Extended formulations, non-negative factorizations and randomized communication protocols

Michele Conforti[1] Yuri Faenza[1] Samuel Fiorini[2]
Roland Grappe[1] Hans Raj Tiwary[2]

[1]Università di Padova

[2]Université Libre de Bruxelles

What is this talk about?

In 91, Mihalis Yannakakis published an important paper.

What is this talk about?

In 91, Mihalis Yannakakis published an important paper.

Theorem (Y91)

Every symmetric extended formulation for the perfect matching polytope has exponential size.

What is this talk about?

In 91, Mihalis Yannakakis published an important paper.

Theorem (Y91)

Every symmetric extended formulation for the perfect matching polytope has exponential size.

What we do:

- Prove a result of the type: *every* $\langle$*other condition*$\rangle$ *extended formulation for the perfect matching polytope has* $\langle$*certain size*$\rangle$

What is this talk about?

In 91, Mihalis Yannakakis published an important paper.

Theorem (Y91)

Every symmetric extended formulation for the perfect matching polytope has exponential size.

What we do:

- Prove a result of the type: *every ⟨other condition⟩ extended formulation for the perfect matching polytope has ⟨certain size⟩*
- Develop a new way of *interpreting* extended formulations

Extension complexity

Getting started

$P \subseteq \mathbb{R}^d$ polytope

$Q \subseteq \mathbb{R}^e$ polyhedron &
$\pi : \mathbb{R}^e \to \mathbb{R}^d$ linear map | define *extension* of P if $\boxed{\pi(Q) = P}$

Extension complexity

Getting started

$P \subseteq \mathbb{R}^d$ polytope

$Q \subseteq \mathbb{R}^e$ polyhedron &
$\pi : \mathbb{R}^e \to \mathbb{R}^d$ linear map | define *extension* of P if $\boxed{\pi(Q) = P}$

Definition

$\mathrm{xc}(P) :=$ *extension complexity* of P

$:=$ min #facets in an extension of P

Extension complexity

Getting started

$P \subseteq \mathbb{R}^d$ polytope

$Q \subseteq \mathbb{R}^e$ polyhedron & $\pi : \mathbb{R}^e \to \mathbb{R}^d$ linear map	define *extension* of P	if	$\boxed{\pi(Q) = P}$

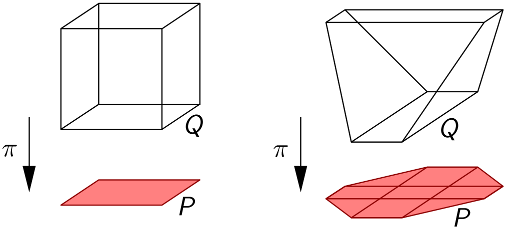
Definition

$\mathrm{xc}(P) :=$ *extension complexity* of P

$:= \min \#$facets in an extension of P

Example:

$\mathrm{xc}(\text{regular 8-gon}) \leqslant 6$



Extension complexity

More examples, and things for your to do list

Known results:

- xc(regular n-gon) $= \Theta(\log n)$ [BN01]
- xc(n-permutohedron) $= \Theta(n \log n)$ [G10]
- xc(spanning tree polytope of K_n) $= O(n^3)$ [M87]
- xc(stable set polytope of perfect graph G) $= n^{O(\log n)}$ [Y91]

Extension complexity

More examples, and things for your to do list

Known results:

- xc(regular n-gon) $= \Theta(\log n)$ [BN01]
- xc(n-permutohedron) $= \Theta(n \log n)$ [G10]
- xc(spanning tree polytope of K_n) $= O(n^3)$ [M87]
- xc(stable set polytope of perfect graph G) $= n^{O(\log n)}$ [Y91]

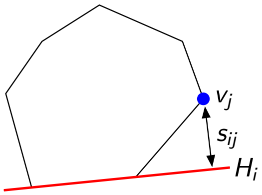
Open questions:

- xc(nonregular n-gon) $=$?
- xc(0/1 polytope P in $\mathbb{R}^d$) at most polynomial in d? [K]
- xc(perfect matching polytope of K_n) $=$?
- ...

Slack matrices

$$\begin{aligned} P &= \operatorname{conv}\{v_1, \dots, v_n\} \\ &= \{x \in \mathbb{R}^d : a_1^T x - b_1 \geqslant 0, \dots, a_m^T x - b_m \geqslant 0\} \end{aligned}$$

non-redundant inner/outer descriptions of P (assuming P full-dim.)



Definition

$S(P) :=$ *slack matrix* of P

$:= m \times n$ matrix (s_{ij}) with $s_{ij} = a_i^T v_j - b_i \geqslant 0$

$=$ slack of jth vertex w.r.t. ith facet

Non-negative rank

$S \in \mathbb{R}_{+}^{m \times n}$

Non-negative rank

$S \in \mathbb{R}_{+}^{m \times n}$

Definition

$$\begin{aligned} \mathrm{rank}_{+}(S) &:= \textit{non-negative rank of } S \\ &:= \min r \text{ s.t. } S = AB \text{ with } A \in \mathbb{R}_{+}^{m \times r} \text{ and } B \in \mathbb{R}_{+}^{r \times n} \\ &= \min r \text{ s.t. } S \text{ is sum of } r \text{ non-neg. rank 1 matrices} \end{aligned}$$

Non-negative rank... of a slack matrix

$S \in \mathbb{R}_{+}^{m \times n}$

Definition

$\text{rank}_{+}(S) :=$ *non-negative rank* of S

$:= \min r$ s.t. $S = AB$ with $A \in \mathbb{R}_{+}^{m \times r}$ and $B \in \mathbb{R}_{+}^{r \times n}$

$= \min r$ s.t. S is sum of r non-neg. rank 1 matrices

Theorem (Y91, FKPT10+)

$$xc(P) = rank_{+}(S(P))$$

Non-negative rank... of a slack matrix

$S \in \mathbb{R}_+^{m \times n}$

Definition

$\text{rank}_+(S) :=$ *non-negative rank* of S

$:= \min r$ s.t. $S = AB$ with $A \in \mathbb{R}_+^{m \times r}$ and $B \in \mathbb{R}_+^{r \times n}$

$= \min r$ s.t. S is sum of r non-neg. rank 1 matrices

Theorem (Y91, FKPT10+)

$$xc(P) = rank_+(S(P))$$

Why like this result?

- elegant mathematical fact
- connects several research fields together
  - polyhedral combinatorics
  - applied linear algebra
  - communication complexity

Deterministic communication protocols

The tale of Alice and Bob

$f : X \times Y \to \{0, 1\}$ boolean function (matrix)

Two players:

- Alice knows $x \in X$
- Bob knows $y \in Y$

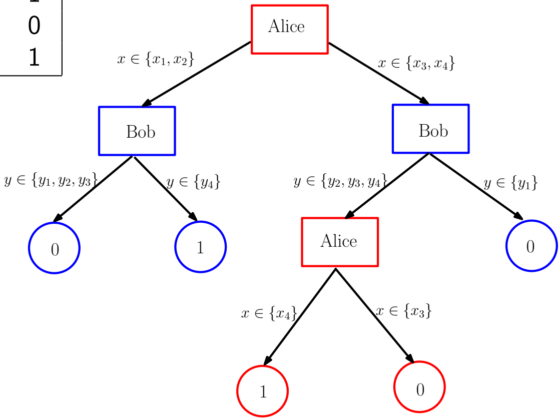
want to *compute* $f(x, y)$ *by exchanging bits*

Goal: Minimize *complexity* $:=$ #bits exchanged

Deterministic communication protocols

An example of a protocol

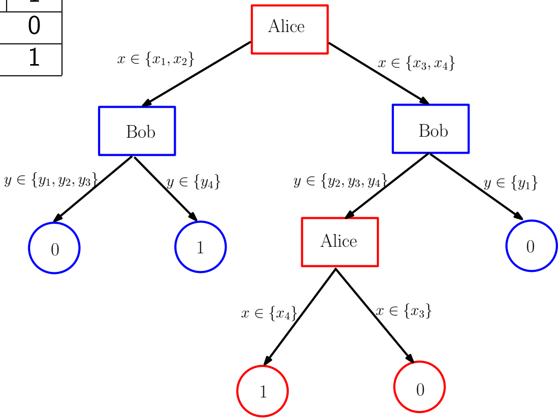
	y_1	y_2	y_3	y_4
x_1	0	0	0	1
x_2	0	0	0	1
x_3	0	0	0	0
x_4	0	1	1	1



Deterministic communication protocols

An example of a protocol

	y_1	y_2	y_3	y_4
x_1	0	0	0	1
x_2	0	0	0	1
x_3	0	0	0	0
x_4	0	1	1	1



Deterministic communication protocols

Cases where Alice and Bob provide an extended formulation

Assume: $P = \mathrm{STAB}(G)$ for G perfect graph

Deterministic communication protocols

Cases where Alice and Bob provide an extended formulation

Assume: $P = \text{STAB}(G)$ for G perfect graph $\implies S(P)$ is binary

Deterministic communication protocols

Cases where Alice and Bob provide an extended formulation

Assume: $P = \text{STAB}(G)$ for G perfect graph $\implies S(P)$ is binary

Slack matrix: rows correspond to cliques K
columns correspond to stable sets S
slack of S w.r.t K is $1 - \chi^S(K) = 1 - |S \cap K|$

Deterministic communication protocols

Cases where Alice and Bob provide an extended formulation

Assume: $P = \text{STAB}(G)$ for G perfect graph $\implies S(P)$ is binary

Slack matrix: rows correspond to cliques K
columns correspond to stable sets S
slack of S w.r.t K is $1 - \chi^S(K) = 1 - |S \cap K|$

$$\boxed{\exists \text{ complexity } c \text{ protocol for computing } S(P) \implies \text{xc}(P) \leqslant 2^c}$$

Deterministic communication protocols

Cases where Alice and Bob provide an extended formulation

Assume: $P = \text{STAB}(G)$ for G perfect graph $\implies$ $S(P)$ is binary

Slack matrix: rows correspond to cliques K
columns correspond to stable sets S
slack of S w.r.t K is $1 - \chi^S(K) = 1 - |S \cap K|$

$$\boxed{\exists \text{ complexity } c \text{ protocol for computing } S(P) \implies \text{xc}(P) \leqslant 2^c}$$

Theorem (Y91)

For all n-vertex perfect graphs G,

$$xc(STAB(G)) \leqslant 2^{O(\log^2 n)} = n^{O(\log n)}$$

Proof.

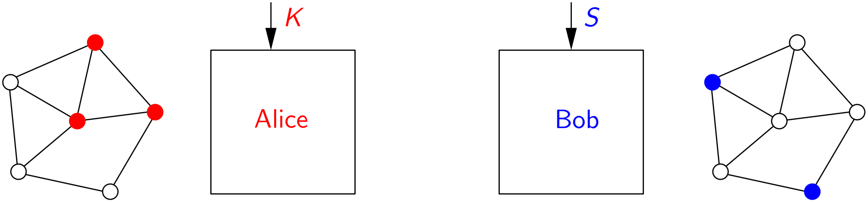
$\exists O(\log^2 n)$ complexity protocol for computing $S(\text{STAB}(G))$ $\square$

Deterministic communication protocols

Perfect without claw

Example: $P = \mathrm{STAB}(G)$ for G claw-free perfect graph

$\exists\ 3\log n + O(1)$ complexity protocol for $S(P) \implies \boxed{\mathrm{xc}(P) = O(n^3)}$

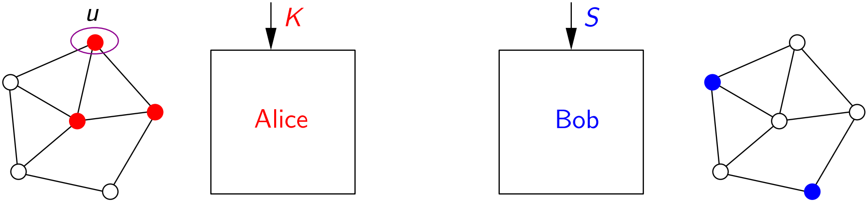


Deterministic communication protocols

Perfect without claw

Example: $P = \text{STAB}(G)$ for G claw-free perfect graph

$\exists\ 3 \log n + O(1)$ complexity protocol for $S(P) \implies \boxed{\text{xc}(P) = O(n^3)}$

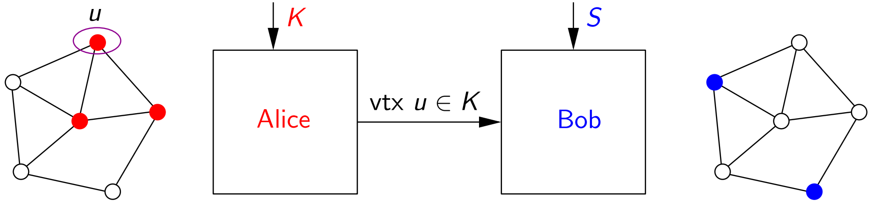


Deterministic communication protocols

Perfect without claw

Example: $P = \text{STAB}(G)$ for G claw-free perfect graph

$\exists\ 3\log n + O(1)$ complexity protocol for $S(P) \implies \boxed{\text{xc}(P) = O(n^3)}$

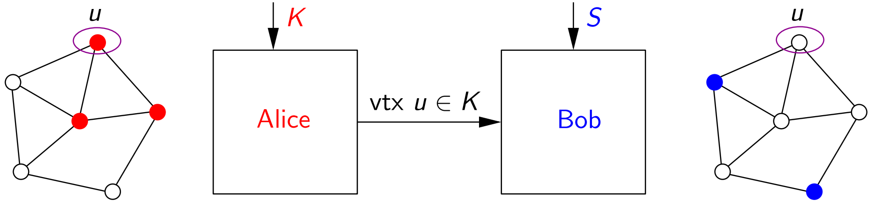


Deterministic communication protocols

Perfect without claw

Example: $P = \mathrm{STAB}(G)$ for G claw-free perfect graph

$\exists\ 3\log n + O(1)$ complexity protocol for $S(P) \implies \boxed{\mathrm{xc}(P) = O(n^3)}$

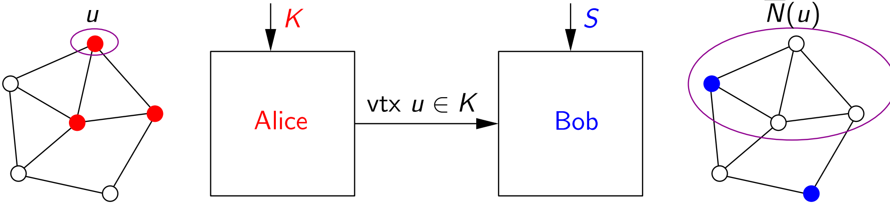


Deterministic communication protocols

Perfect without claw

Example: $P = \mathrm{STAB}(G)$ for G claw-free perfect graph

$\exists\ 3 \log n + O(1)$ complexity protocol for $S(P) \implies \boxed{\mathrm{xc}(P) = O(n^3)}$

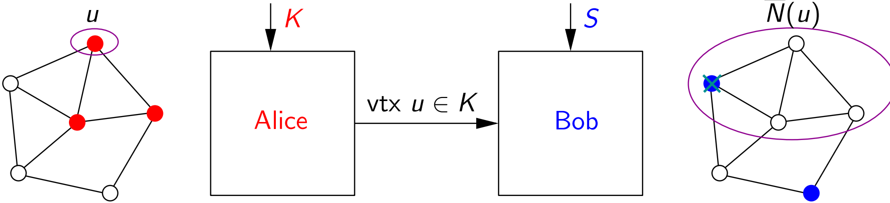


Deterministic communication protocols

Perfect without claw

Example: $P = \text{STAB}(G)$ for G claw-free perfect graph

$\exists\ 3 \log n + O(1)$ complexity protocol for $S(P) \implies \boxed{\text{xc}(P) = O(n^3)}$

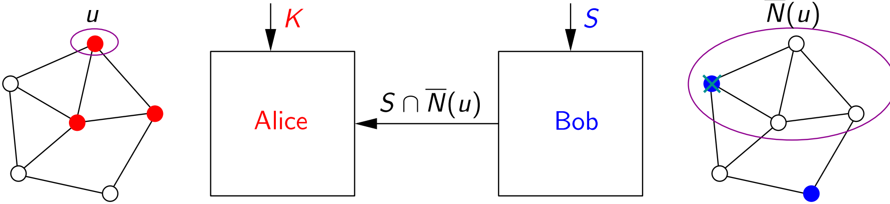


Deterministic communication protocols

Perfect without claw

Example: $P = \mathrm{STAB}(G)$ for G claw-free perfect graph

$\exists\ 3 \log n + O(1)$ complexity protocol for $S(P) \implies \boxed{\mathrm{xc}(P) = O(n^3)}$

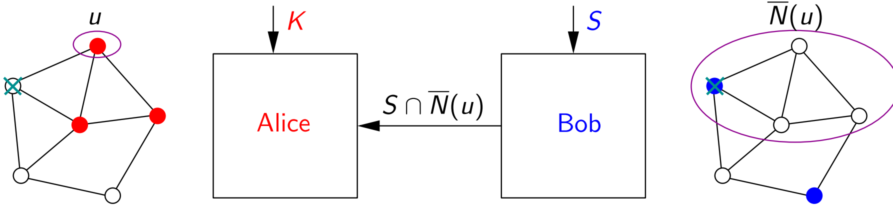


Deterministic communication protocols

Perfect without claw

Example: $P = \mathrm{STAB}(G)$ for G claw-free perfect graph

$\exists\ 3\log n + O(1)$ complexity protocol for $S(P) \implies \boxed{\mathrm{xc}(P) = O(n^3)}$

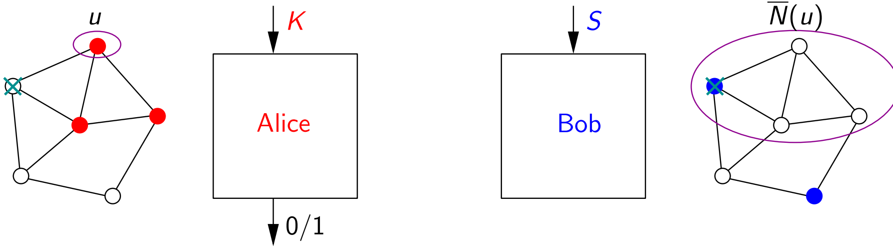


Deterministic communication protocols

Perfect without claw

Example: $P = \mathrm{STAB}(G)$ for G claw-free perfect graph

$\exists\ 3\log n + O(1)$ complexity protocol for $S(P)$ $\implies$ $\boxed{\mathrm{xc}(P) = O(n^3)}$



Deterministic communication protocols

How good are they?

Limitations:

1. binary slack matrices $S(P)$
2. deterministic protocols imply disjoint rectangles

Deterministic communication protocols

How good are they?

Limitations:

1. binary slack matrices $S(P)$
2. deterministic protocols imply disjoint rectangles

Possible solutions:

1. Alice and Bob could output non-binary values
2. randomized protocols would allow overlapping rectangles

Our main result

A tight relationship between extended formulations and communication complexity

Theorem

If c is the minimum complexity of a randomized communication protocol computing $S(P)$ on average, then

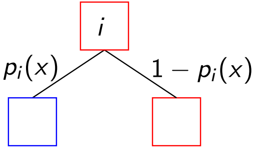
$$\boxed{xc(P) = rank_+(S(P)) = \Theta(2^c)}$$

Randomized communication protocols

Introducing the model

Similar to deterministic protocols except that

- Alice and Bob can use private random bits to make a choice



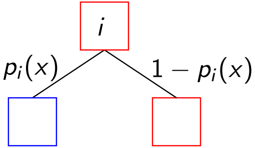
- The value output on given input (x, y) is a random variable

Randomized communication protocols

Introducing the model

Similar to deterministic protocols except that

- Alice and Bob can use private random bits to make a choice



- The value output on given input (x, y) is a random variable

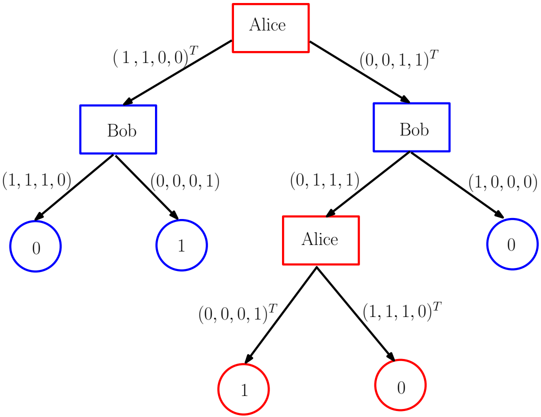
We say that the protocol *computes f on average* if: $\forall (x, y) \in X \times Y$,

$$E[\text{value output by the protocol on input } (x, y)] = f(x, y)$$

Randomized communication protocols

Go randomized

The previous deterministic protocol can be viewed as a randomized protocol where all the probabilities are either zero or one



In general can use arbitrary probabilities, and nonnegative values!

Proof of the theorem

Factorization $\implies$ protocol

Write $S(P) = TU$, where

- $T \in \mathbb{R}_+^{m \times r}$ row stochastic (w.l.o.g.)
- $U \in \mathbb{R}_+^{r \times n}$
- $r = \text{rank}_+(S(P)) + 1$

Proof of the theorem

Factorization $\implies$ protocol

Write $S(P) = TU$, where

$$\left| \begin{array}{l} T \in \mathbb{R}_+^{m \times r} \text{ row stochastic (w.l.o.g.)} \\ U \in \mathbb{R}_+^{r \times n} \\ r = \text{rank}_+(S(P)) + 1 \end{array} \right.$$

Protocol:

- Alice gets row index i, Bob gets column index j
- Alice picks random column index k w.p. t_{ik}, sends it to Bob
- Bob outputs value u_{kj}

Proof of the theorem

Factorization $\Longrightarrow$ protocol

Write $S(P) = TU$, where

$$\left|\begin{array}{l} T \in \mathbb{R}_+^{m\times r} \text{ row stochastic (w.l.o.g.)} \\ U \in \mathbb{R}_+^{r\times n} \\ r = \text{rank}_+(S(P)) + 1 \end{array}\right.$$

Protocol:

- Alice gets row index i, Bob gets column index j
- Alice picks random column index k w.p. t_{ik}, sends it to Bob
- Bob outputs value u_{kj}

Expected value on input (i, j): $\sum_{k=1}^{r} t_{ik} u_{kj} = s_{ij}$

Complexity: $\log \text{rank}_+(S(P)) + O(1)$

New lower bound for the perfect matching polytope of K_n

Now: $P =$ perfect matching polytope of K_n

Slack matrix: rows correspond to odd sets S
columns correspond to matchings M
slack of M w.r.t. S is $|M \cap \delta(S)| - 1$

New lower bound for the perfect matching polytope of K_n

Now: $P =$ perfect matching polytope of K_n

Slack matrix: rows correspond to odd sets S
columns correspond to matchings M
slack of M w.r.t. S is $|M \cap \delta(S)| - 1$

Theorem ("small randomness implies large size")

Consider an extension for P and a corresponding randomized protocol. If the probability that the protocol outputs a non-zero value, given a pair (S, M) with non-zero slack, is at least $p(n) \gg 1/n$, then the protocol has complexity $\Omega(np(n))$ and the extended formulation has size $2^{\Omega(np(n))}$

New lower bound for the perfect matching polytope of K_n

Now: $P =$ perfect matching polytope of K_n

Slack matrix: rows correspond to odd sets S
columns correspond to matchings M
slack of M w.r.t. S is $|M \cap \delta(S)| - 1$

Theorem ("small randomness implies large size")

Consider an extension for P and a corresponding randomized protocol. If the probability that the protocol outputs a non-zero value, given a pair (S, M) with non-zero slack, is at least $p(n) \gg 1/n$, then the protocol has complexity $\Omega(np(n))$ and the extended formulation has size $2^{\Omega(np(n))}$

Particular case: If a non-zero slack is detected with constant probability, then extension has exponential size

New lower bound for the perfect matching polytope of K_n

Now: $P =$ perfect matching polytope of K_n

Slack matrix: rows correspond to odd sets S
columns correspond to matchings M
slack of M w.r.t. S is $|M \cap \delta(S)| - 1$

Theorem ("small randomness implies large size")

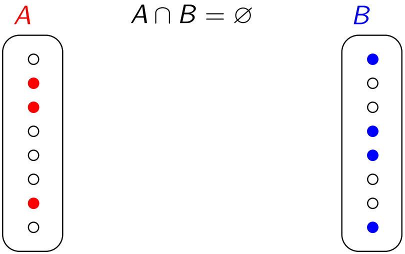
Consider an extension for P and a corresponding randomized protocol. If the probability that the protocol outputs a non-zero value, given a pair (S, M) with non-zero slack, is at least $p(n) \gg 1/n$, then the protocol has complexity $\Omega(np(n))$ and the extended formulation has size $2^{\Omega(np(n))}$

Particular case: If a non-zero slack is detected with constant probability, then extension has exponential size

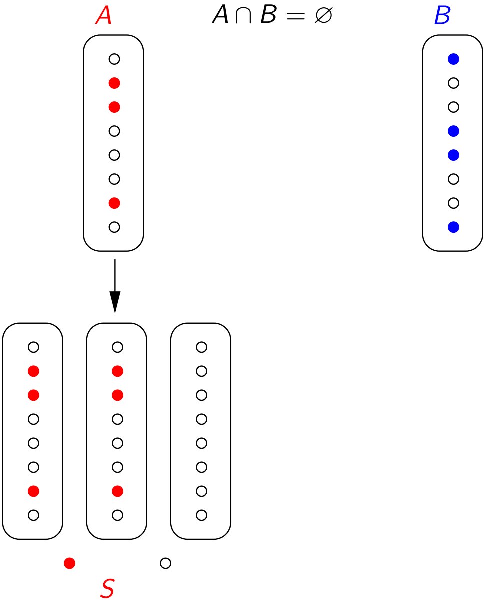
Even more particular case: If the protocol is deterministic, then extension has exponential size

Proof sketch (reduction from SET DISJOINTNESS)

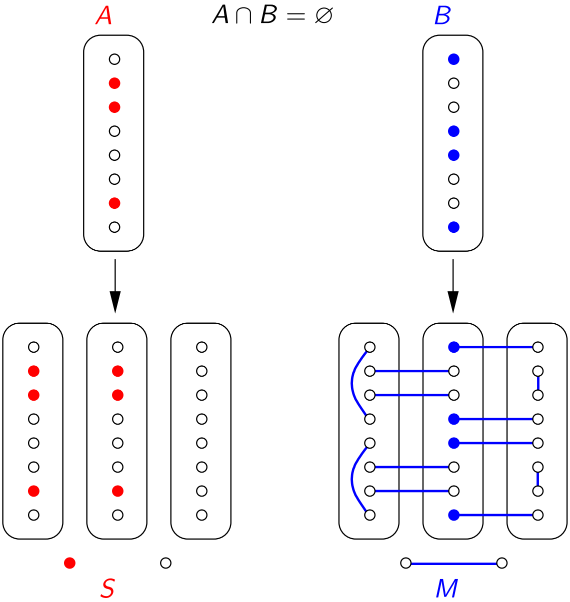
Proof sketch (reduction from SET DISJOINTNESS)



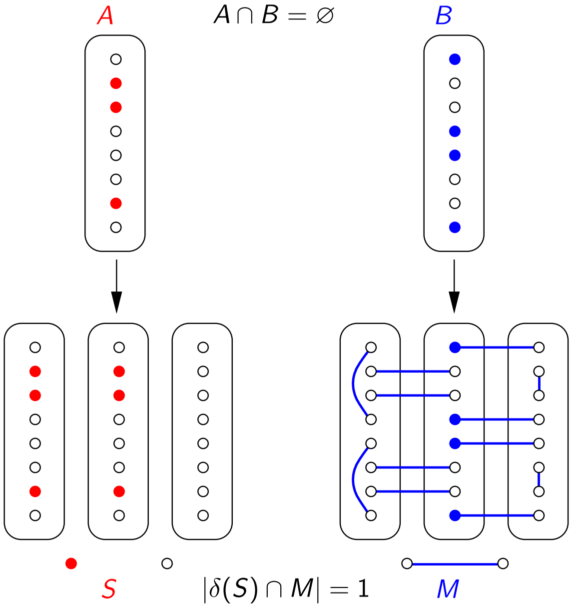
Proof sketch (reduction from SET DISJOINTNESS)



Proof sketch (reduction from SET DISJOINTNESS)

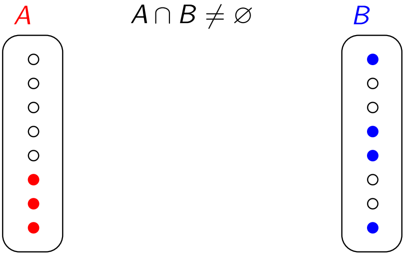


Proof sketch (reduction from SET DISJOINTNESS)

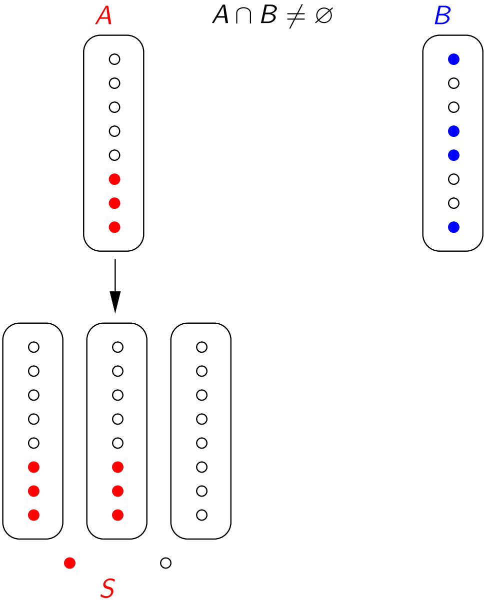


Proof sketch (reduction from SET DISJOINTNESS)

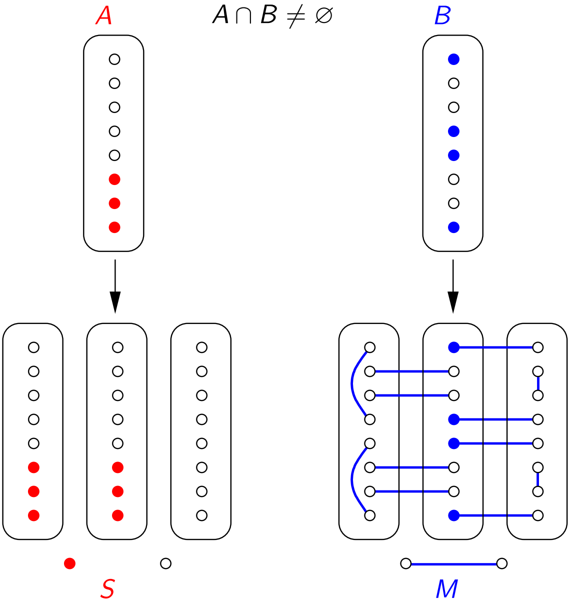
Proof sketch (reduction from SET DISJOINTNESS)



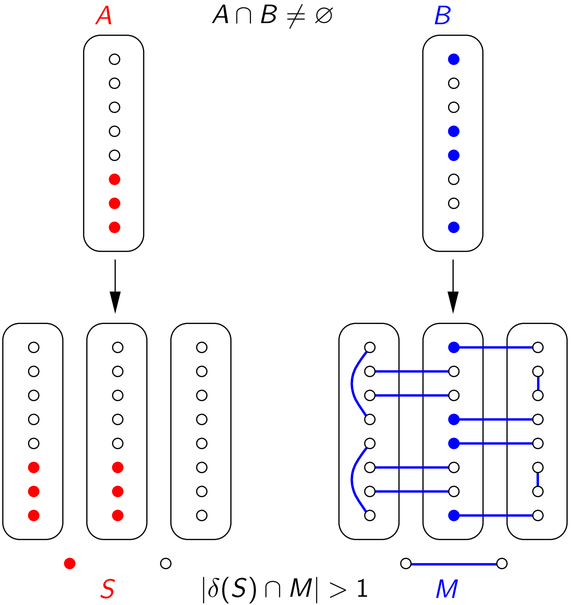
Proof sketch (reduction from SET DISJOINTNESS)



Proof sketch (reduction from SET DISJOINTNESS)



Proof sketch (reduction from SET DISJOINTNESS)



Proof sketch (wrapping up)

Next,

- repeat protocol $\Theta(1/p(n))$ times on pair (S, M), independently
- declare A and B non-disjoint as soon as know that $\text{slack}(S, M) \neq 0$
- otherwise, declare A and B disjoint

Proof sketch (wrapping up)

Next,

- repeat protocol $\Theta(1/p(n))$ times on pair (S, M), independently
- declare A and B non-disjoint as soon as know that $\text{slack}(S, M) \neq 0$
- otherwise, declare A and B disjoint

Finally, use

Fact. computing DISJ with high probability needs $\Omega(n)$ bits

Concluding remarks / questions

- "Small randomness implies large size" also holds for spanning trees!!

Concluding remarks / questions

- "Small randomness implies large size" also holds for spanning trees!!

- $\exists$ nice characterization of polytopes with binary slack matrix?

Concluding remarks / questions

- "Small randomness implies large size" also holds for spanning trees!!
- $\exists$ nice characterization of polytopes with binary slack matrix?
- What is the communication complexity of the relation

$$\{(S, M, e) : S \text{ odd set}, M \text{ perfect matching}, e \in \delta(S) \cap M\} \text{ ?}$$

Concluding remarks / questions

- "Small randomness implies large size" also holds for spanning trees!!
- $\exists$ nice characterization of polytopes with binary slack matrix?
- What is the communication complexity of the relation

$$\{(S, M, e) : S \text{ odd set}, M \text{ perfect matching}, e \in \delta(S) \cap M\} \ ?$$

- ("log rank conjecture"-inspired) For which polytopes do we have

$$\text{randomized-cc}(\text{support of } S(P)) \leqslant \text{poly}(\log \text{rank}_+ S(P)) \ ?$$


Communication
Complexity
Eyal Kushilevitz and Noam Nisan

Thank You!