LECTURE 26

Cutting planes proof system

Input: Inconsistent system of 0-1 linear inequalities

Refutation: Derive $0 \geq 1$

Configuration-style proof

At each derivation step

1. DOWNLOAD axiom constraint
2. Apply INFERENCE rule to constraints in memory
3. ERASE constraint

Inference rules

well, axiom download, technically speaking

Variable axioms $\dfrac{}{x \geq 0}$ $\dfrac{}{-x \geq -1}$

Addition $$\frac{\sum_i a_i x_i \geq A \qquad \sum_i b_i x_i \geq B}{\sum_i (a_i + b_i) x_i \geq A + B}$$

Multiplication $$\frac{\sum_i a_i x_i \geq A}{\sum_i c a_i x_i \geq c A} \quad c \in \mathbb{N}^+$$

Division $$\frac{\sum_i c a_i x_i \geq A}{\sum_i a_i x_i \geq \lceil A/c \rceil}$$

Complexity measures:

Length = # constraints in derivation

Line space = max # constraints in memory

What about $\boxed{\text{magnitude of coefficients?}}$

[Buss & Clote '96] building on [Cook, Coullard & Turán '87]

(a) Cutting planes with division only by fixed $k \geq 2$ is as powerful as general cutting planes (up to polynomial factors)

(b) Suppose coefficients and constants have absolute values $\leq B$ and that cutting planes refutes input in length L. Then $\exists$ refutation in length $O(L^3 \log B)$ with coefficients and constants of absolute value $O(L^2 \cdot B \cdot 2^L)$.

$\boxed{\text{So coefficients need not have more than polynomial \# bits / exponential magnitude}}$

[Dadush & Tiwari '20] proved analogous result for stabbing planes.

OPEN PROBLEM: Possible to bring this down to logarithmic # bits / polynomial magnitude? Buss & Clote state that this was their goal.

Still remains open!

What would separating formulas look like?

Define CP* as cutting planes, but in any derivation the coefficients and constant terms should have size at most polynomial in size of input

i.e., magnitude

= logarithmic # bits

Aside: CP* also defined by requiring integers to have magnitude at most polynomial in input size and exponential in # steps of refutation. Essentially Same definition if we insist on polynomial-length refutations. We will define CP* in terms of input.

Can we prove that there is something CP can do efficiently that CP* cannot?

Yes! [dRMNPRV '20]

There are families of CNF formulas $\{F_n\}_{n=1}^{\infty}$ such that

- Cutting planes refutes F_n in (roughly) quadratic length and constant line space simultaneously.
- CP* cannot refute F_n in subexponential length and subpolynomial line space simultaneously

MAIN TECHNICAL INGREDIENT

Lifting theorem using equality gadget

HIGH-LEVEL IDEA

Take HORN FORMULA: At most 1 positive literal/clause
Can be refuted by deriving unit clauses z_i
in some order in resolution

Make this line-space-efficient in cutting planes
by deriving

$$\sum_{i=0}^{n-1} 2^i z_i = \sum_{i=0}^{n-1} 2^i$$
$$= 2^n - 1$$

(Note that $\boxed{\sum_i a_i z_i = A}$ is syntactic
sugar for

$$\boxed{\begin{array}{l} \sum_i a_i z_i \geq A \\ \sum_i -a_i z_i \geq -A \end{array}}$$
)

Lift formula F with EQUALITY GADGET

$$EQ(x,y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{o/w} \end{cases} \qquad x,y \in \{0,1\}$$

EXAMPLE

$$C = z_1 \vee \overline{z}_2$$

Then $C[EQ] = C \circ EQ =$

$$\begin{aligned} & (x_1 \vee \overline{y}_1 \vee x_2 \vee y_2) \\ \wedge\ & (x_1 \vee \overline{y}_1 \vee \overline{x}_2 \vee \overline{y}_2) \\ \wedge\ & (\overline{x}_1 \vee y_1 \vee x_2 \vee y_2) \\ \wedge\ & (\overline{x}_1 \vee y_1 \vee \overline{x}_2 \vee \overline{y}_2) \end{aligned}$$

(A) Prove that line-space-efficient CP refutation still works for $F \circ EQ$ if F Horn formula

Derive (in)equalities

$$\sum_{i=0}^{n} 2^i (x_i - y_i) = 0 \qquad (*)$$

Whenever, say, z_k followed from

z_i
z_j
$\overline{z}_i \vee \overline{z}_j \vee z_k$

"decode"

$x_i = y_i$
$x_j = y_j$

from $(*)$ and apply to

$(\overline{z}_i \vee \overline{z}_j \vee z_k) \circ EQ$

to derive

$x_k = y_k$

and add to $(*)$. Want to do this length- and space-efficiently (REQUIRES A PROOF)

Yields upper bound for general cutting planes.

(B) Suppose there is a short, line-space-efficient refutation π^* in CP^* of $F_n \circ EQ$ in length L and line space s

CP^* VI

Yields deterministic communication protocol for $\mathrm{Search}(F_n) \circ EQ$ in cost

$$\approx s \log L$$

Alice & Bob can evaluate the inequalities and send number — logarithmic # bits

Prove lifting theorem relating communication complexity D^{cc} with decision tree query complexity D^{dt} by

$$\boxed{D^{cc}(\mathrm{Search}(F) \circ EQ) \gtrsim D^{dt}(\mathrm{Search}(F))}$$

Plug in Horn formulas with large decision tree query cplx — PEBBLING FORMULAS

DONE Right?

Except [Loff & Mukhopadhyay '19] show that such lifting theorem is NOT TRUE for
- equality gadget
- relations/search problems (as opposed to functions)

So instead
- Use equality gadget over non-constant # bits
- Lift Nullstellensatz refutation degree (happens to be = query cplx for pebbling formulas)

$\boxed{EQ_q}: \{0,1\}^q \times \{0,1\}^q \to \{0,1\}$ $\quad$ $\boxed{EQ_q(x,y) = 1 \text{ iff } x = y}$

MAIN LIFTING THEOREM

Suppose that

- F minimally unsatisfiable CNF formula over n variables
- $\mathbb{F}$ any field
- $g : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ any gadget such that

Rank of matrix $(g(x,y))_{x \in \mathcal{X}, y \in \mathcal{Y}}$

$$\boxed{\mathrm{rank}_{\mathbb{F}}(g) \geq \frac{6en}{\mathrm{Deg}^{\mathbb{F}}_{NS}(F \vdash \bot)}}$$

Then

$$\boxed{D^{cc}(\mathrm{Search}(F) \circ g) \geq \mathrm{Deg}^{\mathbb{F}}_{NS}(F \vdash \bot)}$$

UPPER BOUND FOR CP

Suppose that

- G any DAG with constant fan-in & single sink
- $q \in \mathbb{N}^+$, $q = \Theta(\log \log n)$

Rank $2^{\Theta(\log\log n)} = \Theta(\log n)$

Then the formula $\boxed{Peb_G \circ EQ_q}$ has

- $O(n \log \log n)$ variables
- $\tilde{O}(n)$ clauses of width $O(\log \log n)$
- cutting planes refutation in simultaneous length $\tilde{O}(n^2)$ and line space $O(1)$

$\tilde{O}(f(n))$ means $O\left(f(n) (\log(f(n)))^k\right)$ for some constant k

LOWER BOUND FOR CP*

Any CP* refutation of $Peb_G \circ EQ_q$ as above in length L and line space s must satisfy

$$\boxed{s \log L = \Omega(n / \log^2 n)}$$

Equality gadget provides a sweet spot!

- Hard for deterministic communication (which can use CP* proofs)
- Easy for randomized and real communication (otherwise we would get hardness for general cutting planes

SOME OPEN PROBLEMS

1. Size separation for CP vs CP*?
2. Line space lower bounds for CP*?
3. True length-space trade-offs for CP* that do not apply for CP?
4. Direct lower bound proof for parity decision tree query cplx for pebbling formulas

   (PDTs describe an obvious, and maybe optimal, class of protocols for Alice & Bob)

A $\boxed{\text{(TOTAL) SEARCH PROBLEM}}$ is a relation $\mathcal{S} \subseteq \mathcal{I} \times \mathcal{O}$ such that for all $z \in \mathcal{I}$ there exists $o \in \mathcal{O}$ for which $(z, o) \in \mathcal{S}$

Think of this as computational task:
Given z, find o s.t. $(z, o) \in \mathcal{S}$

If $\mathcal{I} = I^n$ has product structure, and $g : \mathcal{X} \times \mathcal{Y} \to I$ is a function (a GADGET), then the $\boxed{\text{COMPOSED / LIFTED SEARCH PROBLEM}}$ $\mathcal{S} \circ g^n \subseteq (\mathcal{X}^n \times \mathcal{Y}^n) \times \mathcal{O}$ is the task, given $x \in \mathcal{X}^n$ and $y \in \mathcal{Y}^n$ to find o s.t. $(g^n(x, y), o) \in \mathcal{S}$ where

$$\boxed{g^n(x, y) = (g(x_1, y_1), g(x_2, y_2), \ldots, g(x_n, y_n))}$$

Our previous lifting theorems worked for any search problem

Now we have to focus on $\boxed{\text{FALSIFIED CLAUSE SEARCH PROBLEM}}$: Given assignment α to (fixed) unsatisfiable CNF formula F, find clause C falsified by α.

Denote this problem $\boxed{\text{Search}(F)}$

Lifted search problems yield natural communication problems

DETERMINISTIC COMMUNICATION PROTOCOL

Two players Alice with input $x \in \mathcal{X}^n$
Bob with input $y \in \mathcal{Y}^n$

Protocol tree Π

- Every internal node v labelled by function
  $f_v^A : \mathcal{X}^n \to \{0,1\}$ (Alice speaks) or
  $f_v^B : \mathcal{Y}^n \to \{0,1\}$ (Bob speaks)
- Every internal node has two outgoing edges labelled 0 and 1, respectively
- Input $\alpha \in \mathcal{X}^n \times \mathcal{Y}^n$ defines path to leaf l_α
- Leaf l_α should be labelled by answer to $S \circ g^n$
- Cost of protocol Π = length of longest path = max # bits communicated
- For problem P, write $\boxed{D^{cc}(P)}$ for minimal cost of any protocol

Given any gadget $g : \{0,1\}^q \times \{0,1\}^q \to \{0,1\}$ and CNF formula F, can define **LIFTED FORMULA** $F[g]$ or $F \circ g$ by

- replace all literals z_i by CNF encoding of $g(x_{i,1}, \ldots, x_{i,q}, y_{i,1}, \ldots, y_{i,q})$
- replace all literals $\overline{z}_i$ by CNF encoding of $\neg g(x_{i,1}, \ldots, x_{i,q}, y_{i,1}, \ldots, y_{i,q})$
- expand all clauses $C \in F$ to CNF in canonical way

OBSERVATION 1

For any unsatisfiable CNF formula F and any gadget g,

$$D^{cc}(\mathrm{Search}(F \circ g)) \geq D^{cc}(\mathrm{Search}(F) \circ g)$$

We will be interested in the RANK of gadgets

For $g : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$, the $\boxed{\text{RANK}}$ of g over the field $\mathbb{F}$, denoted $\mathrm{rank}_{\mathbb{F}}(g)$, is the rank over $\mathbb{F}$ of the matrix with

- rows indexed by $x \in \mathcal{X}$
- columns indexed by $y \in \mathcal{Y}$
- the cell (x,y) containing $g(x,y)$

EXAMPLE The gadget $EQ^q : \{0,1\}^q \times \{0,1\}^q \to \{0,1\}$ defined by

$$EQ^q(x,y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{otherwise} \end{cases}$$

has $\mathrm{rank}_{\mathbb{F}}(EQ^q) = 2^q$ over any field $\mathbb{F}$

LEMMA 2 [HN 12]

If there is a cutting planes refutation $\pi : F \vdash \bot$ in length L, line space s, and coefficients and constant terms (absolute values) bounded by B, where F is over n variables, then

$$D^{cc}(\mathrm{Search}(F)) = O(s \cdot (\log B + \log n) \log L)$$

$\boxed{\text{NULLSTELLENSATZ}}$ [BIKPP '94]

Given field $\mathbb{F}$

Polynomials $\mathcal{P} = \{p_1, \ldots, p_m\}$ over $x_1, \ldots, x_n$

Boolean axioms $x_j^2 - x_j \quad j \in [n]$

a NULLSTELLENSATZ REFUTATION is a sequence of polynomials $q_1, \ldots, q_m, r_1, \ldots, r_n$ s.t. the syntactic equality

$$\boxed{\sum_{i=1}^{m} q_i p_i + \sum_{j=1}^{n} r_j (x_j^2 - x_j) = 1} \quad (*)$$

holds (after cancellations).

Proof system for CNF formulas by translating clauses

$$\underline{C} = \underbrace{\bigvee_{z \in P} z}_{} \vee \underbrace{\bigvee_{z \in N} \overline{z}}_{}$$

to

$$\boxed{p(C) = \prod_{z \in P} (1 - z) \cdot \prod_{z \in N} z}$$

The $\boxed{\text{DEGREE}}$ of a Nullstellensatz refutation is the largest total degree of a left-hand side polynomial in $(*)$

$\boxed{Deg_{NS}^{\mathbb{F}}(F \vdash \bot)}$ = min NS degree of any refutation of F over $\mathbb{F}$

Let $\mathcal{P} \in \mathbb{F}[\vec{z}]$ be set of polynomials and $d \in \mathbb{N}^+$

A $\boxed{d\text{-DESIGN}}$ for $\mathcal{P}$ is a mapping D of polynomials in $\mathbb{F}[\vec{z}]$ of degree $\leq d$ to $\mathbb{F}$ such that

(1) D is linear

(2) $D(1) = 1$

(3) $D(q p_i) = 0$ for all $p_i \in \mathcal{P}$ and all q such that $Deg(q p_i) \leq d$

(4) $D(z_i^2 q) = D(z_i q)$ for all q s.t. $Deg(q) \leq d-1$

THEOREM 3 [Buss '98]

Suppose $\mathcal{P} \in \mathbb{F}[\vec{z}]$ is such that $z_i^2 - z_i \in \mathcal{P}$ for all z_i. Then $\boxed{Deg_{NS}^{\mathbb{F}}(\mathcal{P} \vdash 1) > d}$ if and only if $\mathcal{P}$ has a d-design.

THEOREM 4 [dRMNPRV '20]

For any single-sink DAG G and any field $\mathbb{F}$ it holds that $\boxed{Deg_{NS}^{\mathbb{F}}(Peb_G \vdash 1)}$ coincides with the reversible pebbling price of G.

Proof sketch Let $V(G) = \{1, 2, \ldots, n\}$

Identify $S \subseteq [n]$ with $z_S = \prod_{i \in S} z_i$

For fixed $d \in \mathbb{N}^+$, define

$D(z_S) = 1$ is pebble configuration reachable from $\emptyset$ by reversible pebbling in space $\leq d$

$D(z_S) = 0$ otherwise

This is a d-design iff reversible pebbling price of $G > d$.

Just for the record, $\boxed{Peb_G}$ is the set of polynomials

- $1 - z_s$ for each source vertex s
- $(1 - z_v) \prod_{u \in pred(v)} u$ for non-source vertex v with immediate predecessors $pred(v)$
- z_t for the sink / target vertex t
- and also $z_v^2 - z_v$ for all v

In what remains of this lecture, focus on Nullstellensatz lifting theorem that yields lower bound for CP^*

LIFTING THEOREM, FULL VERSION

Suppose

- F unsatisfiable k-CNF formula over n variables
- $\mathbb{F}$ any field
- $g : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ with $R_g = rank_{\mathbb{F}}(g) \geq 3$
- $D_F = Deg_{NS}^{\mathbb{F}}(F \vdash \bot)$

Then

$$\boxed{D^{cc}(Search(F) \circ g^n) \geq D_F \log\left(\frac{D_F \cdot R_g}{en}\right) - \frac{4n \log e}{R_g} - \log k}$$

Earlier version can be shown to follow from this

Main non-obvious point: We should have
$\log k \leq k \leq D_F$ if clauses of width k are needed in NS certificate

Lifting theorem relies (very) heavily on [Pitassi-Robere '18]

Let $\mathcal{U}, \mathcal{V}$ be sets

A $\boxed{\text{COMBINATORIAL RECTANGLE}}$ R in $\mathcal{U} \times \mathcal{V}$ is a set $R = A \times B$ for $A \subseteq \mathcal{U}$, $B \subseteq \mathcal{V}$

A $\boxed{\text{RECTANGLE PARTITION}}$ $\mathcal{P} = \{ R_i \mid i \in [t] \}$ of $\mathcal{U} \times \mathcal{V}$ is a set of rectangles such that for all $(u,v) \in \mathcal{U} \times \mathcal{V}$ $\exists! \, R_i \in \mathcal{P}$ s.t. $(u,v) \in R_i$

FACT Deterministic communication protocol splits input space into rectangle partition.

A $\boxed{\text{RECTANGLE COVER}}$ $\mathcal{R} = \{ R_i \mid i \in [t] \}$ of $\mathcal{U} \times \mathcal{V}$ is a set of rectangles such that $\mathcal{U} \times \mathcal{V} \subseteq \bigcup_{i \in [t]} R_i$

Given $\mathcal{U} \times \mathcal{V}$ matrix A and rectangle R in $\mathcal{U} \times \mathcal{V}$, let $\boxed{A|R}$ be submatrix of A induced on R

DEFINITION (Razborov)

Let $\mathcal{R}$ be rectangle cover of $\mathcal{U} \times \mathcal{V}$ and A $\mathcal{U} \times \mathcal{V}$ matrix over $\mathbb{F}$. Then the $\mathbb{F}$-RANK MEASURE of $\mathcal{R}$ at A is

$$\boxed{\mu_{\mathbb{F}}(\mathcal{R}, A) = \frac{\operatorname{rank}_{\mathbb{F}}(A)}{\max_{R \in \mathcal{R}} \operatorname{rank}_{\mathbb{F}}(A|R)}}$$

Can be used to show lower bounds for deterministic communication (and several other computational models — see Robere's PhD thesis in 2018)

Notation If $A \subseteq \mathcal{I}^n$ write A_j / A_J for projection of A to coordinate j / coordinates J

DEFINITION

Let F unsatisfiable CNF formula over n variables

$g : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ gadget

For $C \in F$, say that combinatorial rectangle $R \subseteq \mathcal{X}^n \times \mathcal{Y}^n$ is $\boxed{\text{C-STRUCTURED}}$ if

(1) $g^n(x,y)$ falsifies C $\quad \forall (x,y) \in R$

(2) $\forall z_i \notin Vars(C) \quad R_i = \mathcal{X} \times \mathcal{Y}$

A rectangle cover is $\boxed{\text{F-STRUCTURED}}$ if all rectangles in it are C-structured for $C \in F$

LEMMA 5

Let

- F unsatisfiable CNF formula over n variables
- $g : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ gadget
- $\mathbb{F}$ field

Then

$$\boxed{D^{cc}(Search(F) \circ g) \geq \max_A \min_R \log \mu_{\mathbb{F}}(R, A)}$$

where A ranges over $\mathcal{X}^n \times \mathcal{Y}^n$ matrices over $\mathbb{F}$ and R over F-structured rectangle covers

<u>Proof</u> Let Π be protocol solving $Search(F) \circ g$ and let P be induced monochromatic rectangle partition

Every $R = A \times B$ in P labelled by $C \in F$

For all $(x,y) \in R$ $\quad g^n(x,y)$ falsifies C

We have
$$A \subseteq \mathcal{X}^{Vars(C)} \times \mathcal{X}^{[n] \setminus Vars(C)}$$
$$B \subseteq \mathcal{Y}^{Vars(C)} \times \mathcal{Y}^{[n] \setminus Vars(C)}$$

(overloading z_i and index i)

Let $A' = A_{Vars(C)} \times \mathcal{X}^{[n] \setminus Vars(C)}$

$B' = B_{Vars(C)} \times \mathcal{Y}^{[n] \setminus Vars(C)}$

$R' = A' \times B'$

Then $R' \supseteq R$ and R' is C-structured

Let $\mathcal{R}$ be F-structured rectangle cover obtained in this way.

[Razborov '90] proved that if $\mathcal{P}$ rectangle partition and $\mathcal{R}$ rectangle cover such that $\forall R \in \mathcal{P}\ \exists R'(R) \in \mathcal{R}$ for which $R \subseteq R'(R)$, then

$$\begin{aligned} \mathrm{rank}_{\mathbb{F}}(A) &\leq \sum_{R \in \mathcal{P}} \mathrm{rank}_{\mathbb{F}}(A \mid R) \\ &\leq \sum_{R \in \mathcal{P}} \mathrm{rank}_{\mathbb{F}}(A \mid R'(R)) \\ &\leq |\mathcal{P}| \max_{R' \in \mathcal{R}} \mathrm{rank}_{\mathbb{F}}(A \mid R') \end{aligned}$$

or, in other words,

$$|\mathcal{P}| \geq \frac{\mathrm{rank}_{\mathbb{F}}(A)}{\max_{R' \in \mathcal{R}} \mathrm{rank}_{\mathbb{F}}(A \mid R')} = \mu_{\mathbb{F}}(\mathcal{R}, A)$$

A binary tree with T leaves has height $\geq \log T$, so

$$D^{cc}(\mathrm{Search}(F) \circ g) \geq \log |\mathcal{P}| \geq \log \mu_{\mathbb{F}}(\mathcal{R}, A)$$

as claimed

$\blacksquare$

For a clause $C \in F$ over $z_1, \ldots, z_n$, the $\boxed{\text{CERTIFICATE}}$ of C $Cert(C)$ is the smallest partial assignment π falsifying C.

$$\boxed{Cert(F) = \{ Cert(C) \mid C \in F \}}$$

For $\alpha \in \{0,1\}^n$, α AGREES with $\pi \in Cert(F)$ if $\pi(z_i) = \alpha_i$ if $\pi(z_i) \neq *$

If F is unsatisfiable, every α agrees with some $\pi \in Cert(F)$

DEFINITION

Let

- F unsatisfiable CNF formula
- $\mathbb{F}$ field
- $p \in \mathbb{F}[\vec{z}]$ multilinear polynomial

The $\boxed{\mathbb{F}\text{-ALGEBRAIC GAP COMPLEXITY}}$ of F at p is

$$\boxed{gap_{\mathbb{F}}(F, p) = Deg(p) - \max_{\pi \in Cert(F)} Deg(p \restriction_{\pi})}$$

The $\mathbb{F}$-algebraic gap complexity of F is

$$\boxed{gap_{\mathbb{F}}(F) = \max \{ gap_{\mathbb{F}}(F, p) \mid Deg(p) = n \}}$$

THEOREM 6 [Pitassi & Robere '18]

For any unsatisfiable formula F over n variables and any field $\mathbb{F}$, it holds that

$$\boxed{Deg^{\mathbb{F}}_{NS}(F \vdash \bot) = gap_{\mathbb{F}}(F)}$$

This theorem would be worth a separate lecture...

Example

$F \doteq \overline{z}_1 \wedge \overline{z}_2 \wedge \dots \wedge \overline{z}_n \wedge (z_1 \vee z_2 \vee \dots \vee z_n)$

Choose $p = OR_n$

$$= 1 - \prod_{i=1}^{n} (1 - z_i)$$

$deg(p) = n$

For $C_i = \overline{z}_i$ $\pi_i := Cert(C_i) = \{z_i \mapsto 1\}$

$p \restriction \pi_i = 1$ $\quad Deg(p \restriction \pi_i) = 0$

For $C_{n+1} = (z_1 \vee \dots \vee z_n)$ $\pi_{n+1} = Cert(C_{n+1})$

$$= \{ z_i \mapsto 0 \mid i \in [n] \}$$

$p \restriction \pi_{n+1} = 0$ $\quad$ again $Deg(p \restriction \pi_{n+1}) = 0$

So $gap_{\pi}(F, p) = n$

(QUESTION: Why can't we use F to get our lower bound?)

What we want to do now

1. Take polynomial P witnessing gap complexity
2. Compose $\boxed{p \circ g^n}$ to get $\mathcal{X}^n \times \mathcal{Y}^n$ matrix A
3. Argue that A yields large $\mathbb{F}$-rank measure
4. This establishes communication complexity lower bound

Let $p \in \mathbb{F}[\vec{z}]$ multilinear polynomial

With notation inspired by Fourier analysis, write

$$\boxed{p = \sum_{S \subseteq [n]} \hat{p}(S) \prod_{i \in S} z_i}$$

For $g : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$, define lifted polynomial

$$\boxed{p \circ g^n(x,y) = \sum_{S \subseteq [n]} \hat{p}(S) \prod_{i \in S} g(x_i, y_i)}$$

Overload notation to view $p \circ g^n$ as $\mathcal{X}^n \times \mathcal{Y}^n$ matrix with entry $(x^*, y^*) \in \mathcal{X}^n \times \mathcal{Y}^n$ containing $p \circ g^n(x^*, y^*)$

THEOREM 7 [dRMNPRV '20]

For any multilinear $p \in \mathbb{F}[\vec{z}]$ and any $g : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ with $\mathrm{rank}_{\mathbb{F}}(g) \geq 2$ it holds that

$$\boxed{\sum_{S : \hat{p}(S) \neq 0} (\mathrm{rank}_{\mathbb{F}}(g) - 2)^{|S|} \leq \mathrm{rank}_{\mathbb{F}}(p \circ g^n) \leq \sum_{S : \hat{p}(S) \neq 0} \mathrm{rank}_{\mathbb{F}}(g)^{|S|}}$$

Closely follows ideas in [PR '18] with minor but crucial twists

Lemma 5 together with the next theorem yields the main lifting theorem

THEOREM 8 [dRMNPRV'20]

Let

- F unsatisfiable k-CNF formula over n variables
- $\mathbb{F}$ any field
- $g : \mathcal{X} \times \mathcal{Y} \to \{0,1\}$ gadget with $\mathrm{rank}_{\mathbb{F}}(g) \geq 3$

Then there is a $\mathcal{X}^n \times \mathcal{Y}^n$ matrix A over $\mathbb{F}$ such that for any F-structured rectangle cover $\mathcal{R}$ of $\mathcal{X}^n \times \mathcal{Y}^n$ it holds that

$$\boxed{\mu_{\mathbb{F}}(\mathcal{R}, A) \geq \frac{1}{k}\left(\frac{\mathrm{Deg}^{\mathbb{F}}_{NS}(F \vdash \bot) \cdot \mathrm{rank}_{\mathbb{F}}(g)}{en}\right)^{\mathrm{Deg}^{\mathbb{F}}_{NS}(F \vdash \bot)} \exp\left(\frac{-4n}{\mathrm{rank}_{\mathbb{F}}(g)}\right)}$$

Proof sketch

Fix p s.t. $\mathrm{gap}_{\mathbb{F}}(F) = \mathrm{gap}_{\mathbb{F}}(F, p)$

Let $\boxed{A = p \circ g^n}$ and analyze

$$\mu_{\mathbb{F}}(\mathcal{R}, p \circ g^n) = \boxed{\frac{\mathrm{rank}_{\mathbb{F}}(p \circ g^n)}{\max\limits_{R \in \mathcal{R}}\left(\mathrm{rank}_{\mathbb{F}}(p \circ g^n \restriction R)\right)}}$$

using Thm 7.

We get numerator $\geq \left(\mathrm{rank}_{\mathbb{F}}(g) - 2\right)^n \quad (\dagger)$

For denominator, argue that for C-structured R

$$p \circ g^n \restriction R = \begin{pmatrix} M & \cdots & M \\ \vdots & \ddots & \vdots \\ M & \cdots & M \end{pmatrix}$$

for $M = (p \restriction_{\pi}) \circ g^{[n] \setminus \mathrm{Vars}(C)}$ for $\pi = \mathrm{Cert}(C)$

But then

$$\mathrm{rank}_{\mathbb{F}}(p \circ g^n|R) = \mathrm{rank}_{\mathbb{F}}\left((p|_\pi) \circ g^{[n] \setminus \mathrm{Vars}(C)}\right)$$

$$\leq \sum_{S : \widehat{p|_\pi}(S) \neq 0} \mathrm{rank}_{\mathbb{F}}(g)^{|S|}$$

and $\widehat{p|_\pi}(S) \neq 0$ only for $|S| \leq n - \mathrm{gap}_{\mathbb{F}}(F)$

We can choose p so that

$$\hat{p}(S) = 0 \qquad \text{for } |S| < n - \mathrm{gap}_{\mathbb{F}}(F)$$

since monomials of this low degree don't affect the algebraic gap

$\pi = \mathrm{Cert}(C)$ assigns $\leq k$ variables, so can bring down degree by $\leq k$

Surviving monomials in $p|_\pi$

- None of degree $> n - \mathrm{gap}_{\mathbb{F}}(F)$
- At most $\binom{n}{\mathrm{gap}_{\mathbb{F}}(F) - i}$ of degree $n - \mathrm{gap}_{\mathbb{F}}(F)$ for $i = 0, \ldots, k$
- None of degree $< n - \mathrm{gap}_{\mathbb{F}}(F) - k$

So

$$\sum_{S : \widehat{p|_\pi}(S) \neq 0} (\mathrm{rank}_{\mathbb{F}}(g))^{|S|} \leq \sum_{i=0}^{k} \binom{n}{\mathrm{gap}_{\mathbb{F}}(F) - i} \mathrm{rank}_{\mathbb{F}}(g)^{n - \mathrm{gap}_{\mathbb{F}}(F) - i}$$

$$\leq k \binom{n}{\mathrm{gap}_{\mathbb{F}}(F)} \mathrm{rank}_{\mathbb{F}}(g)^{n - \mathrm{gap}_{\mathbb{F}}(F)} \qquad (\neq)$$

Work on (†) and (‡) to get

$$\mu_F(R, p\circ g^n) \geq \frac{1}{k}\left(\frac{\mathrm{gap}_F(F)\cdot \mathrm{rank}_F(g)}{en}\right)^{\mathrm{gap}_F(F)} \exp\left(\frac{-4n}{\mathrm{rank}_F(g)}\right)$$

and then use that Thm 6 says that

$$\mathrm{gap}_F(F) = \mathrm{Deg}^F_{NS}(F \vdash \bot)$$

$\blacksquare$