1\}^n\}$

Why is that (a sub-matrix of) the slack matrix?

$$(1-a^Tb)^2 = 1 - 2a^Tb + (a^Tb)^2$$

$$= 1 - 2\langle \text{diag}(b), aa^T \rangle + \langle bb^T, aa^T \rangle$$

$$\Rightarrow \quad 1 \geq \langle 2\text{diag}(b) - bb^T, aa^T \rangle$$

What is the slack?

The Construction of [Fiorini et al]

correlation polytope: $P_{corr} = \text{conv}\{aa^T | a \text{ in } \{0,1\}^n\}$

Why is that (a sub-matrix of) the slack matrix?

$$(1-a^Tb)^2 = 1 - 2a^Tb + (a^Tb)^2$$

$$= 1 - 2\langle \text{diag}(b), aa^T \rangle + \langle bb^T, aa^T \rangle$$

$$\Rightarrow \quad 1 \geq \langle 2\text{diag}(b) - bb^T, aa^T \rangle$$

What is the slack? $(1-a^Tb)^2$

A Hard Distribution

A Hard Distribution

Let $T = \{(a,b) \mid a^T b = 0\}$, $|T| = 3^n$

A Hard Distribution

Let $T = \{(a,b) \mid a^T b = 0\}$, $|T| = 3^n$

Recall: $S_{a,b}=(1-a^T b)^2$, so $S_{a,b}=1$ for all pairs in T

A Hard Distribution

Let $T = \{(a,b) \mid a^T b = 0\}$, $|T| = 3^n$

Recall: $S_{a,b}=(1-a^T b)^2$, so $S_{a,b}=1$ for all pairs in T

How does the sampling procedure **specialize** to this case? (Recall it generates (a,b) unif. from T)

A Hard Distribution

Let $T = \{(a,b) \mid a^T b = 0\}$, $|T| = 3^n$

Recall: $S_{a,b}=(1-a^T b)^2$, so $S_{a,b}=1$ for all pairs in T

How does the sampling procedure **specialize** to this case? (Recall it generates (a,b) unif. from T)

Sampling Procedure:

A Hard Distribution

Let $T = \{(a,b) \mid a^T b = 0\}$, $|T| = 3^n$

Recall: $S_{a,b}=(1-a^T b)^2$, so $S_{a,b}=1$ for all pairs in T

How does the sampling procedure **specialize** to this case? (Recall it generates (a,b) unif. from T)

Sampling Procedure:

- Let R_i be the sum of $M_i(a,b)$ over (a,b) in T and let R be the sum of R_i

A Hard Distribution

Let $T = \{(a,b) \mid a^T b = 0\}$, $|T| = 3^n$

Recall: $S_{a,b}=(1-a^T b)^2$, so $S_{a,b}=1$ for all pairs in T

How does the sampling procedure **specialize** to this case? (Recall it generates (a,b) unif. from T)

Sampling Procedure:

- Let R_i be the sum of $M_i(a,b)$ over (a,b) in T and let R be the sum of R_i
- Choose i with probability R_i/R

A Hard Distribution

Let $T = \{(a,b) \mid a^T b = 0\}$, $|T| = 3^n$

Recall: $S_{a,b}=(1-a^T b)^2$, so $S_{a,b}=1$ for all pairs in T

How does the sampling procedure **specialize** to this case? (Recall it generates (a,b) unif. from T)

Sampling Procedure:

- Let R_i be the sum of $M_i(a,b)$ over (a,b) in T and let R be the sum of R_i
- Choose i with probability R_i/R
- Choose (a,b) with probability $M_i(a,b)/R_i$

Entropy Accounting 101

Entropy Accounting 101

Sampling Procedure:

- Let R_i be the sum of $M_i(a,b)$ over (a,b) in T and let R be the sum of R_i
- Choose i with probability R_i/R
- Choose (a,b) with probability $M_i(a,b)/R_i$

Entropy Accounting 101

Sampling Procedure:

- Let R_i be the sum of $M_i(a,b)$ over (a,b) in T and let R be the sum of R_i
- Choose i with probability R_i/R
- Choose (a,b) with probability $M_i(a,b)/R_i$

Total Entropy:

$n \log_2 3 \leq$

Entropy Accounting 101

Sampling Procedure:

- Let R_i be the sum of $M_i(a,b)$ over (a,b) in T and let R be the sum of R_i
- Choose i with probability R_i/R
- Choose (a,b) with probability $M_i(a,b)/R_i$

Total Entropy:

$$n \log_2 3 \leq \underbrace{\quad\quad}_{\text{choose } i} + \underbrace{\quad\quad\quad\quad}_{\text{choose } (a,b) \text{ conditioned on } i}$$

Entropy Accounting 101

Sampling Procedure:

- Let R_i be the sum of $M_i(a,b)$ over (a,b) in T and let R be the sum of R_i
- Choose i with probability R_i/R
- Choose (a,b) with probability $M_i(a,b)/R_i$

Total Entropy:

$$n \log_2 3 \leq \underbrace{\log_2 r}_{\text{choose } i} + \underbrace{\qquad\qquad}_{\text{choose (a,b) conditioned on } i}$$

Entropy Accounting 101

Sampling Procedure:

- Let R_i be the sum of $M_i(a,b)$ over (a,b) in T and let R be the sum of R_i
- Choose i with probability R_i/R
- Choose (a,b) with probability $M_i(a,b)/R_i$

Total Entropy:

$$n \log_2 3 \leq \underbrace{\log_2 r}_{\text{choose } i} + \underbrace{(1-\delta) n \log_2 3}_{\text{choose } (a,b) \text{ conditioned on } i}$$

Entropy Accounting 101

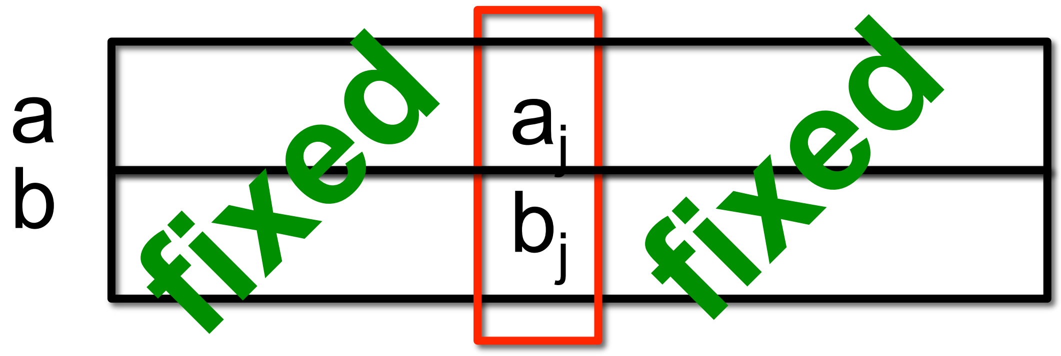
Sampling Procedure:

- Let R_i be the sum of $M_i(a,b)$ over (a,b) in T and let R be the sum of R_i
- Choose i with probability R_i/R
- Choose (a,b) with probability $M_i(a,b)/R_i$

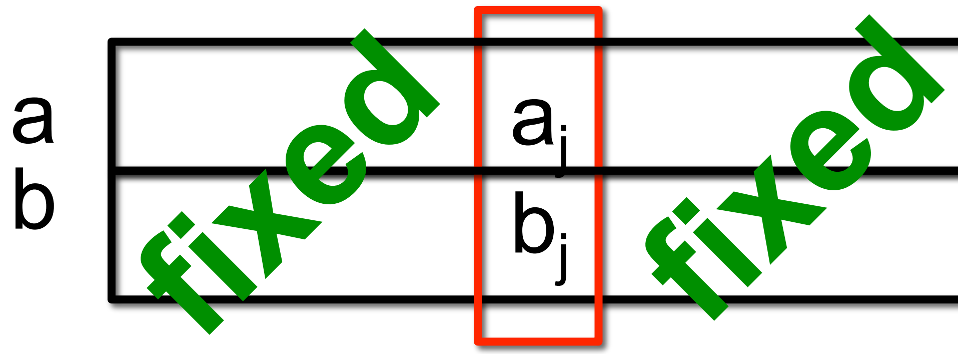
Total Entropy:

$$n \log_2 3 \le \underbrace{\log_2 r}_{\text{choose } i} + \underbrace{(1-\delta) n \log_2 3}_{\text{choose (a,b) conditioned on } i} \quad (?)$$

Suppose that a_{-j} and b_{-j} are **fixed**

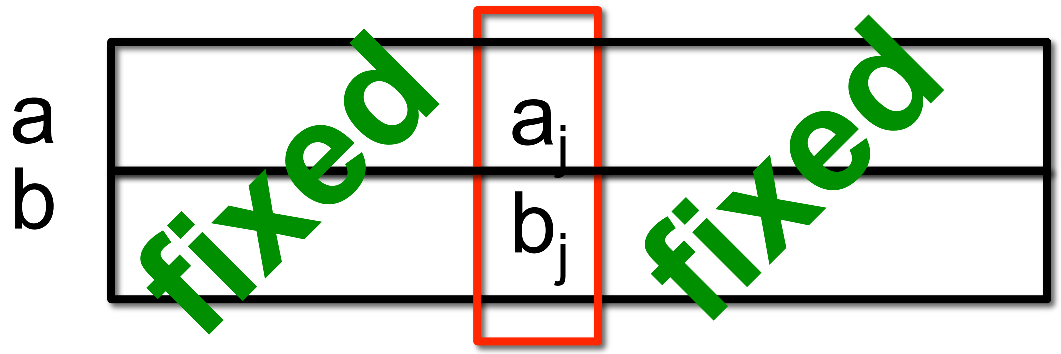


Suppose that a_{-j} and b_{-j} are **fixed**



M_i restricted to (a_{-j}, b_{-j})

Suppose that a_{-j} and b_{-j} are **fixed**



M_i restricted to (a_{-j},b_{-j})

	$(\ldots b_j=0\ldots)$	$(\ldots b_j=1\ldots)$
$(a_{1..j-1},a_j=0,a_{j+1\ldots n})$	$M_i(a,b)$	$M_i(a,b)$
$(a_{1..j-1},a_j=1,a_{j+1\ldots n})$	$M_i(a,b)$	$M_i(a,b)$

	$(\ldots b_j=0\ldots)$	$(\ldots b_j=1\ldots)$
$(a_{1..j-1},a_j=0,a_{j+1\ldots n})$	$M_i(a,b)$	$M_i(a,b)$
$(a_{1..j-1},a_j=1,a_{j+1\ldots n})$	$M_i(a,b)$	$M_i(a,b)$

If $a_j=1$, $b_j=1$ then $a^T b = 1$, hence $M_i(a,b) = 0$

	$(\ldots b_j=0\ldots)$	$(\ldots b_j=1\ldots)$
$(a_{1..j-1},a_j=0,a_{j+1\ldots n})$	$M_i(a,b)$	$M_i(a,b)$
$(a_{1..j-1},a_j=1,a_{j+1\ldots n})$	$M_i(a,b)$	$M_i(a,b)$

If $a_j=1$, $b_j=1$ then $a^Tb = 1$, hence $M_i(a,b) = 0$

	$(\ldots b_j=0\ldots)$	$(\ldots b_j=1\ldots)$
$(a_{1..j-1},a_j=0,a_{j+1\ldots n})$	$M_i(a,b)$	$M_i(a,b)$
$(a_{1..j-1},a_j=1,a_{j+1\ldots n})$	$M_i(a,b)$	**zero**

If $a_j=1$, $b_j=1$ then $a^Tb = 1$, hence $M_i(a,b) = 0$

But $rank(M_i)=1$, hence there must be another zero in either the same row or column

	$(\ldots b_j=0\ldots)$	$(\ldots b_j=1\ldots)$
$(a_{1..j-1}, a_j=0, a_{j+1\ldots n})$	$M_i(a,b)$	$M_i(a,b)$
$(a_{1..j-1}, a_j=1, a_{j+1\ldots n})$	$M_i(a,b)$	**zero**

If $a_j=1$, $b_j=1$ then $a^Tb = 1$, hence $M_i(a,b) = 0$

But $rank(M_i)=1$, hence there must be another zero in either the same row or column

	$(\ldots b_j=0\ldots)$	$(\ldots b_j=1\ldots)$
$(a_{1..j-1},a_j=0,a_{j+1\ldots n})$	$M_i(a,b)$	$M_i(a,b)$
$(a_{1..j-1},a_j=1,a_{j+1\ldots n})$	**zero**	**zero**

If $a_j=1$, $b_j=1$ then $a^Tb = 1$, hence $M_i(a,b) = 0$

But $rank(M_i)=1$, hence there must be another zero in either the same row or column

$H(a_j,b_j \mid i, a_{-j}, b_{-j}) \leq 1 < \log_2 3$

	$(\ldots b_j=0\ldots)$	$(\ldots b_j=1\ldots)$
$(a_{1..j-1},a_j=0,a_{j+1\ldots n})$	$M_i(a,b)$	$M_i(a,b)$
$(a_{1..j-1},a_j=1,a_{j+1\ldots n})$	**zero**	**zero**

Entropy Accounting 101

Generate uniformly random (a,b) in T:

- Let R_i be the sum of $M_i(a,b)$ over (a,b) in T and let R be the sum of R_i
- Choose i with probability R_i/R
- Choose (a,b) with probability $M_i(a,b)/R_i$

Total Entropy:

$$n \log_2 3 \leq \underbrace{\log_2 r}_{\text{choose } i} + \underbrace{}_{\text{choose (a,b) conditioned on } i}$$

Entropy Accounting 101

Generate uniformly random (a,b) in T:

- Let R_i be the sum of $M_i(a,b)$ over (a,b) in T and let R be the sum of R_i
- Choose i with probability R_i/R
- Choose (a,b) with probability $M_i(a,b)/R_i$

Total Entropy:

$$n \log_2 3 \leq \underbrace{\log_2 r}_{\text{choose } i} + \underbrace{n}_{\text{choose } (a,b) \text{ conditioned on } i}$$

Outline

Part I: Tools for Extended Formulations

- Yannakakis's Factorization Theorem
- The Rectangle Bound
- A Sampling Argument

Part II: Applications

- Correlation Polytope
- Approximating the Correlation Polytope
- Matching Polytope

Outline

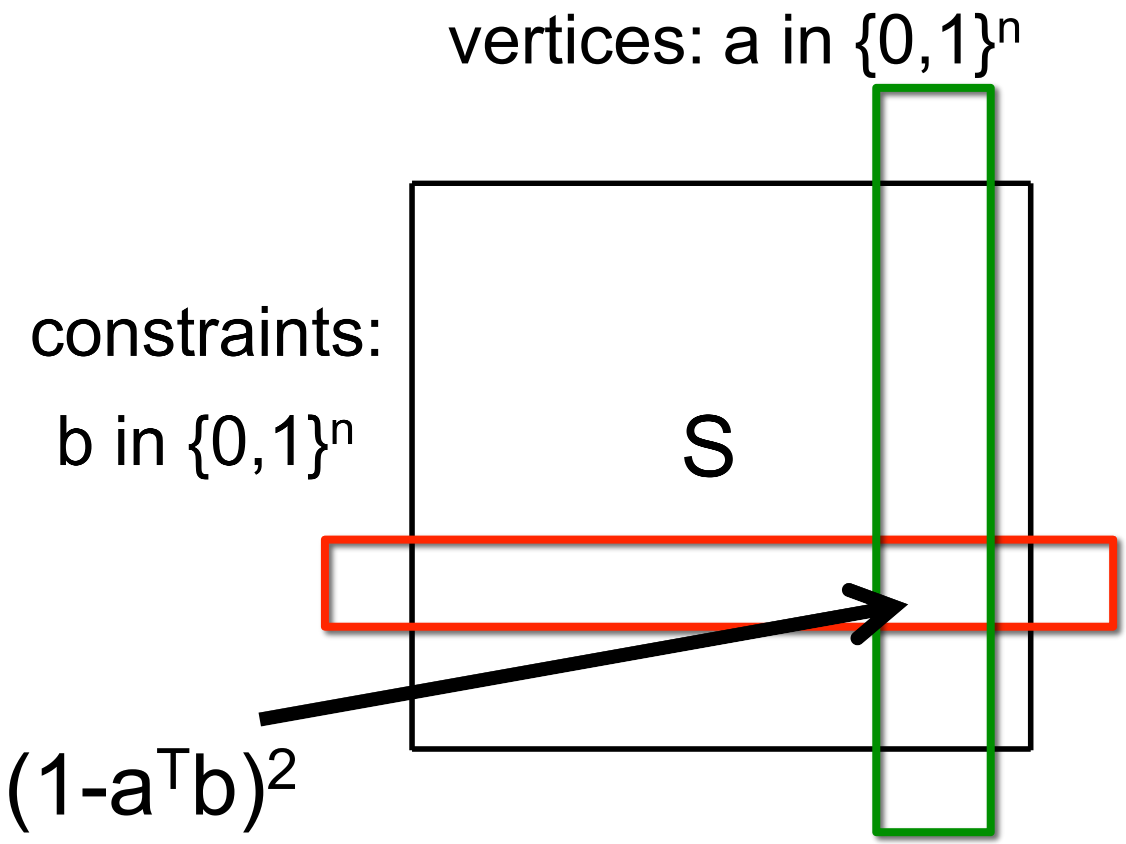
Part I: Tools for Extended Formulations

- Yannakakis's Factorization Theorem
- The Rectangle Bound
- A Sampling Argument

Part II: Applications

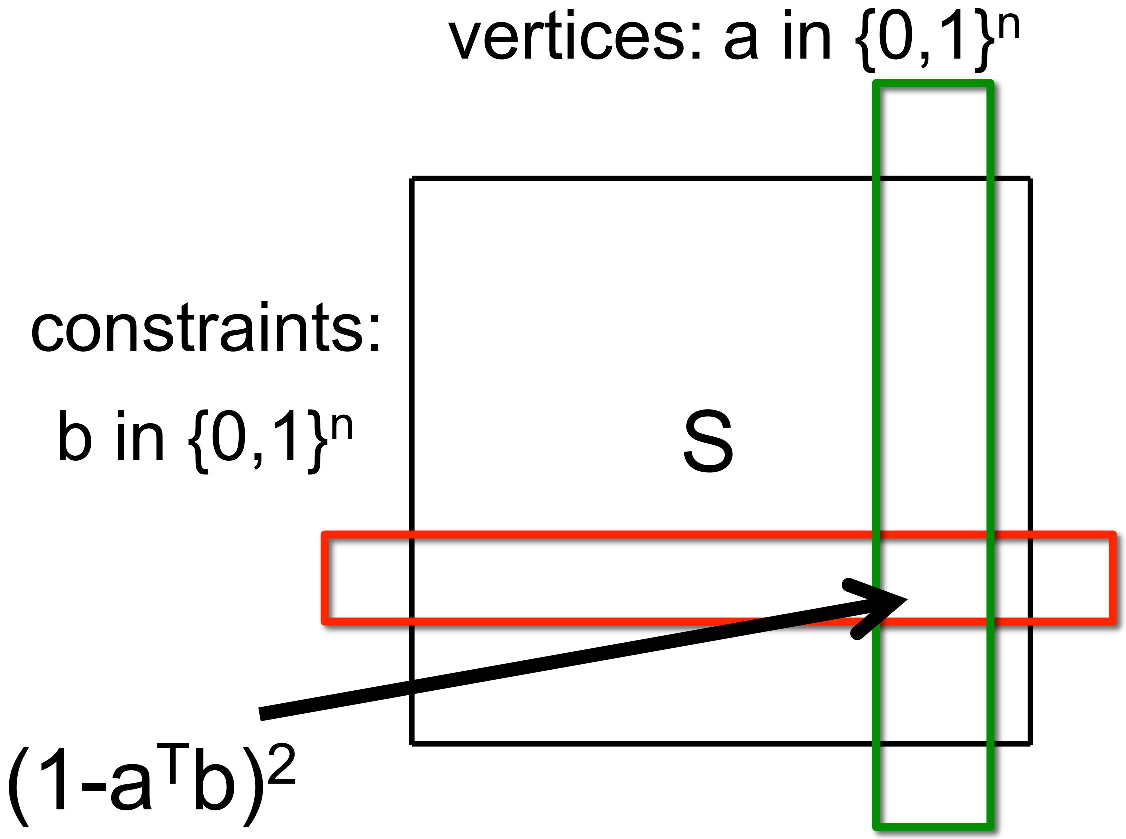
- Correlation Polytope
- **Approximating the Correlation Polytope**
- Matching Polytope

Approximate EFs [Braun et al]



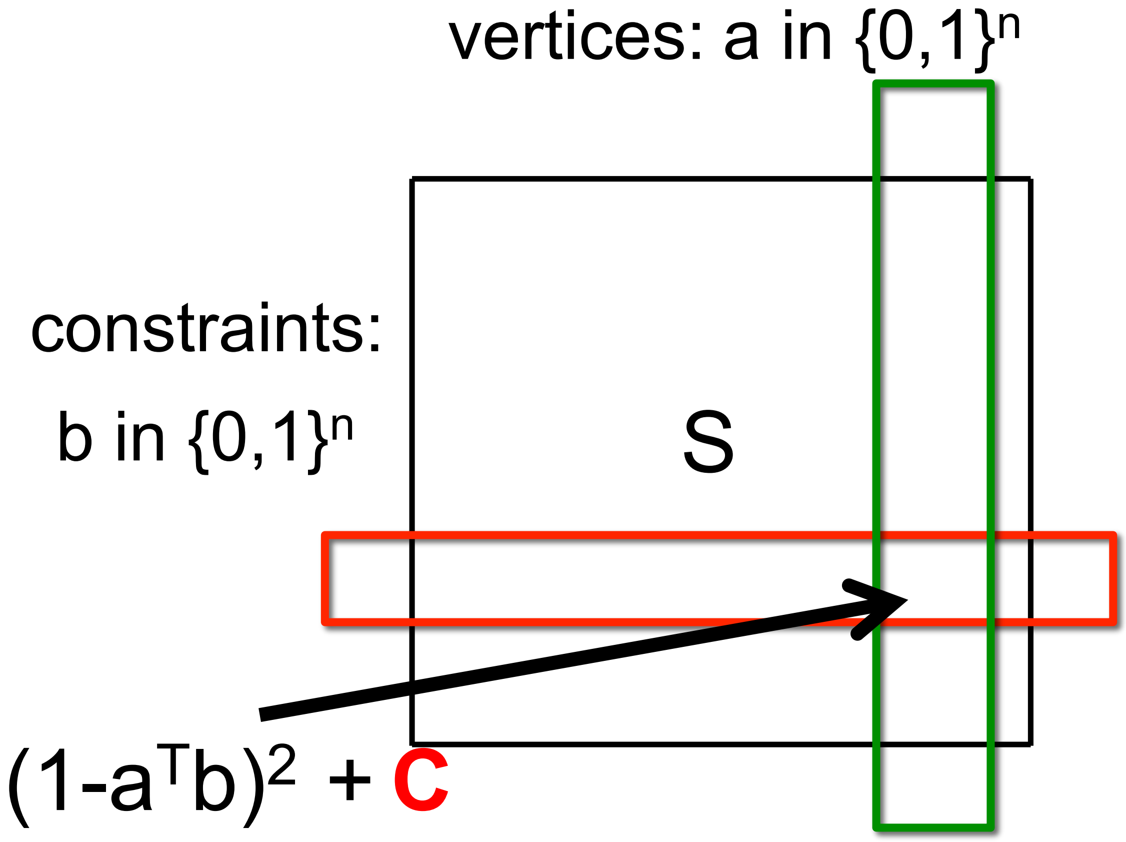
Approximate EFs [Braun et al]

Is there a K (with small xc) s.t. $P_{corr} \subset K \subset (C+1)P_{corr}$?



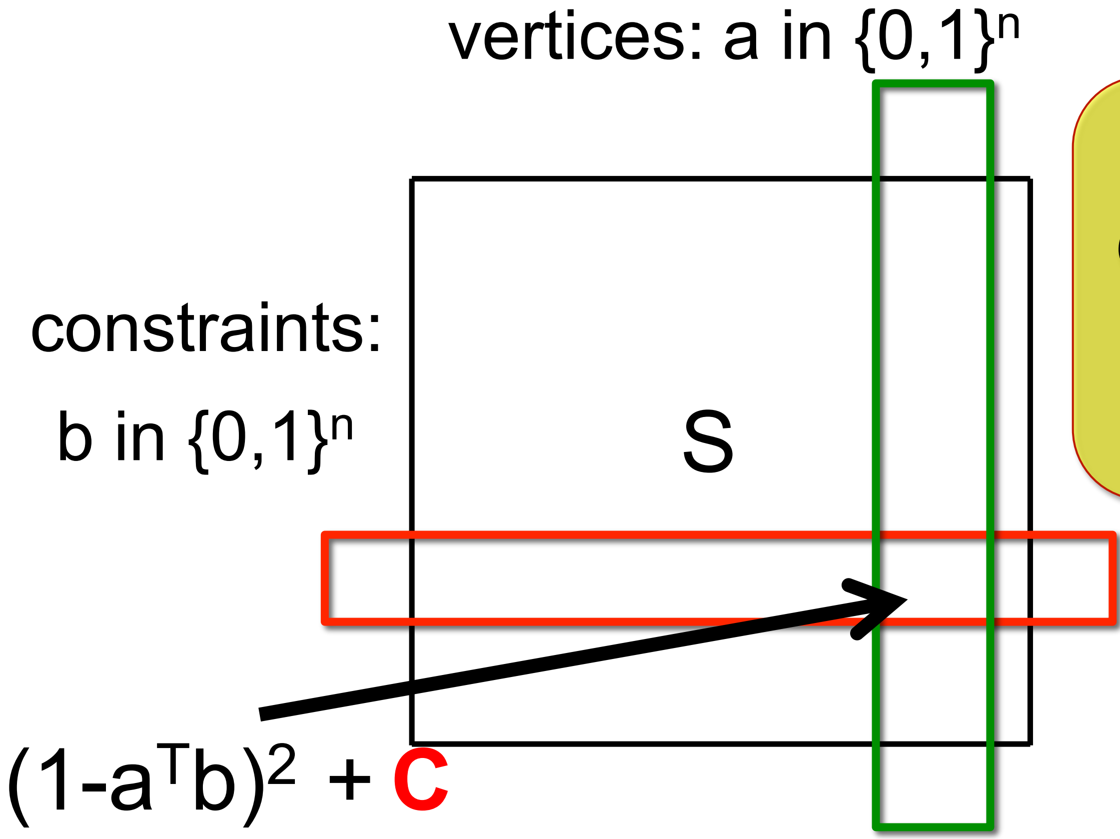
Approximate EFs [Braun et al]

Is there a K (with small xc) s.t. $P_{corr} \subset K \subset (C+1)P_{corr}$?



Approximate EFs [Braun et al]

Is there a K (with small xc) s.t. $P_{corr} \subset K \subset (C+1)P_{corr}$?



New Goal:
Output the answer to UDISJ with prob. at least ½ + 1/2(C+1)

Is the correlation polytope hard to approximate for large values of C?

Analogy: Is UDISJ hard to compute with prob. $\frac{1}{2}+1/2(C+1)$ for large values of C?

Is the correlation polytope hard to approximate for large values of C?

Analogy: Is UDISJ hard to compute with prob. $\frac{1}{2}+1/2(C+1)$ for large values of C?

There is a natural barrier at $C = \sqrt{n}$ for proving l.b.s:

Is the correlation polytope hard to approximate for large values of C?

Analogy: Is UDISJ hard to compute with prob. $\frac{1}{2}+1/2(C+1)$ for large values of C?

There is a natural barrier at $C = \sqrt{n}$ for proving l.b.s:

Claim: If UDISJ can be computed with prob. $\frac{1}{2}+1/2(C+1)$ using $o(n/C^2)$ bits, then UDISJ can be computed with prob. $\frac{3}{4}$ using $o(n)$ bits

Is the correlation polytope hard to approximate for large values of C?

Analogy: Is UDISJ hard to compute with prob. $\frac{1}{2}+1/2(C+1)$ for large values of C?

There is a natural barrier at $C = \sqrt{n}$ for proving l.b.s:

Claim: If UDISJ can be computed with prob. $\frac{1}{2}+1/2(C+1)$ using $o(n/C^2)$ bits, then UDISJ can be computed with prob. ¾ using $o(n)$ bits

Proof: Run the protocol $O(C^2)$ times and take the majority vote

Is the correlation polytope hard to approximate for large values of C?

Analogy: Is UDISJ hard to compute with prob. $\frac{1}{2}+1/2(C+1)$ for large values of C?

There is a natural barrier at $C = \sqrt{n}$ for proving l.b.s:

Is the correlation polytope hard to approximate for large values of C?

Analogy: Is UDISJ hard to compute with prob. $\frac{1}{2}+1/2(C+1)$ for large values of C?

There is a natural barrier at $C = \sqrt{n}$ for proving l.b.s:

Corollary [from K-S]: Computing UDISJ with probability $\frac{1}{2}+1/2(C+1)$ requires $\Omega(n/C^2)$ bits

Is the correlation polytope hard to approximate for large values of C?

Analogy: Is UDISJ hard to compute with prob. $\frac{1}{2}+1/2(C+1)$ for large values of C?

There is a natural barrier at $C = \sqrt{n}$ for proving l.b.s:

Corollary [from K-S]: Computing UDISJ with probability $\frac{1}{2}+1/2(C+1)$ requires $\Omega(n/C^2)$ bits

Theorem [B-M]: Computing UDISJ with probability $\frac{1}{2}+1/2(C+1)$ requires $\Omega(n/C)$ bits

Is the correlation polytope hard to approximate for large values of C?

Analogy: Is UDISJ hard to compute with prob. $\frac{1}{2}+1/2(C+1)$ for large values of C?

There is a natural barrier at $C = \sqrt{n}$ for proving l.b.s:

Theorem [B-M]: Any EF that approximates clique within n^{1-eps} has size $\exp(n^{eps})$

Theorem [B-M]: Computing UDISJ with probability $\frac{1}{2}+1/2(C+1)$ requires $\Omega(n/C)$ bits

Outline

Part I: Tools for Extended Formulations

- Yannakakis's Factorization Theorem
- The Rectangle Bound
- A Sampling Argument

Part II: Applications

- Correlation Polytope
- Approximating the Correlation Polytope
- Matching Polytope

Outline

Part I: Tools for Extended Formulations

- Yannakakis's Factorization Theorem
- The Rectangle Bound
- A Sampling Argument

Part II: Applications

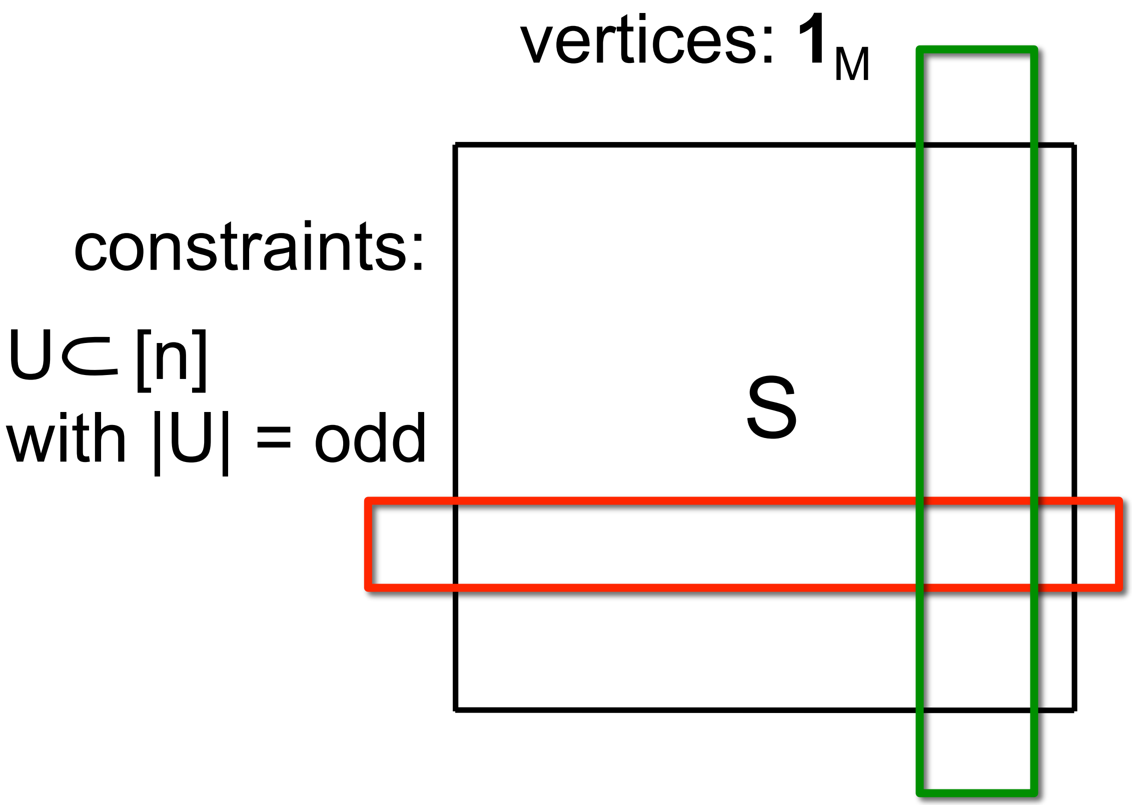
- Correlation Polytope
- Approximating the Correlation Polytope
- **Matching Polytope**

The Matching Polytope [Edmonds]

P_{PM}= conv{$\mathbf{1}_M$| M is a perfect matching in K_n}

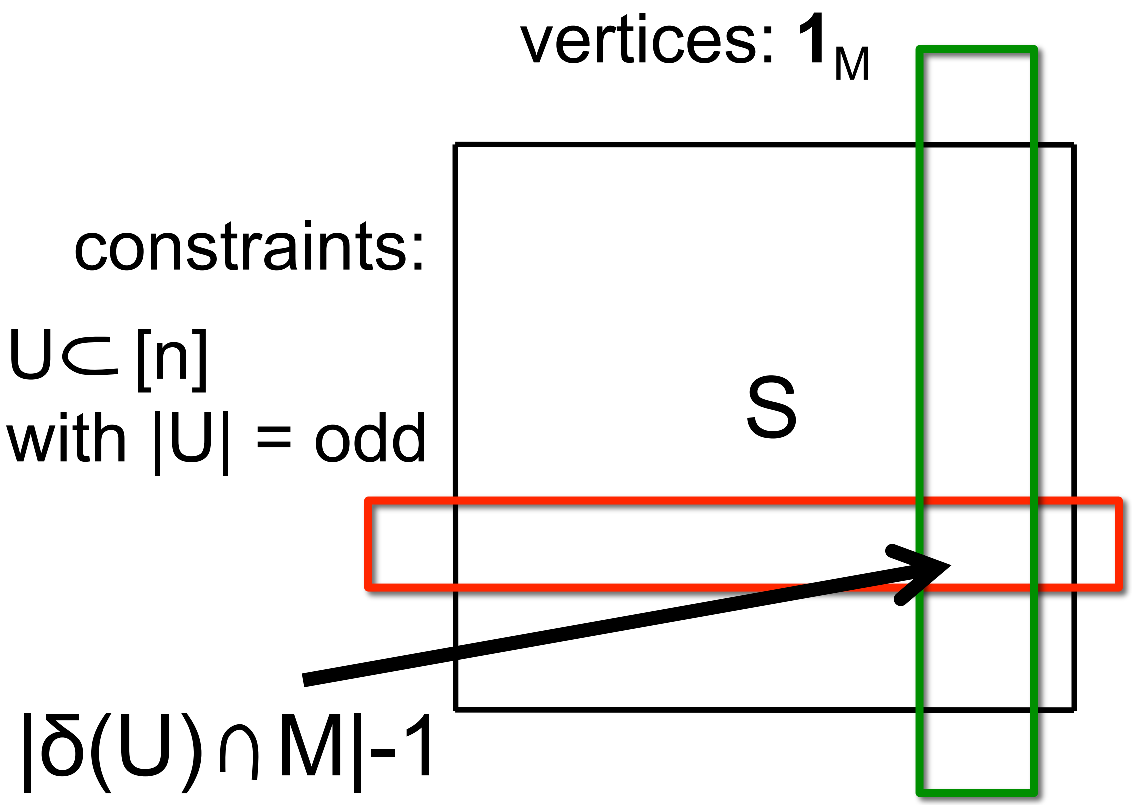
The Matching Polytope [Edmonds]

$P_{PM} = \text{conv}\{\mathbf{1}_M \mid M \text{ is a perfect matching in } K_n\}$



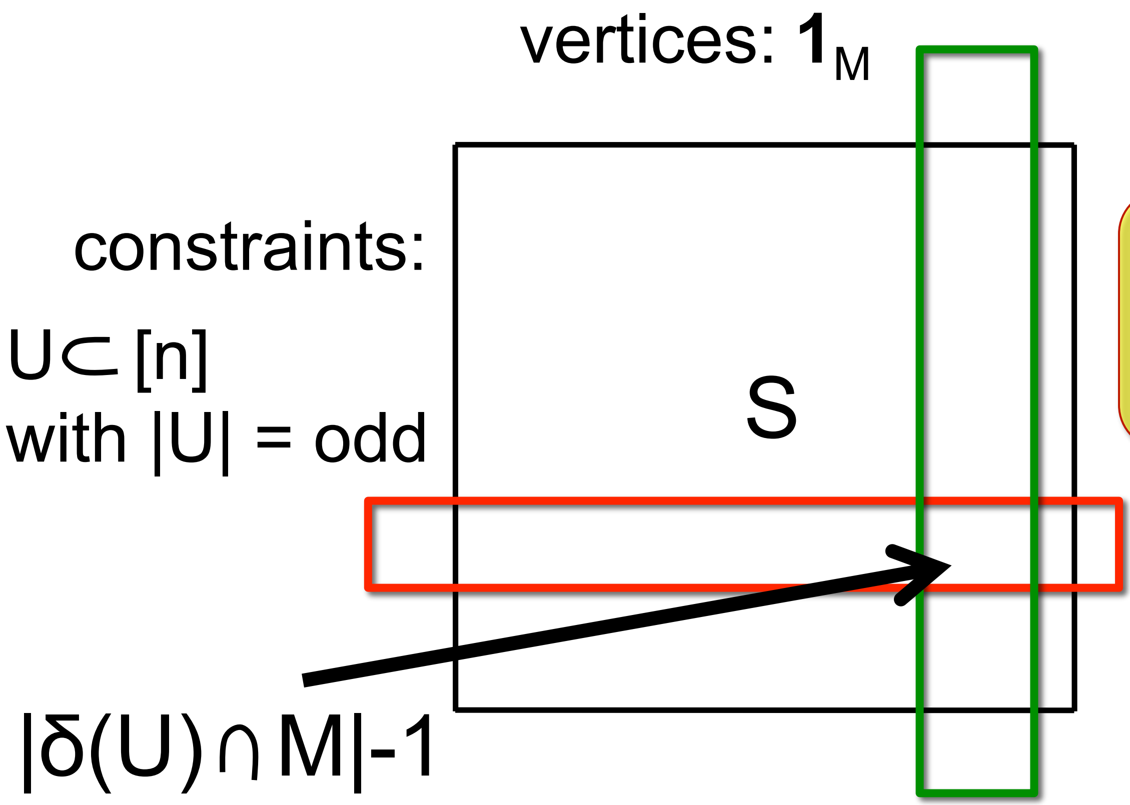
The Matching Polytope [Edmonds]

P_{PM}= conv{$\mathbf{1}_M$| M is a perfect matching in K_n}



The Matching Polytope [Edmonds]

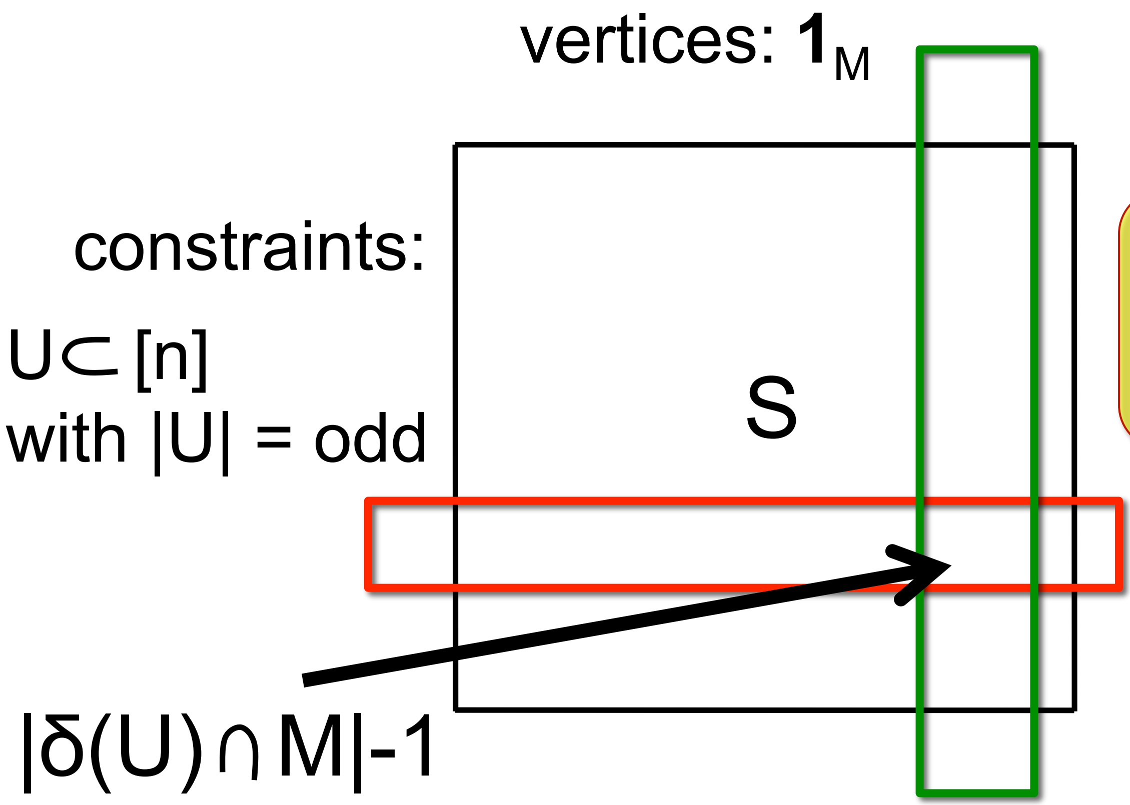
P_{PM}= conv{$\mathbf{1}_M$| M is a perfect matching in K_n}



Is there a small rectangle covering?

The Matching Polytope [Edmonds]

P_{PM}= conv{$\mathbf{1}_M$| M is a perfect matching in K_n}



Is there a small rectangle covering?

Yes! Just guess two edges in M, crossing the cut

Hyperplane Separation Lemma

[Rothvoss] attributed to **[Fiorini]:**

Hyperplane Separation Lemma

[Rothvoss] attributed to **[Fiorini]:**

Lemma: For slack matrix S, any matrix W:

$$\text{rank}^+(S) \geq \frac{\langle S, W \rangle}{\|S\|_\infty \, \alpha}$$

where $\alpha = \max \langle W, R \rangle$ s.t. R is rank one, entries in [0,1]

Hyperplane Separation Lemma

[Rothvoss] attributed to **[Fiorini]:**

Lemma: For slack matrix S, any matrix W:

$$\mathrm{rank}^+(S) \geq \frac{\langle S, W \rangle}{\|S\|_\infty \, \alpha}$$

where $\alpha = \max \langle W, R \rangle$ s.t. R is rank one, entries in [0,1]

Proof:

$$\langle W, S \rangle = \Sigma \, \|R_i\|_\infty \langle W, R_i / \|R_i\|_\infty \rangle \leq \alpha \, r \, \|S\|_\infty$$

Theorem [Rothvoss '13]: Any EF for perfect matching has size $2^{\Omega(n)}$ (same for TSP)

Theorem [Rothvoss '13]: Any EF for perfect matching has size $2^{\Omega(n)}$ (same for TSP)

How do we choose W?

Theorem [Rothvoss '13]: Any EF for perfect matching has size $2^{\Omega(n)}$ (same for TSP)

How do we choose W?

$$W_{U,M} = \begin{cases} -\infty & \text{if } |\delta(U) \cap M| = 1 \\ 1/Q_3 & \text{if } |\delta(U) \cap M| = 3 \\ -1/Q_k & \text{if } |\delta(U) \cap M| = k \\ 0 & \text{else} \end{cases}$$

Theorem [Rothvoss '13]: Any EF for perfect matching has size $2^{\Omega(n)}$ (same for TSP)

How do we choose W?

$$W_{U,M} = \begin{cases} -\infty & \text{if } |\delta(U) \cap M| = 1 \\ 1/Q_3 & \text{if } |\delta(U) \cap M| = 3 \\ -1/Q_k & \text{if } |\delta(U) \cap M| = k \\ 0 & \text{else} \end{cases}$$

Proof is a substantial modification to Razborov's rectangle corruption lemma

Summary:

- Extended formulations and Yannakakis' factorization theorem

Summary:

- Extended formulations and Yannakakis' factorization theorem
- **Lower bound techniques:** rectangle bound, information complexity, hyperplane separation

Summary:

- Extended formulations and Yannakakis' factorization theorem
- **Lower bound techniques:** rectangle bound, information complexity, hyperplane separation
- **Applications:** connections between correlation polytope and disjointness,

Summary:

- Extended formulations and Yannakakis' factorization theorem
- **Lower bound techniques:** rectangle bound, information complexity, hyperplane separation
- **Applications:** connections between correlation polytope and disjointness,
- **Open question:** Can we prove lower bounds against general SDPs?

Any Questions?

Summary:

- Extended formulations and Yannakakis' factorization theorem
- **Lower bound techniques:** rectangle bound, information complexity, hyperplane separation
- **Applications:** connections between correlation polytope and disjointness,
- **Open question:** Can we prove lower bounds against general SDPs?

Thanks!