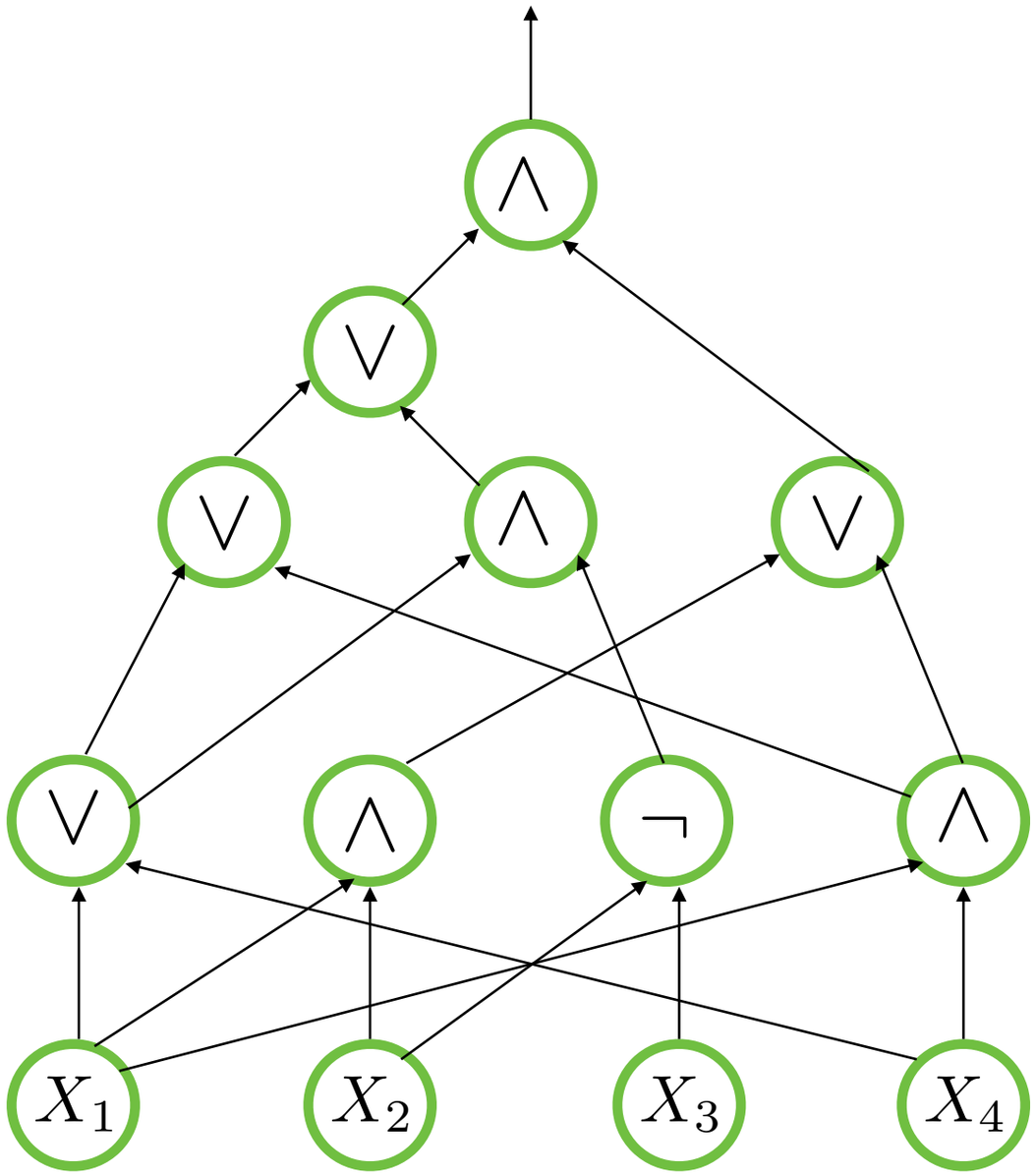
A quadratic lower bound for homogeneous algebraic branching programs

Mrinal Kumar

Speaker - Ramprasad Saptharishi

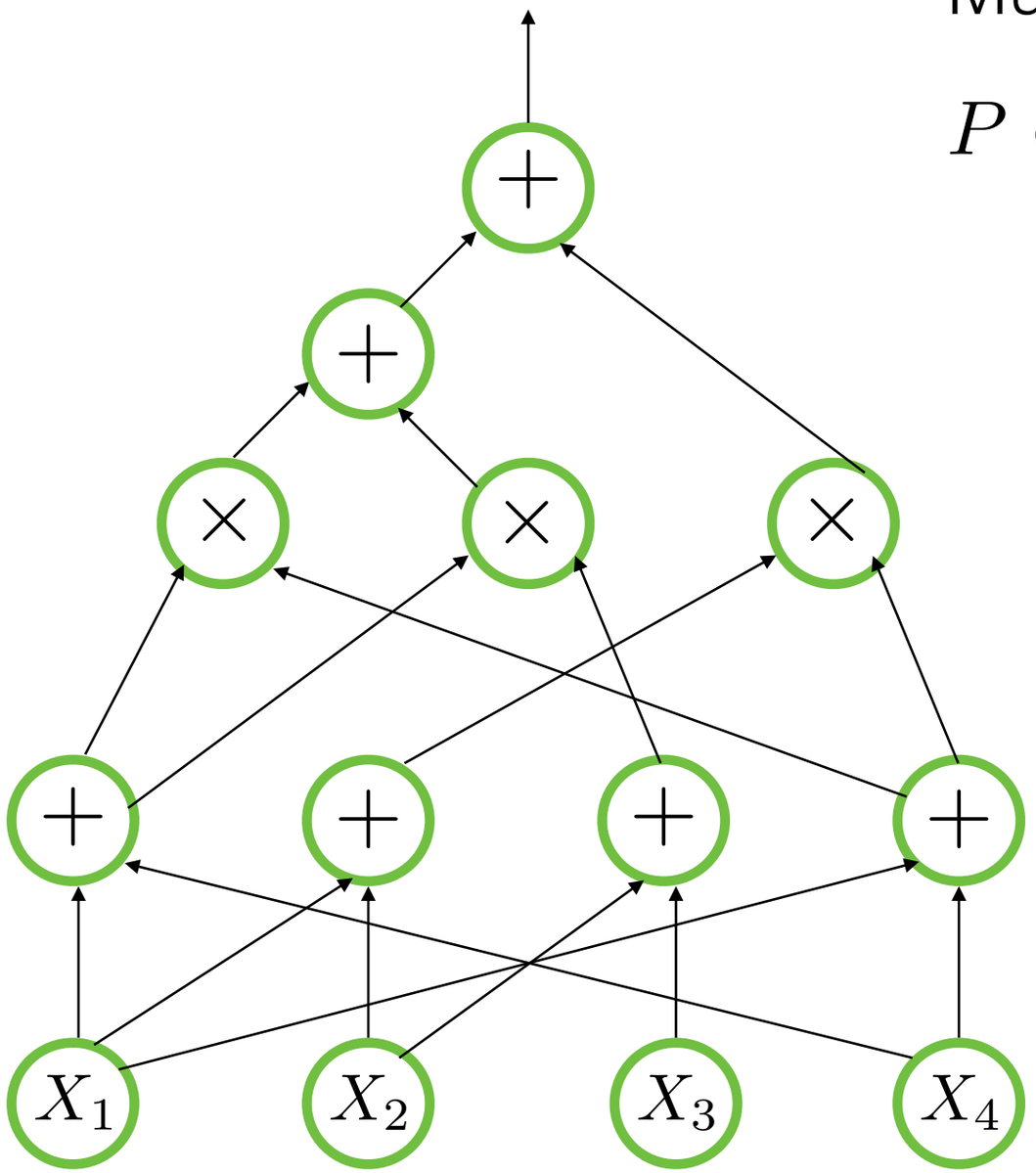
Boolean circuits



Boolean circuits

$f : \{0,1\}^n \to \{0,1\}$

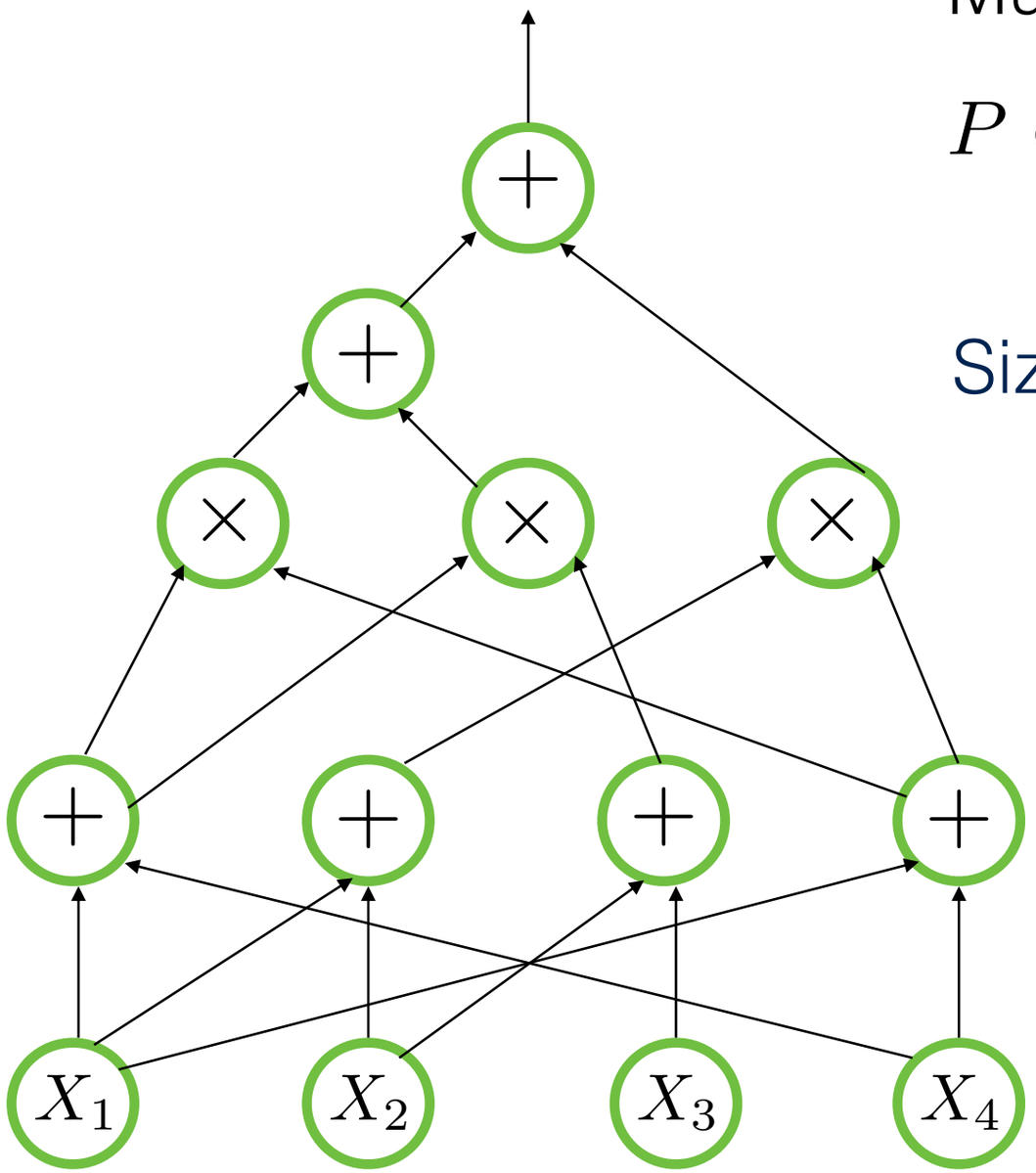
Arithmetic circuits



Multivariate polynomial

$$P \in \mathbb{F}[X_1, X_2, \ldots, X_n]$$

Arithmetic circuits

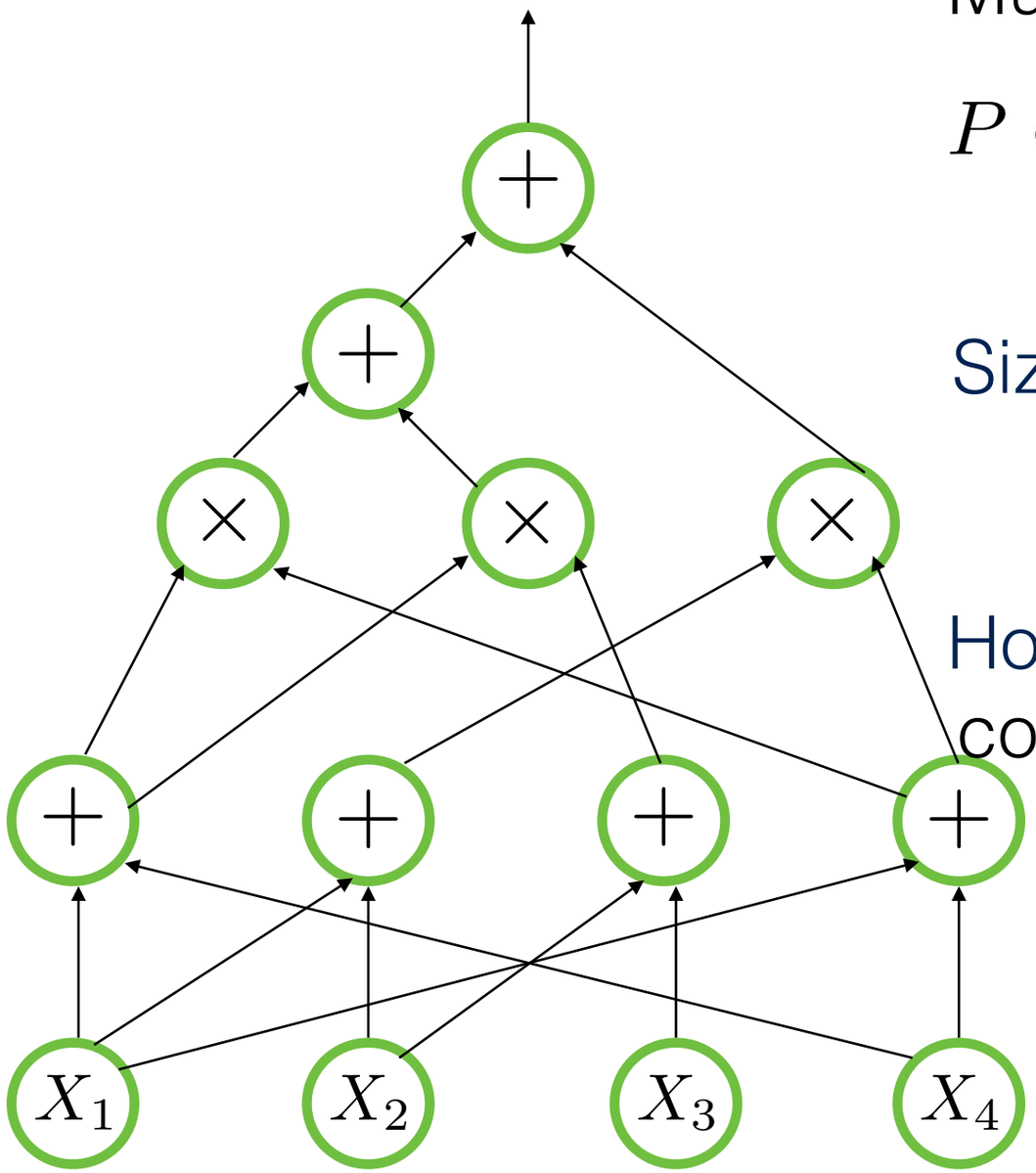


Multivariate polynomial

$$P \in \mathbb{F}[X_1, X_2, \ldots, X_n]$$

Size - number of gates

Arithmetic circuits



Multivariate polynomial

$P \in \mathbb{F}[X_1, X_2, \ldots, X_n]$

Size - number of gates

Homogeneous - every gate computes a homogeneous polynomial

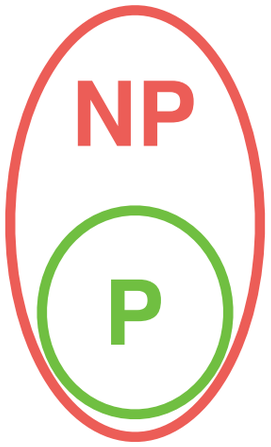
Algebraic complexity classes

- VP - Polynomials computed by poly(n) size, poly(n) degree arithmetic circuits (e.g Determinant)

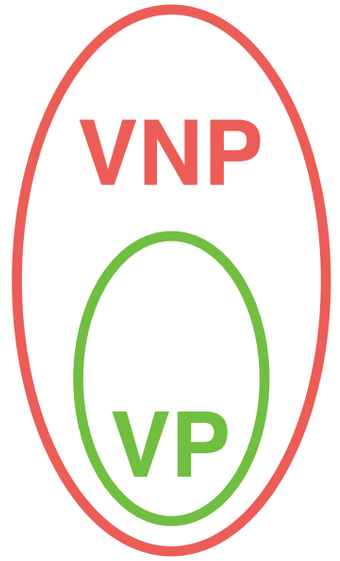
- VNP - Family of explicit polynomials (e.g Permanent)

Cook's vs Valiant's hypothesis

P vs NP

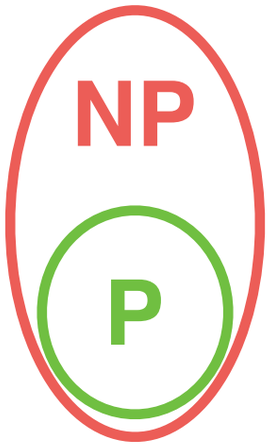


VP vs VNP

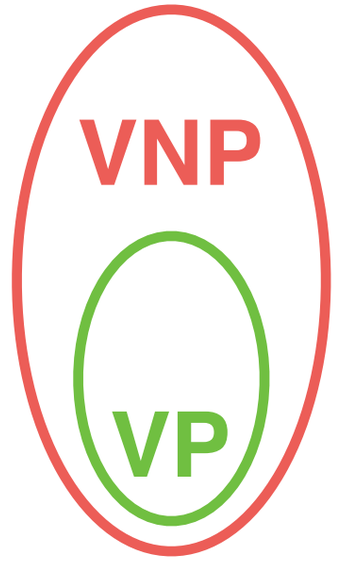


Cook's vs Valiant's hypothesis

P vs NP



VP vs VNP



[Burgisser] Under GRH, VP = VNP implies non-uniform P = non-uniform NP.

Lower bounds for arithmetic circuits

Are there explicit polynomial families of which cannot be computed by polynomial sized arithmetic circuits ?

General lower bounds

Theorem [Strassen - 73, Baur & Strassen - 83]

Any arithmetic circuit which computes the polynomial $X_1^d + X_2^d + \ldots + X_n^d$ has at least $\Omega(n \log d)$ gates.

General lower bounds

Theorem [Strassen - 73, Baur & Strassen - 83]

Any arithmetic circuit which computes the polynomial $X_1^d + X_2^d + \ldots + X_n^d$ has at least $\Omega(n \log d)$ gates.

Continues to be the best lower bound known for general circuits.

General lower bounds

Theorem [Strassen - 73, Baur & Strassen - 83]

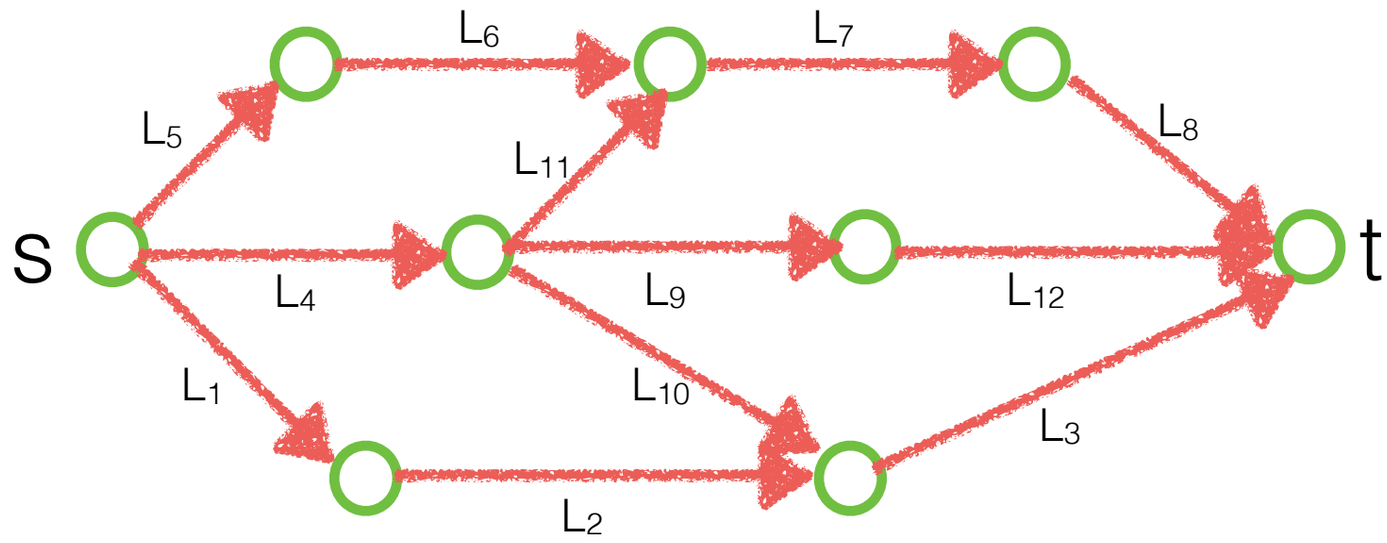
Any arithmetic circuit which computes the polynomial $X_1^d + X_2^d + \ldots + X_n^d$ has at least $\Omega(n \log d)$ gates.

Continues to be the best lower bound known for general circuits.

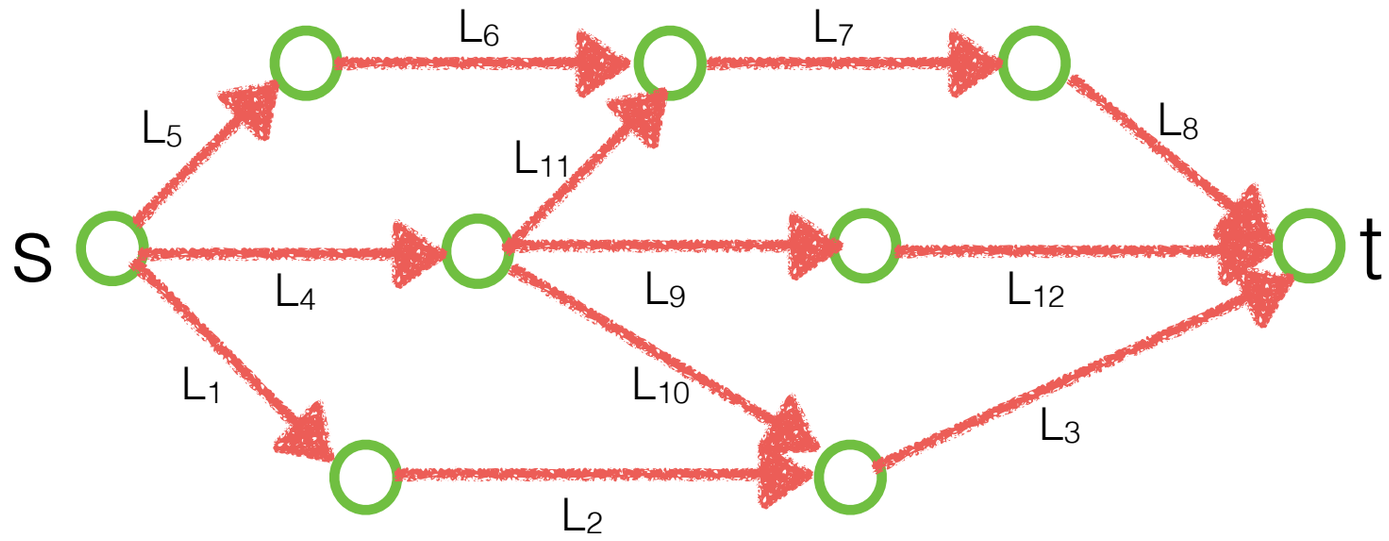
Lack of progress for this question motivates an interest in weaker models - arithmetic formula, bounded depth circuits etc.

Algebraic branching programs (ABP)

Algebraic branching programs

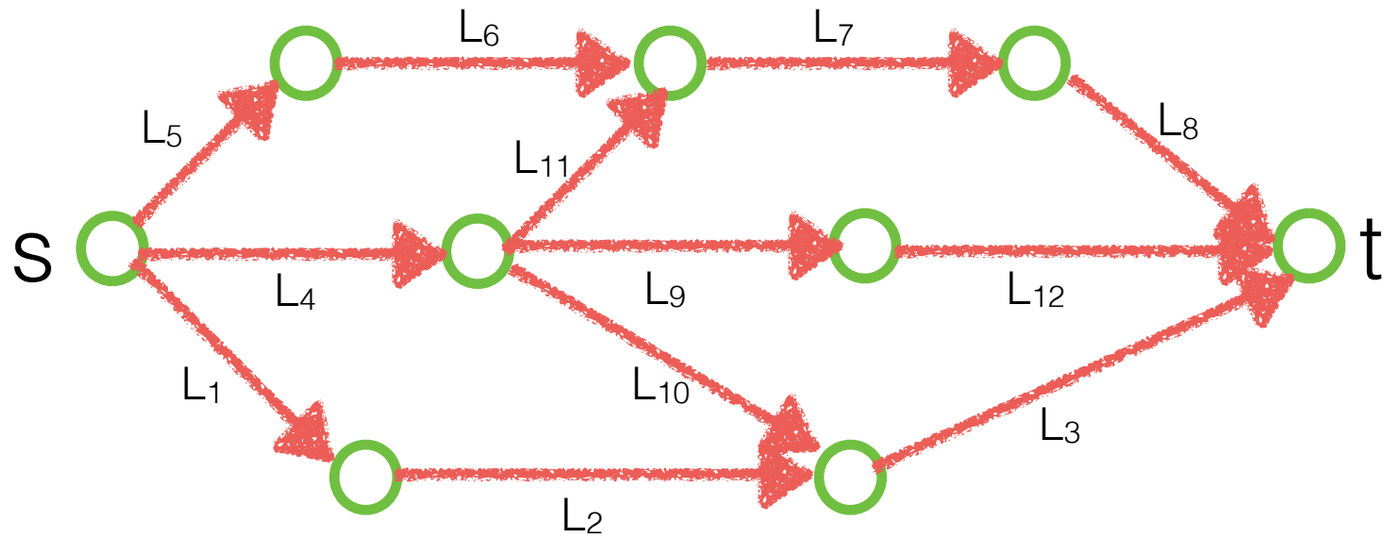


Algebraic branching programs



Each edge weight is an affine form

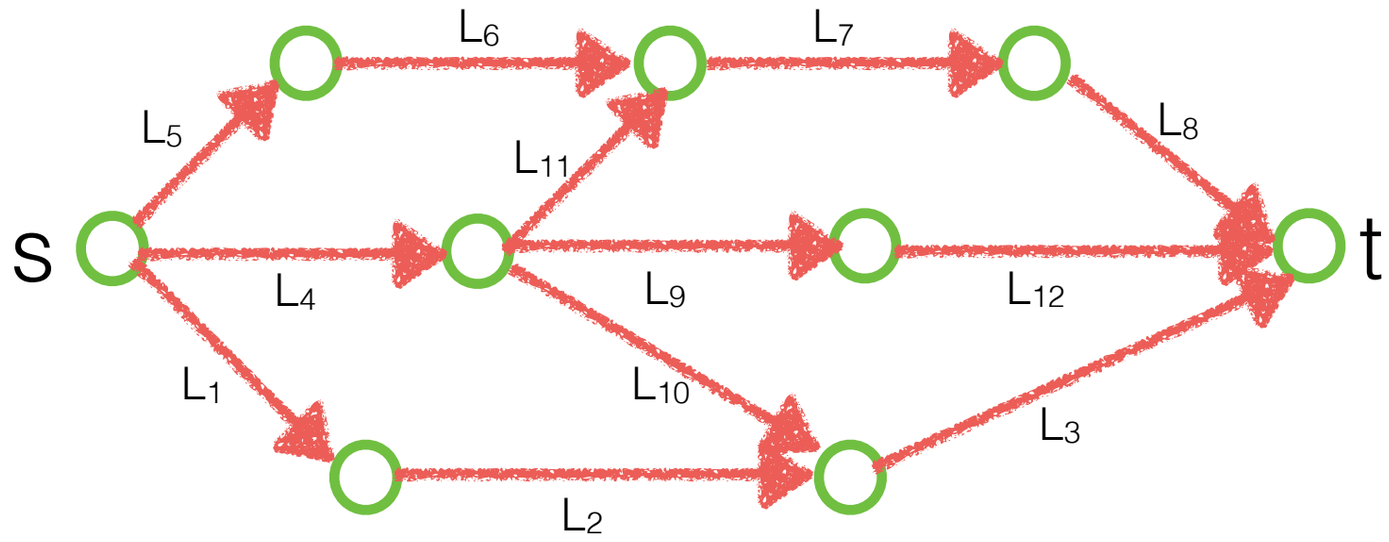
Algebraic branching programs



Each edge weight is an affine form

Weight of a path = product of edge weights in the path

Algebraic branching programs

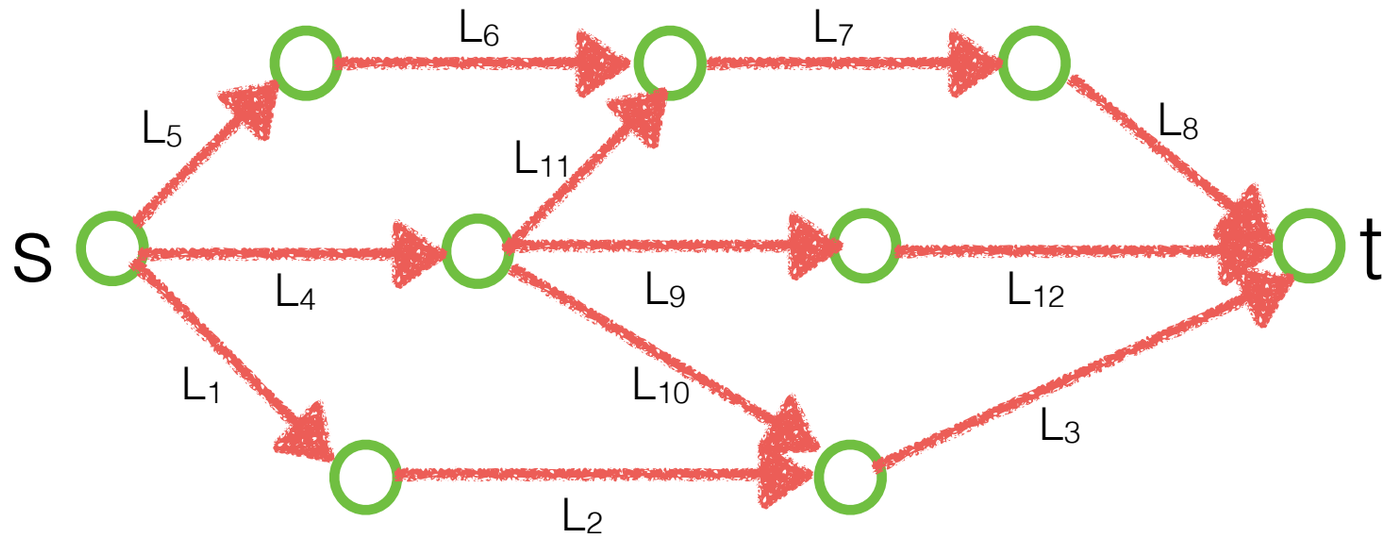


Each edge weight is an affine form

Weight of a path = product of edge weights in the path

Polynomial computed = sum of weights of all s-t paths

Algebraic branching programs



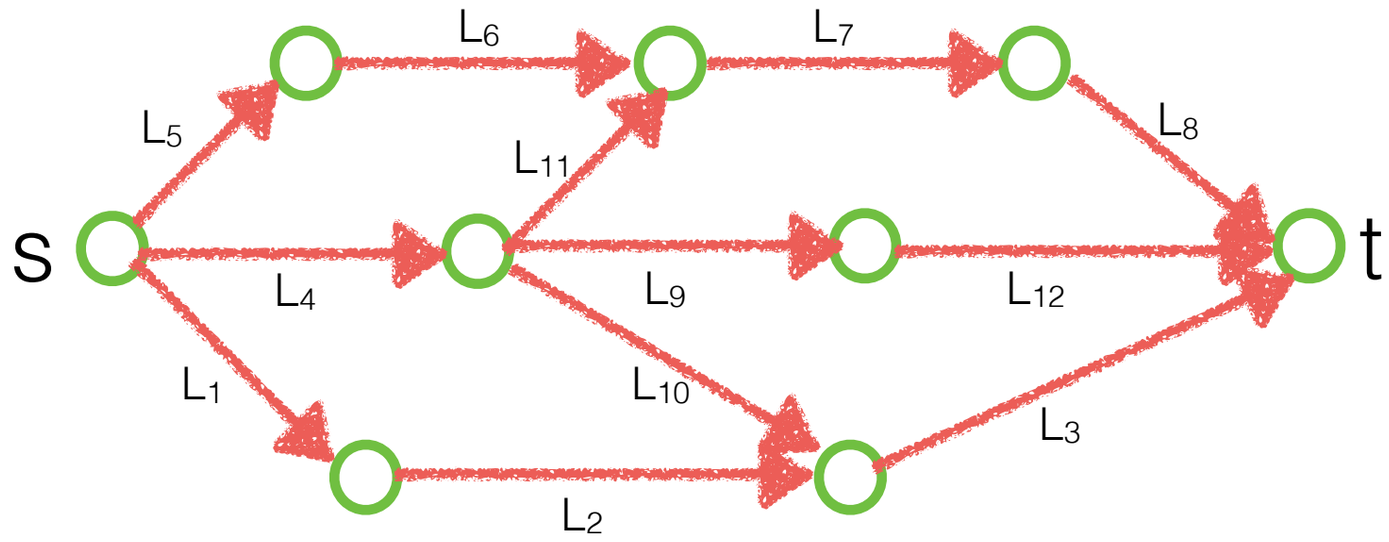
Each edge weight is an affine form

Weight of a path = product of edge weights in the path

Polynomial computed = sum of weights of all s-t paths

Formulas $\leq$ ABPs $\leq$ Circuits

Algebraic branching programs



Each edge weight is an affine form

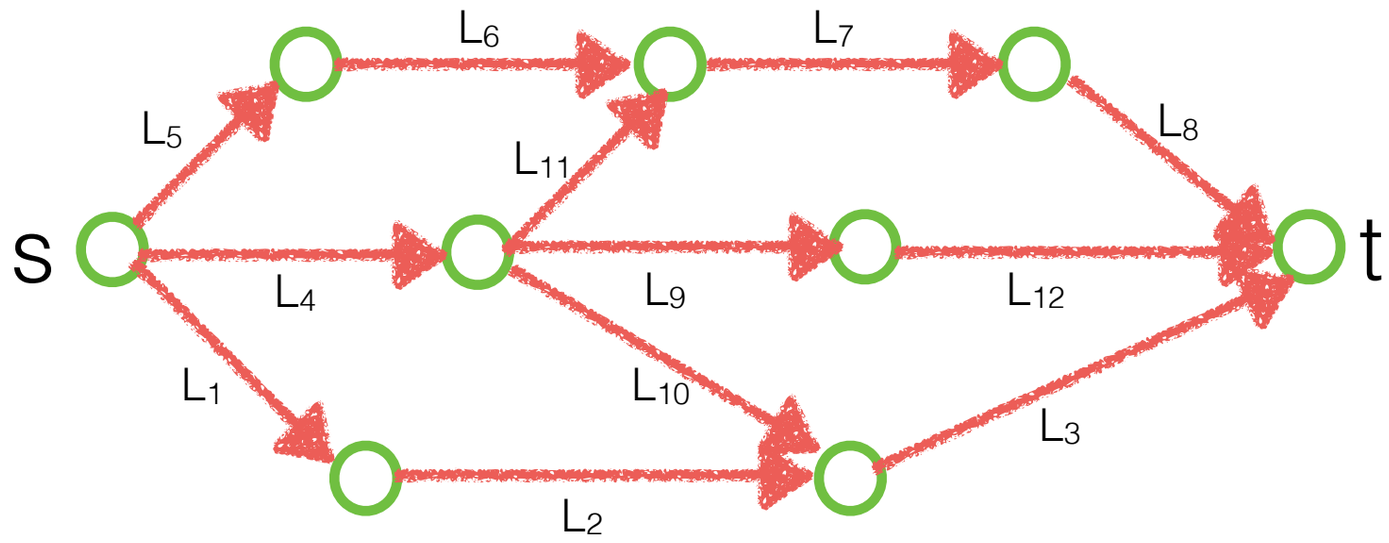
Weight of a path = product of edge weights in the path

Polynomial computed = sum of weights of all s-t paths

Formulas $\leq$ ABPs $\leq$ Circuits

So, proving better ABP and formula lower bounds could be an easier task.

Algebraic branching programs



Each edge weight is an affine form

Weight of a path = product of edge weights in the path

Polynomial computed = sum of weights of all s-t paths

Formulas $\leq$ ABPs $\leq$ Circuits

So, proving better ABP and formula lower bounds could be an easier task.

And we do know nearly quadratic lower bounds for arithmetic formula!

Formula lower bounds

Theorem [Kalorkoti - 85]

Any arithmetic formula which computes the determinant of an n dimensional symbolic matrix has at least $\Omega(n^3)$ leaves.

Formula lower bounds

Theorem [Kalorkoti - 85]

Any arithmetic formula which computes the determinant of an n dimensional symbolic matrix has at least $\Omega(n^3)$ leaves.

Shpilka-Yehudayoff subsequently observed that this yields nearly quadratic lower bounds for another explicit polynomial.

Formula lower bounds

Theorem [Kalorkoti - 85]

Any arithmetic formula which computes the determinant of an n dimensional symbolic matrix has at least $\Omega(n^3)$ leaves.

Shpilka-Yehudayoff subsequently observed that this yields nearly quadratic lower bounds for another explicit polynomial.

However, the lower bounds for ABPs continued to be just $\Omega(n \log d)$.

Results

Homogeneous ABP lower bounds

Theorem

Any homogeneous ABP which computes the polynomial $X_1^d + X_2^d + \ldots + X_n^d$ has at least $\Omega(nd)$ vertices.

Homogeneous ABP lower bounds

Theorem

Any homogeneous ABP which computes the polynomial $X_1^d + X_2^d + \ldots + X_n^d$ has at least $\Omega(nd)$ vertices.

For d = n (which is in the 'right' range of parameters), we get a quadratic lower bound on the number of vertices.

Homogeneous ABP lower bounds

Theorem

Any homogeneous ABP which computes the polynomial $X_1^d + X_2^d + \ldots + X_n^d$ has at least $\Omega(nd)$ vertices.

For d = n (which is in the ‘right’ range of parameters), we get a quadratic lower bound on the number of vertices.

It is not clear if anything super linear was known prior to this on the number of vertices.

Homogeneous ABP lower bounds

Theorem

Any homogeneous ABP which computes the polynomial $X_1^d + X_2^d + \ldots + X_n^d$ has at least $\Omega(nd)$ vertices.

For d = n (which is in the ‘right’ range of parameters), we get a quadratic lower bound on the number of vertices.

It is not clear if anything super linear was known prior to this on the number of vertices.

Baur-Strassen seems to imply a super linear lower bound only on the number of edges.

Also gives new proofs for….

Theorem [Str 73, BS 83]

Any homogeneous arithmetic circuit which computes the polynomial $X_1^d + X_2^d + \ldots + X_n^d$ has at least $\Omega(n \log d)$ gates.

Also gives new proofs for….

Theorem [Str 73, BS 83]

Any homogeneous arithmetic circuit which computes the polynomial $X_1^d + X_2^d + \ldots + X_n^d$ has at least $\Omega(n \log d)$ gates.

Theorem [Mignon-Ressayre 04, Yabe 16]

The determinantal complexity of $X_1^d + X_2^d + \ldots + X_n^d$ is at least $n/2$.(n over reals)

Also gives new proofs for….

Theorem [Str 73, BS 83]

Any homogeneous arithmetic circuit which computes the polynomial $X_1^d + X_2^d + \ldots + X_n^d$ has at least $\Omega(n \log d)$ gates.

Theorem [Mignon-Ressayre 04, Yabe 16]

The determinantal complexity of $X_1^d + X_2^d + \ldots + X_n^d$ is at least $n/2$.(n over reals)

The original proofs of the above two lower bounds seemed quite different.

Also gives new proofs for….

Theorem [Str 73, BS 83]

Any homogeneous arithmetic circuit which computes the polynomial $X_1^d + X_2^d + \ldots + X_n^d$ has at least $\Omega(n \log d)$ gates.

Theorem [Mignon-Ressayre 04, Yabe 16]

The determinantal complexity of $X_1^d + X_2^d + \ldots + X_n^d$ is at least $n/2$.(n over reals)

The original proofs of the above two lower bounds seemed quite different.

Our proofs are short, intuitive and essentially the same.

Proof sketches

Overview of the proofs

Two simple steps :

Overview of the proofs

Two simple steps :

- There are ‘many’ disjoint s-t vertex cuts in a homogeneous ABP.

Overview of the proofs

Two simple steps :

- There are ‘many’ disjoint s-t vertex cuts in a homogeneous ABP.
- Each of the cuts has ‘many’ vertices if the ABP computes $P_{n,d} = X_1^d + X_2^d + \ldots + X_n^d$.

Overview of the proofs

Two simple steps :

- There are ‘many’ disjoint s-t vertex cuts in a homogeneous ABP.
- Each of the cuts has ‘many’ vertices if the ABP computes $P_{n,d} = X_1^d + X_2^d + \ldots + X_n^d$.

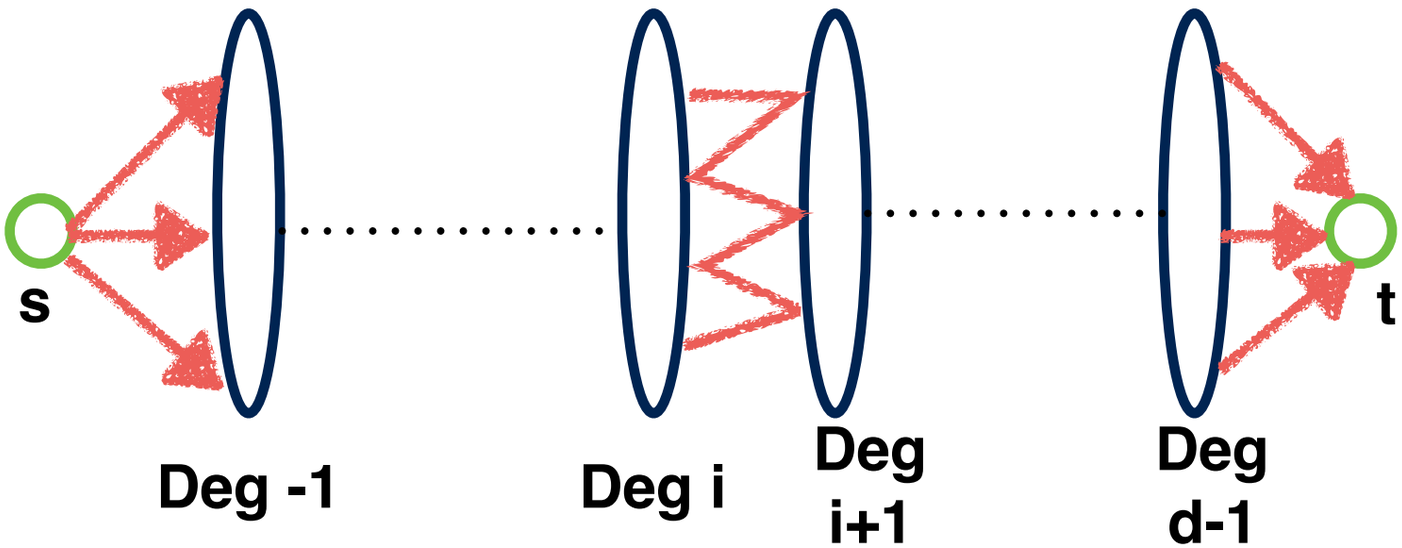
We work over complex numbers for rest of the talk.

Step 1 : Many disjoint s-t vertex cuts

A homogeneous ABP can be assumed to look layered.

Step 1 : Many disjoint s-t vertex cuts

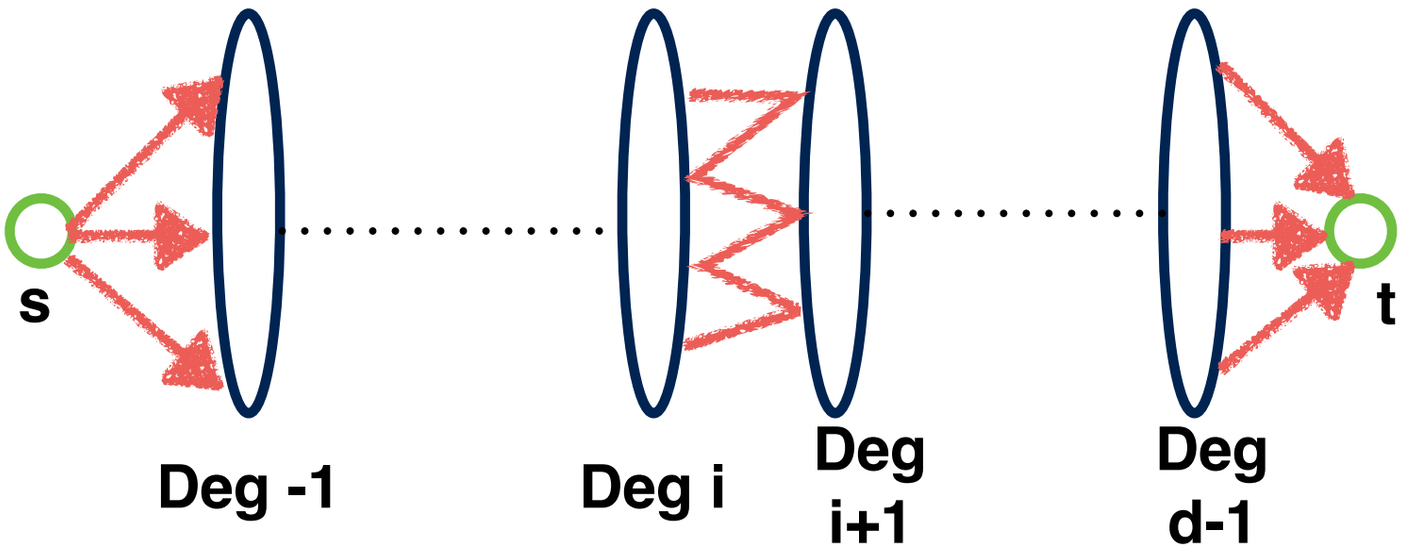
A homogeneous ABP can be assumed to look layered.



Step 1 : Many disjoint s-t vertex cuts

A homogeneous ABP can be assumed to look layered.

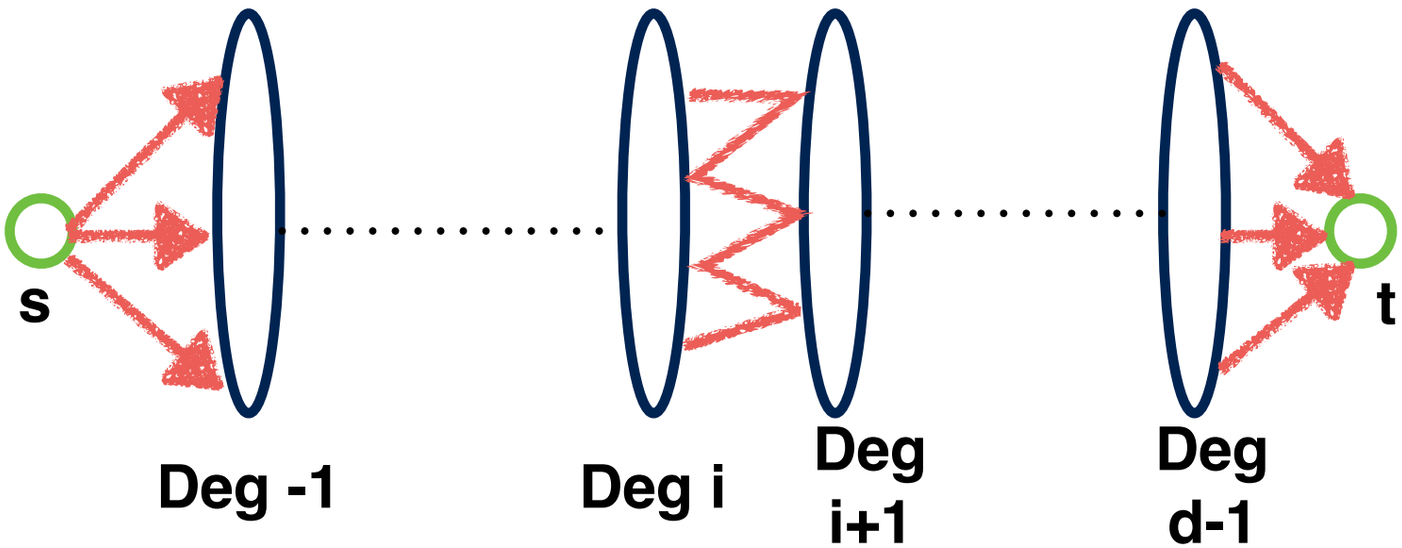
Layer i = vertices of degree i.



Step 1 : Many disjoint s-t vertex cuts

A homogeneous ABP can be assumed to look layered.

Layer i = vertices of degree i.

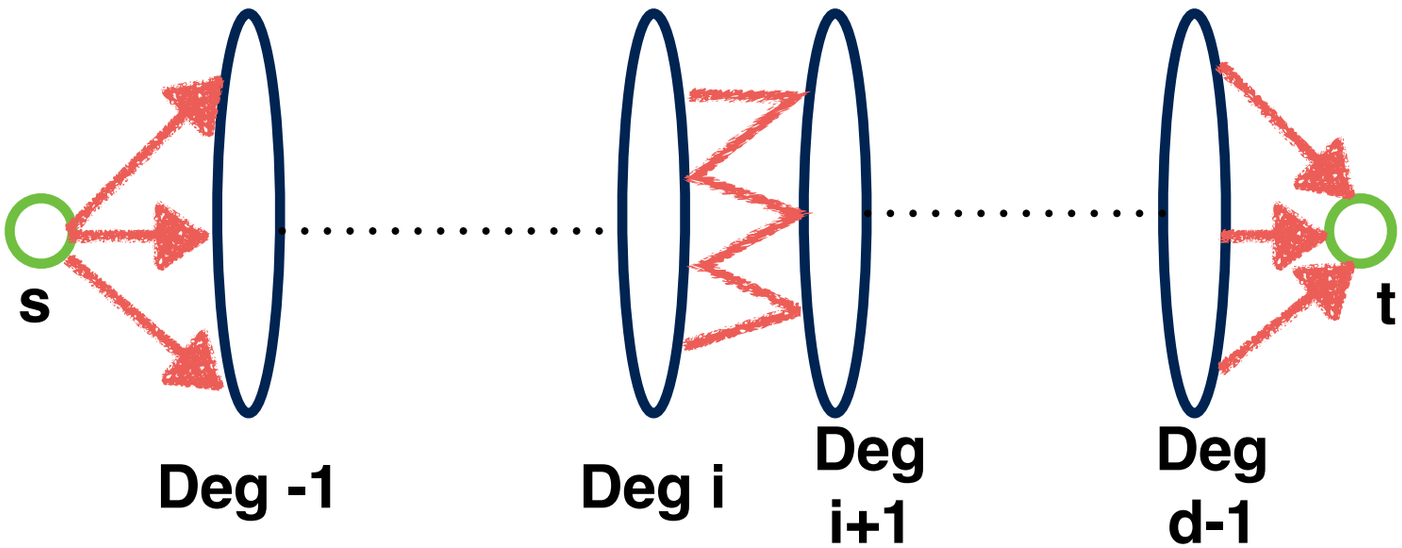


Edges only between successive layers.

Step 1 : Many disjoint s-t vertex cuts

A homogeneous ABP can be assumed to look layered.

Layer i = vertices of degree i.

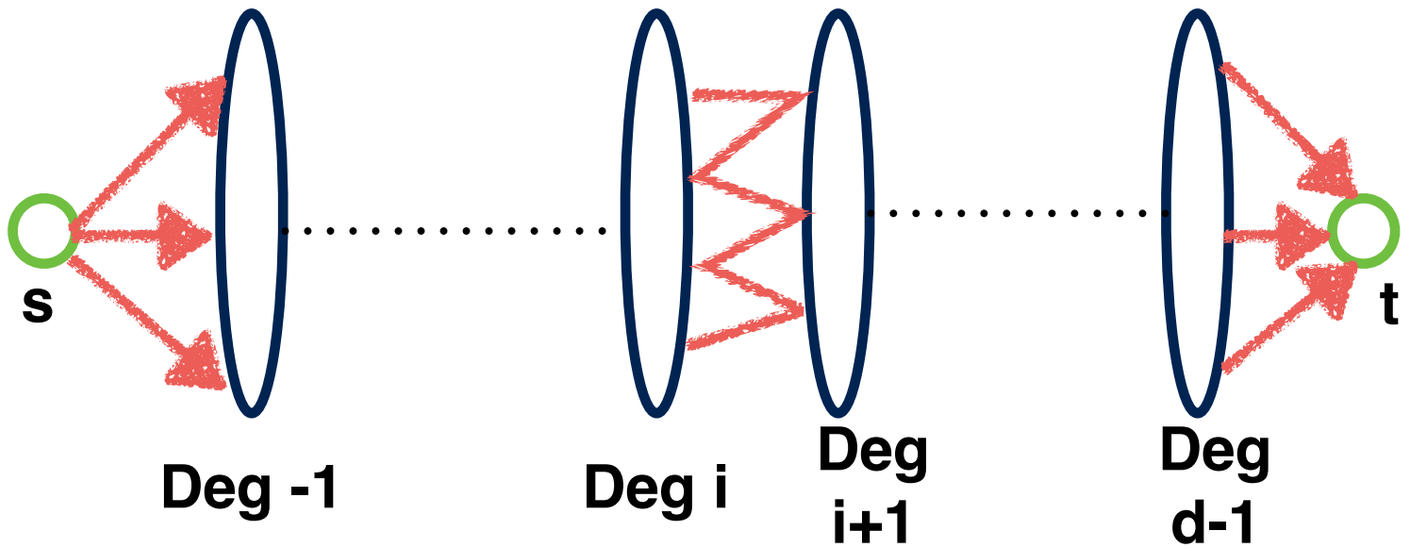


Edges only between successive layers.

So, each layer in a s-t vertex cut.

Step 1 : Many disjoint s-t vertex cuts

A homogeneous ABP can be assumed to look layered.



Layer i = vertices of degree i.

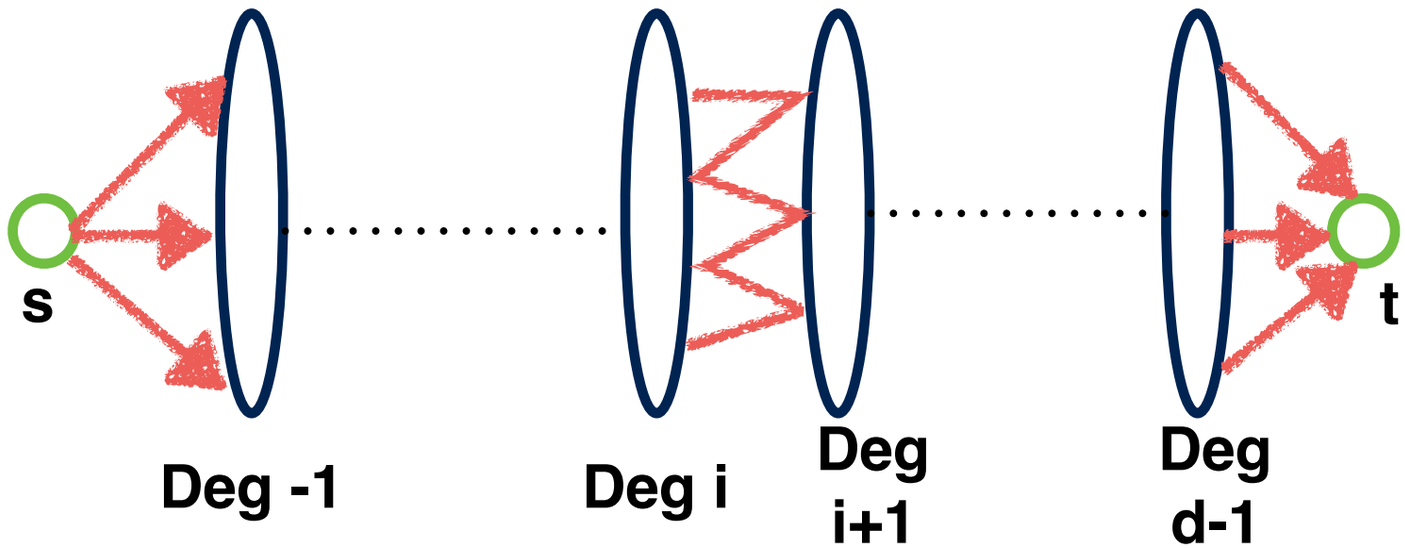
Edges only between successive layers.

So, each layer in a s-t vertex cut.

Since each edge label is affine, there are at least d-1 layers.

Step 1 : Many disjoint s-t vertex cuts

A homogeneous ABP can be assumed to look layered.



Layer i = vertices of degree i.

Edges only between successive layers.

So, each layer in a s-t vertex cut.

Since each edge label is affine, there are at least d-1 layers.

So, d-1 disjoint s-t vertex cuts.

Step 2 : Each s-t cut must be large

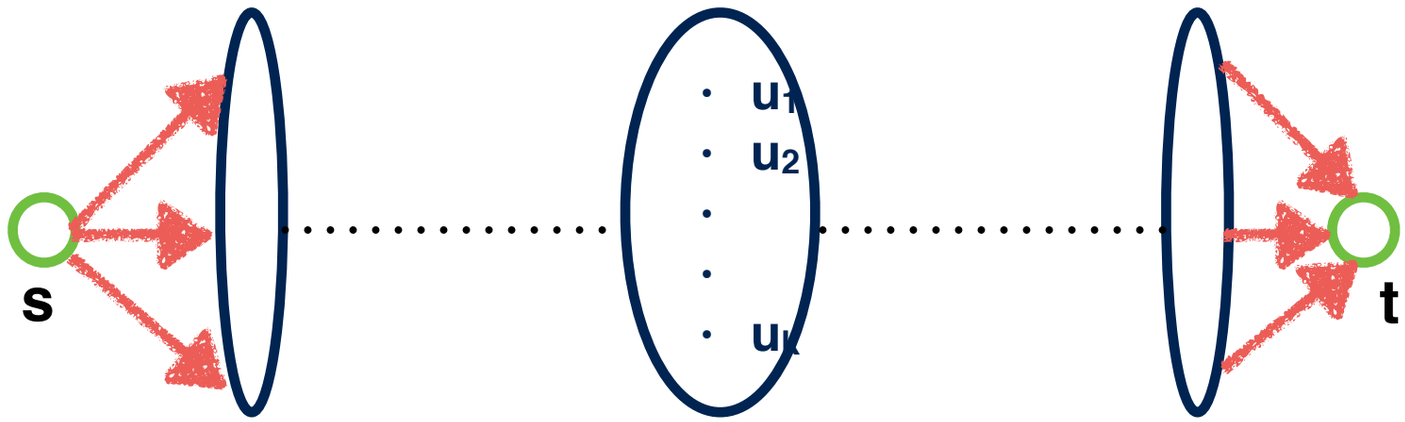
Lemma : Each s-t vertex cut in a homogeneous ABP which computes the $X_1^d + X_2^d + \ldots + X_n^d$ must have at least $n/2$ vertices.

Proof : Let $u_1, u_2, \ldots, u_k$ be the vertices in the cut.

Step 2 : Each s-t cut must be large

Lemma : Each s-t vertex cut in a homogeneous ABP which computes the $X_1^d + X_2^d + \ldots + X_n^d$ must have at least $n/2$ vertices.

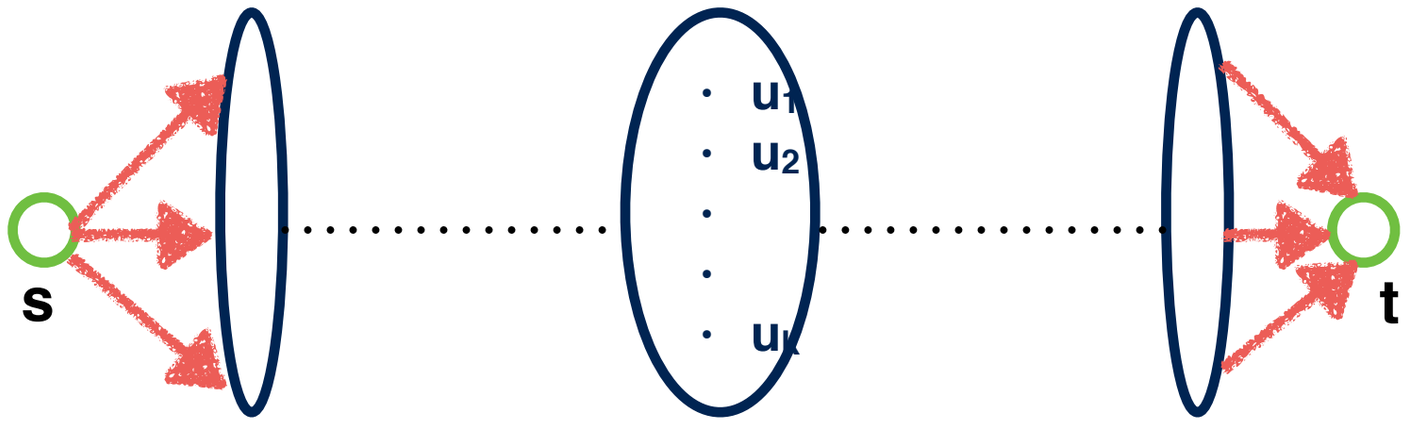
Proof : Let $u_1, u_2, \ldots, u_k$ be the vertices in the cut.



Step 2 : Each s-t cut must be large

Lemma : Each s-t vertex cut in a homogeneous ABP which computes the $X_1^d + X_2^d + \ldots + X_n^d$ must have at least $n/2$ vertices.

Proof : Let $u_1, u_2, \ldots, u_k$ be the vertices in the cut.



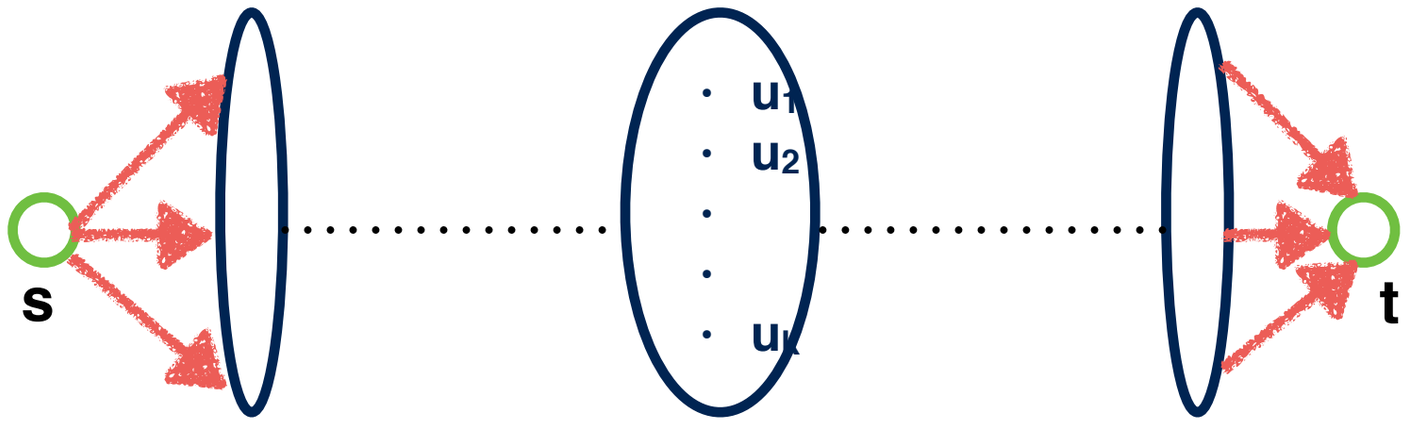
Then, there are homogeneous polynomials $v_1, v_2, \ldots, v_k$ such that

$$X_1^d + X_2^d + \ldots + X_n^d = \sum_{i=1}^{k} u_i v_i$$

Step 2 : Each s-t cut must be large

Lemma : Each s-t vertex cut in a homogeneous ABP which computes the $X_1^d + X_2^d + \ldots + X_n^d$ must have at least $n/2$ vertices.

Proof : Let $u_1, u_2, \ldots, u_k$ be the vertices in the cut.



Then, there are homogeneous polynomials $v_1, v_2, \ldots, v_k$ such that

$$X_1^d + X_2^d + \ldots + X_n^d = \sum_{i=1}^{k} u_i v_i$$

If every such identity implied many summands, then we would be done.

Step 2 : Each s-t cut must be large

Claim [Kayal] : For homogeneous polynomials u_1 , u_2, …., u_k and v_1 , v_2, …., v_k , if

$$X_1^d + X_2^d + \ldots + X_n^d = \sum_{i=1}^{k} u_i v_i$$

Then, $k \geq n/2$.

Step 2 : Each s-t cut must be large

Claim [Kayal] : For homogeneous polynomials u_1 , u_2, …., u_k and v_1 , v_2, …., v_k , if

$$X_1^d + X_2^d + \ldots + X_n^d = \sum_{i=1}^{k} u_i v_i$$

Then, $k \geq n/2$.

Proof : Consider the set of common zeroes of u_1 , u_2, …., u_k and v_1 , v_2, …., v_k.

Step 2 : Each s-t cut must be large

Claim [Kayal] : For homogeneous polynomials u_1 , u_2, …., u_k and v_1 , v_2, …., v_k , if

$$X_1^d + X_2^d + \ldots + X_n^d = \sum_{i=1}^{k} u_i v_i$$

Then, $k \geq n/2$.

Proof : Consider the set of common zeroes of u_1 , u_2, …., u_k and v_1 , v_2, …., v_k.
This is a non-empty variety (homogeneity), of dimension at least $n - 2k$.

Step 2 : Each s-t cut must be large

Claim [Kayal] : For homogeneous polynomials u_1 , u_2, …., u_k and v_1 , v_2, …., v_k , if

$$X_1^d + X_2^d + \ldots + X_n^d = \sum_{i=1}^{k} u_i v_i$$

Then, $k \geq n/2$.

Proof : Consider the set of common zeroes of u_1 , u_2, …., u_k and v_1 , v_2, …., v_k.
This is a non-empty variety (homogeneity), of dimension at least $n - 2k$.
So, the dimension of zeroes of multiplicity at least two of $P_{n,d}$ is at least $n - 2k$.

Step 2 : Each s-t cut must be large

Claim [Kayal] : For homogeneous polynomials u_1 , u_2, …., u_k and v_1 , v_2, …., v_k , if

$$X_1^d + X_2^d + \ldots + X_n^d = \sum_{i=1}^{k} u_i v_i$$

Then, $k \geq n/2$.

Proof : Consider the set of common zeroes of u_1 , u_2, …., u_k and v_1 , v_2, …., v_k.
This is a non-empty variety (homogeneity), of dimension at least $n - 2k$.
So, the dimension of zeroes of multiplicity at least two of $P_{n,d}$ is at least $n - 2k$.

But, the first order derivatives of $P_{n,d}$ are $X_1^{d-1}, X_2^{d-1}, \ldots, X_n^{d-1}$ whose only common zero is the origin. So, $n - 2k \leq 0$.

- For homogeneous circuits, the proof is essentially the same, but there are about log d disjoint cuts.

- For homogeneous circuits, the proof is essentially the same, but there are about log d disjoint cuts.

- So, the lower bound obtained is n.log d.

- For homogeneous circuits, the proof is essentially the same, but there are about log d disjoint cuts.

- So, the lower bound obtained is n.log d.

- Homogeneity is not essential to the proofs. Formal degree at most d suffices.

- For homogeneous circuits, the proof is essentially the same, but there are about log d disjoint cuts.

- So, the lower bound obtained is n.log d.

- Homogeneity is not essential to the proofs. Formal degree at most d suffices.

- We have an appropriate generalization of Kayal's lemma.

Open problems

- Can we get rid of the homogeneity restriction and prove similar lower bounds for ABPs of arbitrary formal degree?

Open problems

- Can we get rid of the homogeneity restriction and prove similar lower bounds for ABPs of arbitrary formal degree?
    - Not clear if there are a lot of disjoint cuts.

Open problems

- Can we get rid of the homogeneity restriction and prove similar lower bounds for ABPs of arbitrary formal degree?
    - Not clear if there are a lot of disjoint cuts.
- Prove super-quadratic lower bounds for homogeneous ABPs?

Open problems

- Can we get rid of the homogeneity restriction and prove similar lower bounds for ABPs of arbitrary formal degree?
    - Not clear if there are a lot of disjoint cuts.
- Prove super-quadratic lower bounds for homogeneous ABPs?
    - Perhaps, we could start with trying to prove improved lower bounds for homogeneous formulas.

Thanks!