Average-Case Hardness in Proof Complexity
(a biased survey)

Susanna F. de Rezende

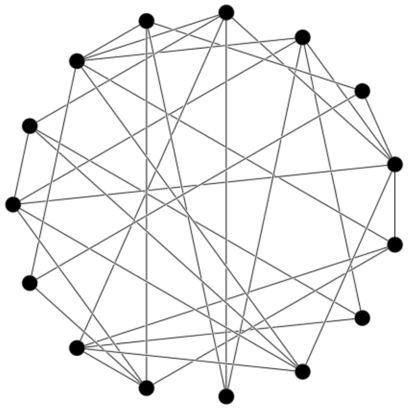
Institute of Mathematics of the
Czech Academy of Sciences

September 16, 2021

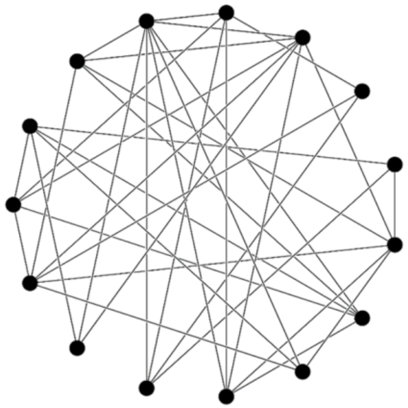
Planted clique problem

Erdős-Rényi random graph $G \sim \mathcal{G}(n, 1/2)$

w.h.p. largest clique has size $\omega(G) \approx 2 \log n$

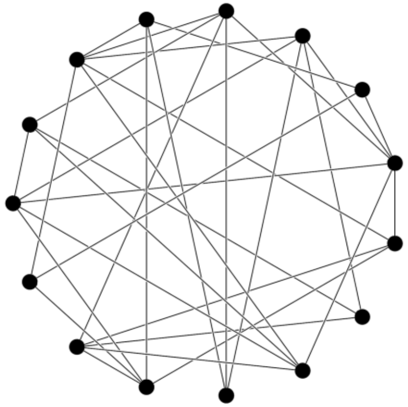


$G' = G + K_k$ where $G \sim \mathcal{G}(n, 1/2)$ and $k > 2 \log n$

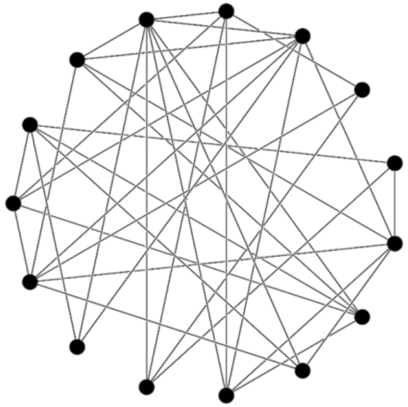


Planted clique problem

Erdős-Rényi random graph $G \sim \mathcal{G}(n, 1/2)$
w.h.p. largest clique has size $\omega(G) \approx 2 \log n$



$G' = G + K_k$ where $G \sim \mathcal{G}(n, 1/2)$ and $k > 2 \log n$

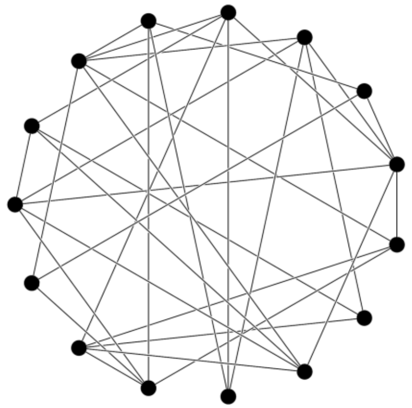


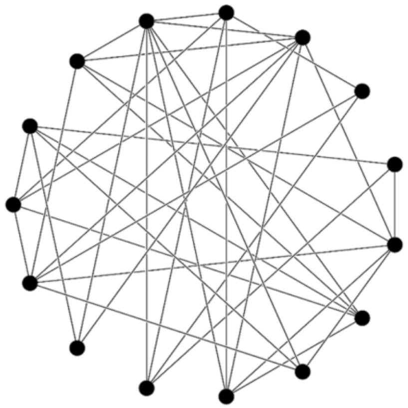
- Brute-force algorithm proves G has no clique of size $k > \omega(G)$ in time $n^{O(\omega(G))}$

Planted clique problem

Erdős-Rényi random graph $G \sim \mathcal{G}(n, 1/2)$
w.h.p. largest clique has size $\omega(G) \approx 2 \log n$

$G' = G + K_k$ where $G \sim \mathcal{G}(n, 1/2)$ and $k > 2 \log n$



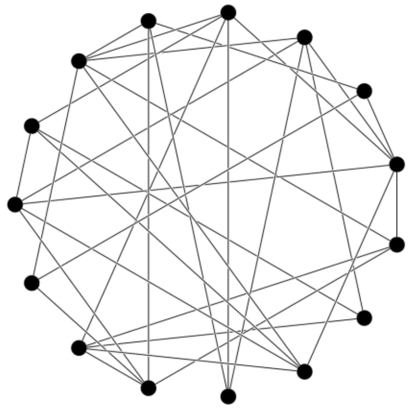


- Brute-force algorithm proves G has no clique of size $k > \omega(G)$ in time $n^{O(\omega(G))}$
- $\exists$ algorithm that distinguishes both distributions in poly-time?

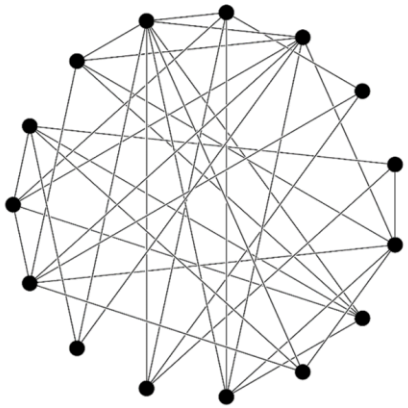
Planted clique problem

Erdős-Rényi random graph $G \sim \mathcal{G}(n, 1/2)$

w.h.p. largest clique has size $\omega(G) \approx 2\log n$



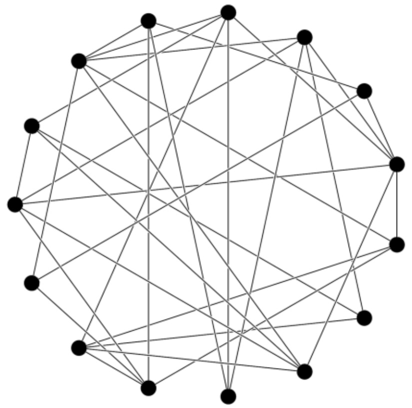
$G' = G + K_k$ where $G \sim \mathcal{G}(n, 1/2)$ and $k > 2\log n$



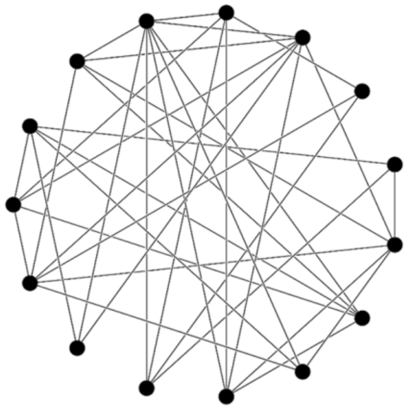
- Brute-force algorithm proves G has no clique of size $k > \omega(G)$ in time $n^{O(\omega(G))}$
- $\exists$ algorithm that distinguishes both distributions in poly-time?
- $\exists$ algorithm that proves G has no k-clique for $\omega(G) < k < \sqrt{n}$ in poly-time?

Planted clique problem

Erdős-Rényi random graph $G \sim \mathcal{G}(n, 1/2)$
w.h.p. largest clique has size $\omega(G) \approx 2 \log n$



$G' = G + K_k$ where $G \sim \mathcal{G}(n, 1/2)$ and $k > 2 \log n$

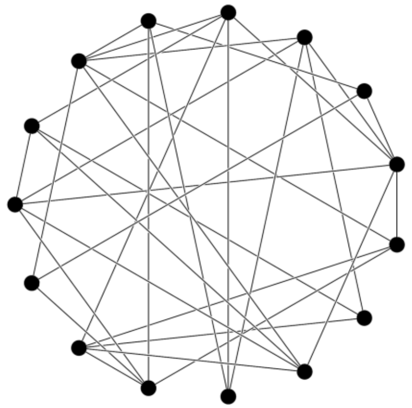


- Brute-force algorithm proves G has no clique of size $k > \omega(G)$ in time $n^{O(\omega(G))}$
- $\exists$ algorithm that distinguishes both distributions in poly-time?
- $\exists$ algorithm that proves G has no k-clique for $\omega(G) < k < \sqrt{n}$ in poly-time?
- Many results for planted clique actually prove lower bound for refutation problem

Planted clique problem

Erdős-Rényi random graph $G \sim \mathcal{G}(n, 1/2)$

w.h.p. largest clique has size $\omega(G) \approx 2 \log n$

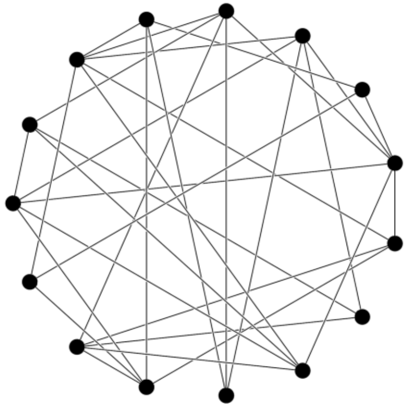


- Brute-force algorithm proves G has no clique of size $k > \omega(G)$ in time $n^{O(\omega(G))}$
- $\exists$ algorithm that distinguishes both distributions in poly-time?
- $\exists$ algorithm that proves G has no k-clique for $\omega(G) < k < \sqrt{n}$ in poly-time?
- Many results for planted clique actually prove lower bound for refutation problem

Clique problem: prove G has no k-clique

Erdős-Rényi random graph $G \sim \mathcal{G}(n, 1/2)$

w.h.p. largest clique has size $\omega(G) \approx 2 \log n$

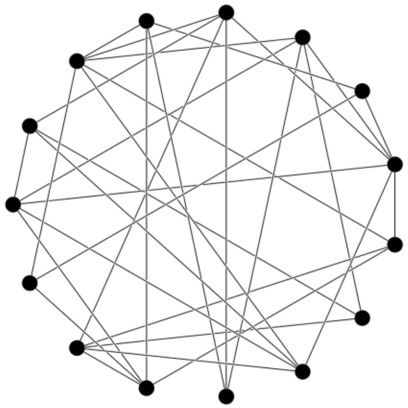


- Brute-force algorithm proves G has no clique of size $k > \omega(G)$ in time $n^{O(\omega(G))}$
- $\exists$ algorithm that proves G has no k-clique for $\omega(G) < k < \sqrt{n}$ in poly-time?

Clique problem: prove G has no k-clique

Erdős-Rényi random graph $G \sim \mathcal{G}(n, 1/2)$

w.h.p. largest clique has size $\omega(G) \approx 2 \log n$

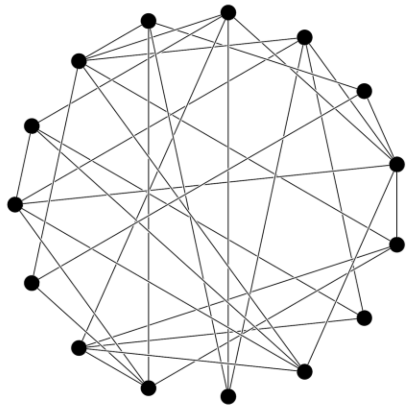


- Brute-force algorithm proves G has no clique of size $k > \omega(G)$ in time $n^{O(\omega(G))}$
- $\exists$ algorithm that proves G has no k-clique for $\omega(G) < k < \sqrt{n}$ in poly-time?
- $\exists$ proof that G has no k-clique for $\omega(G) < k < \sqrt{n}$ in poly-size?

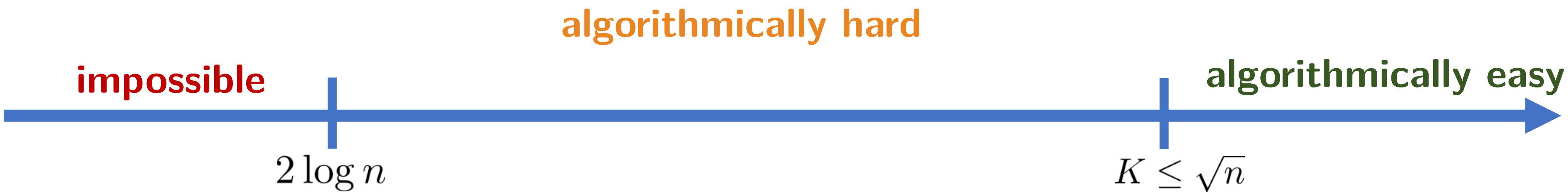
Clique problem: prove G has no k-clique

Erdős-Rényi random graph $G \sim \mathcal{G}(n, 1/2)$

w.h.p. largest clique has size $\omega(G) \approx 2\log n$



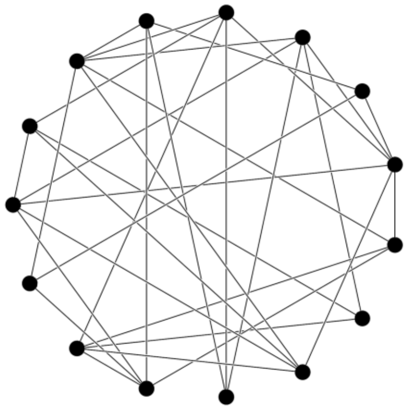
- Brute-force algorithm proves G has no clique of size $k > \omega(G)$ in time $n^{O(\omega(G))}$
- $\exists$ algorithm that proves G has no k-clique for $\omega(G) < k < \sqrt{n}$ in poly-time?
- $\exists$ proof that G has no k-clique for $\omega(G) < k < \sqrt{n}$ in poly-size?



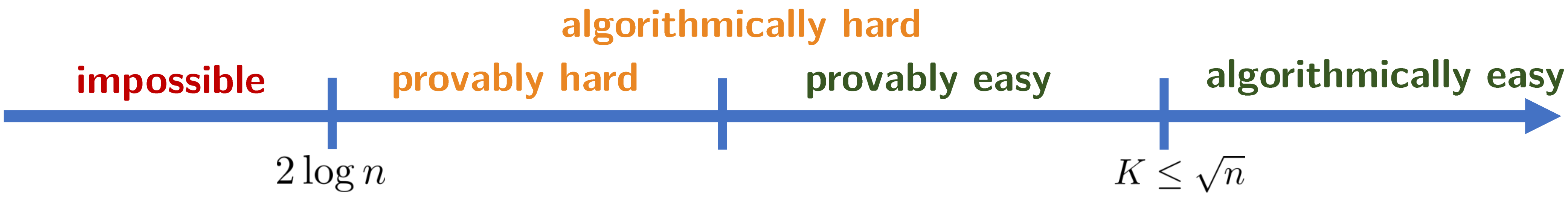
Clique problem: prove G has no k-clique

Erdős-Rényi random graph $G \sim \mathcal{G}(n, 1/2)$

w.h.p. largest clique has size $\omega(G) \approx 2\log n$



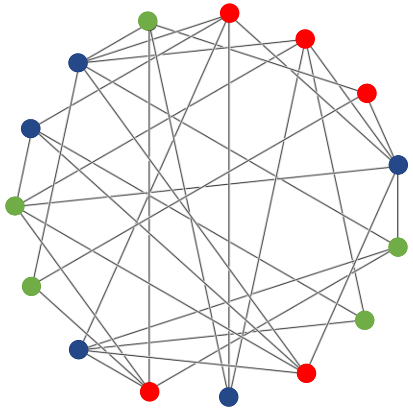
- Brute-force algorithm proves G has no clique of size $k > \omega(G)$ in time $n^{O(\omega(G))}$
- $\exists$ algorithm that proves G has no k-clique for $\omega(G) < k < \sqrt{n}$ in poly-time?
- $\exists$ proof that G has no k-clique for $\omega(G) < k < \sqrt{n}$ in poly-size?



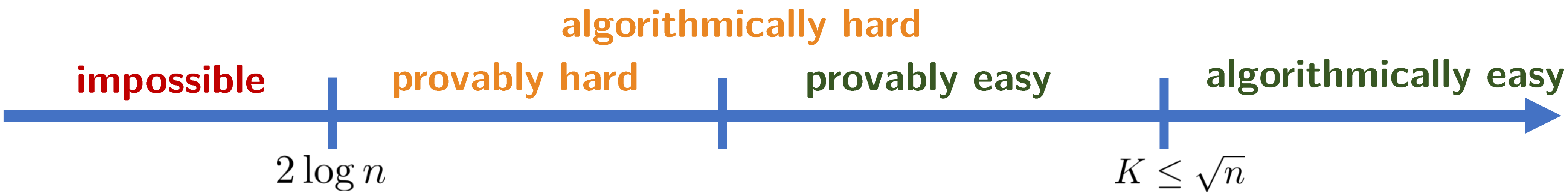
Clique problem: prove G has no k-clique

Erdős-Rényi random graph $G \sim \mathcal{G}(n, 1/2)$

w.h.p. largest clique has size $\omega(G) \approx 2 \log n$



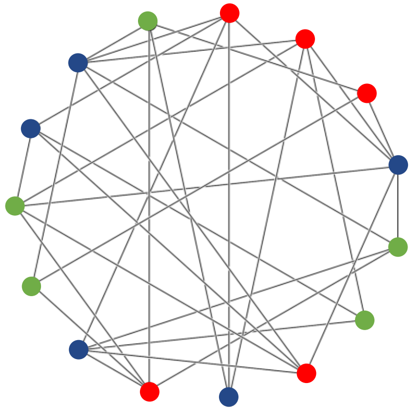
- Brute-force algorithm proves G has no clique of size $k > \omega(G)$ in time $n^{O(\omega(G))}$
- $\exists$ algorithm that proves G has no k-clique for $\omega(G) < k < \sqrt{n}$ in poly-time?
- $\exists$ proof that G has no k-clique for $\omega(G) < k < \sqrt{n}$ in poly-size?



Clique problem: prove G has no k-clique

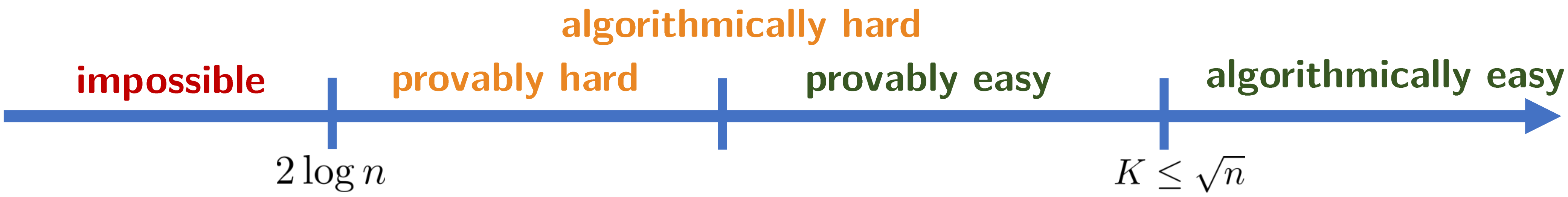
Erdős-Rényi random graph $G \sim \mathcal{G}(n, 1/2)$

w.h.p. largest clique has size $\omega(G) \approx 2 \log n$



$\omega(G) \leq \theta(G) \leq \chi(G)$

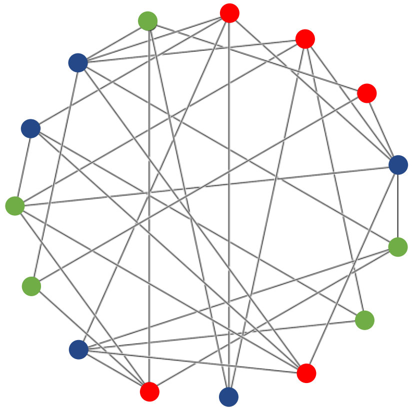
- Brute-force algorithm proves G has no clique of size $k > \omega(G)$ in time $n^{O(\omega(G))}$
- $\exists$ algorithm that proves G has no k-clique for $\omega(G) < k < \sqrt{n}$ in poly-time?
- $\exists$ proof that G has no k-clique for $\omega(G) < k < \sqrt{n}$ in poly-size?



Clique problem: prove G has no k-clique

Erdős-Rényi random graph $G \sim \mathcal{G}(n, 1/2)$

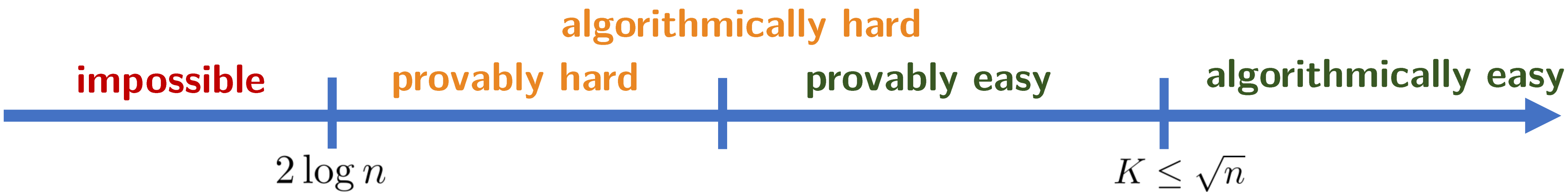
w.h.p. largest clique has size $\omega(G) \approx 2 \log n$



Focus on $k = \omega(G) + 1$

$\omega(G) \leq \theta(G) \leq \chi(G)$

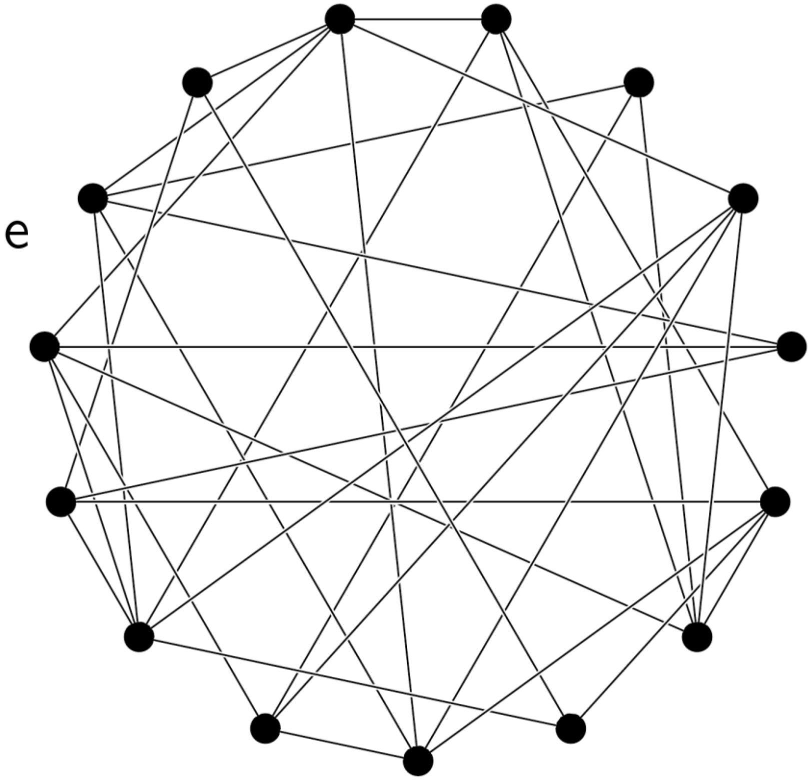
- Brute-force algorithm proves G has no clique of size $k > \omega(G)$ in time $n^{O(\omega(G))}$
- $\exists$ algorithm that proves G has no k-clique for $\omega(G) < k < \sqrt{n}$ in poly-time?
- $\exists$ proof that G has no k-clique for $\omega(G) < k < \sqrt{n}$ in poly-size?



Clique problem: prove G has no k-clique

- Formula Clique(G, k)
states that G has a k clique

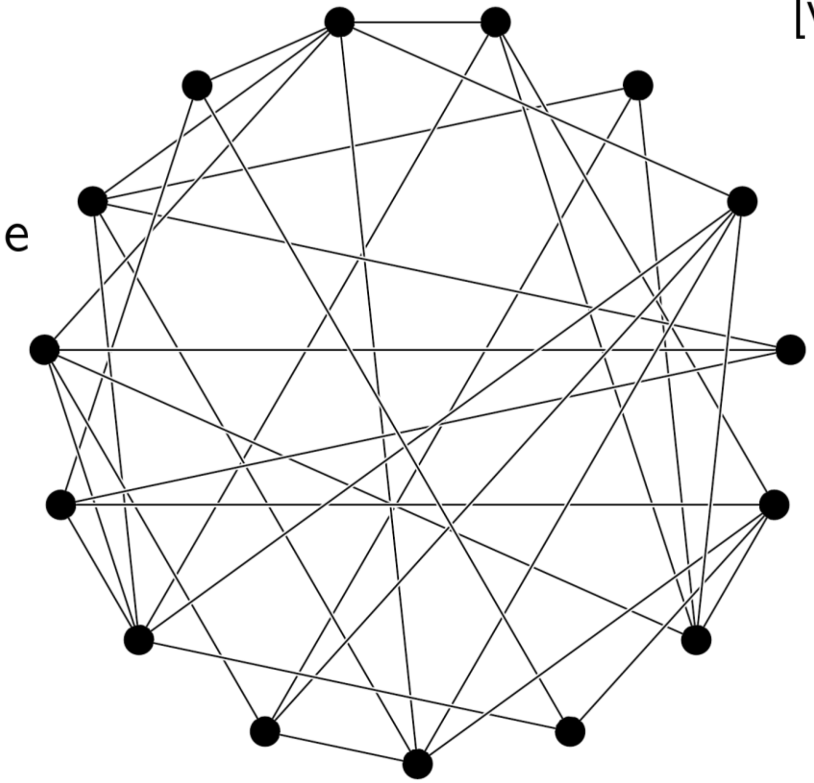
- Prove Clique(G, k) is unsatisfiable



Clique problem: prove G has no k-clique

- Formula Clique(G, k) states that G has a k clique

- Prove Clique(G, k) is unsatisfiable



Variable x_{vi} = [vertex v is ith member of clique]

There are k clique members

$$\bigvee_{v \in V} x_{vi} \qquad \forall i \in [k]$$

A vertex can only be once in clique

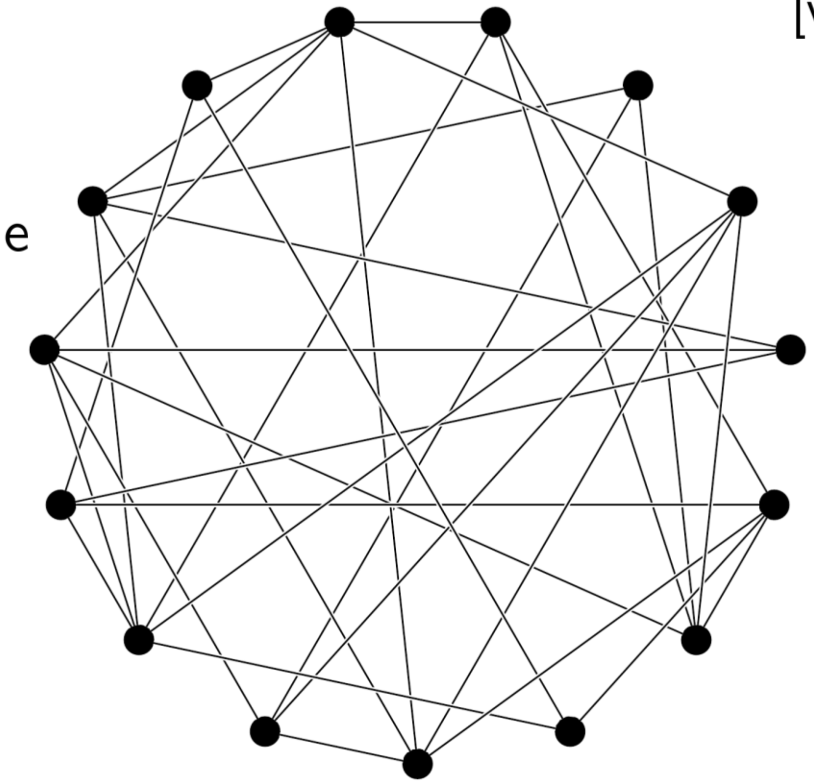
$$\overline{x}_{vi} \vee \overline{x}_{vi'} \qquad \forall v \in V,\ \forall i \neq i' \in [k]$$

Non-neighbors are not both in clique

$$\overline{x}_{ui} \vee \overline{x}_{vi'} \qquad \forall (u, v) \notin E,\ \forall i, i' \in [k]$$

Clique problem: prove G has no k-clique

- Formula Clique(G,k) states that G has a k clique
- Prove Clique(G,k) is unsatisfiable
- Lower bound size of refutation?



Variable x_{vi} = [vertex v is ith member of clique]

There are k clique members

$$\bigvee_{v \in V} x_{vi} \qquad \forall i \in [k]$$

A vertex can only be once in clique

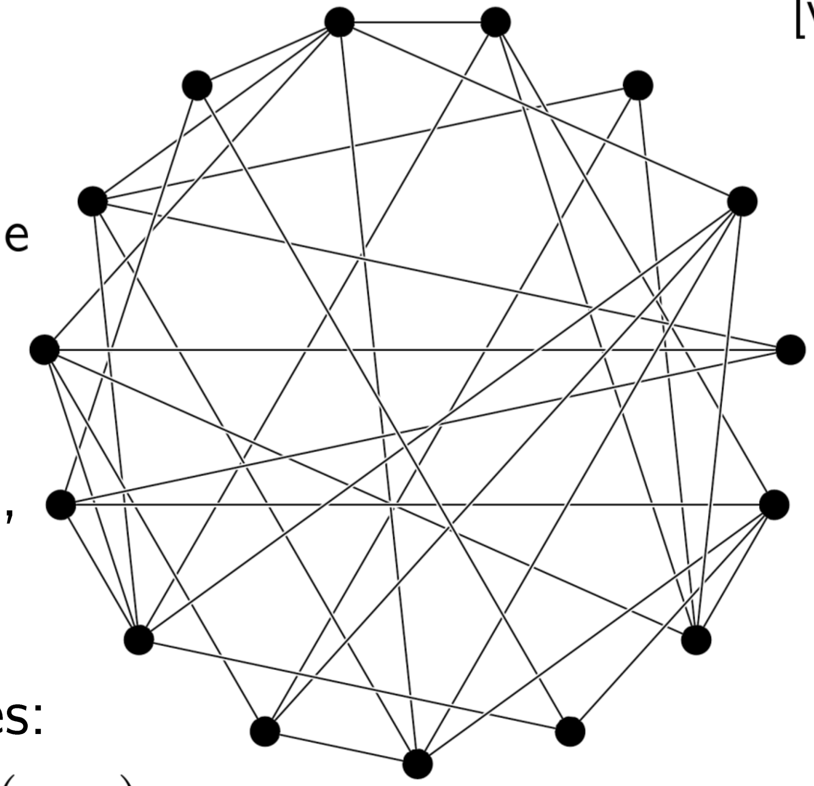
$$\overline{x}_{vi} \vee \overline{x}_{vi'} \qquad \begin{array}{c}\forall v \in V\\ \forall i \neq i' \in [k]\end{array}$$

Non-neighbors are not both in clique

$$\overline{x}_{ui} \vee \overline{x}_{vi'} \qquad \begin{array}{c}\forall (u,v) \notin E\\ \forall i, i' \in [k]\end{array}$$

Clique problem: prove G has no k-clique

- Formula Clique(G, k) states that G has a k clique
- Prove Clique(G, k) is unsatisfiable
- Lower bound size of refutation?
- Can we show that "brute-force", size $n^{\Theta(k)}$ refutation is optimal?
- Natural candidate hard instances: Erdős-Rényi random graph $G \sim \mathcal{G}(n, p)$
  p close to k-clique threshold



Variable x_{vi} = [vertex v is ith member of clique]

There are k clique members

$$\bigvee_{v \in V} x_{vi} \qquad \forall i \in [k]$$

A vertex can only be once in clique

$$\overline{x}_{vi} \vee \overline{x}_{vi'} \qquad \forall v \in V, \ \forall i \neq i' \in [k]$$

Non-neighbors are not both in clique

$$\overline{x}_{ui} \vee \overline{x}_{vi'} \qquad \forall (u, v) \notin E, \ \forall i, i' \in [k]$$

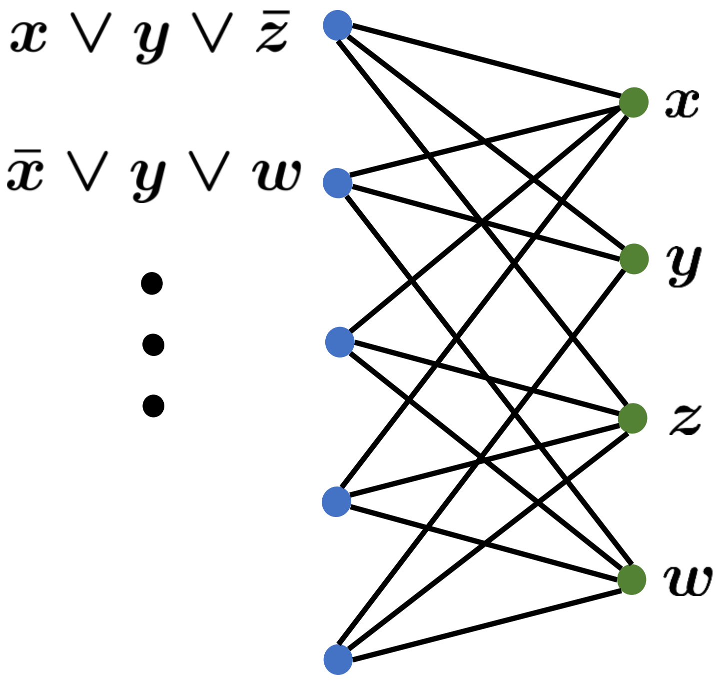
Other candidate hard formulas

Other candidate hard formulas

- Random k-SAT
  - Choose each of the $2^k\binom{n}{k}$ possible clauses with probability p
  - Choose m clauses of $2^k\binom{n}{k}$ possible clauses uniformly at random

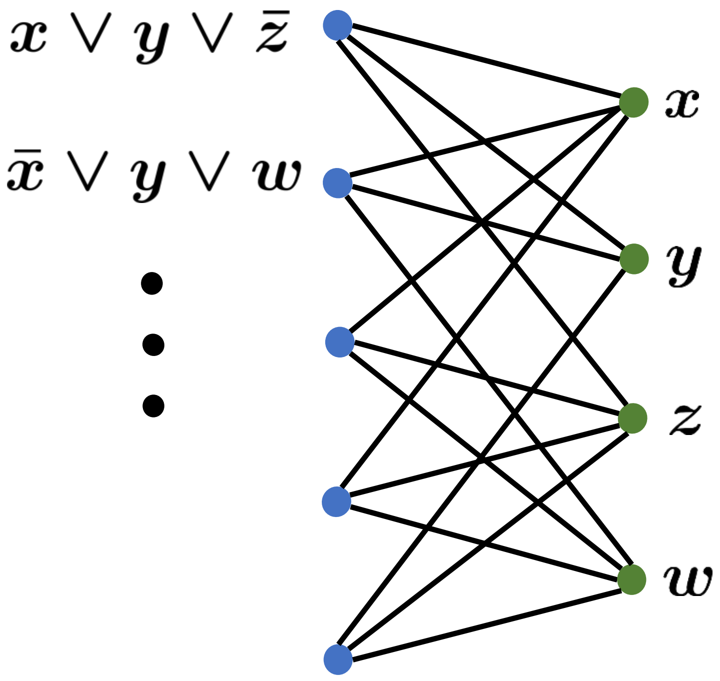
Other candidate hard formulas

- Random k-SAT
  - Choose each of the $2^k\binom{n}{k}$ possible clauses with probability p
  - Choose m clauses of $2^k\binom{n}{k}$ possible clauses uniformly at random
- Clause-variable incidence graph (here, no signs)



Other candidate hard formulas

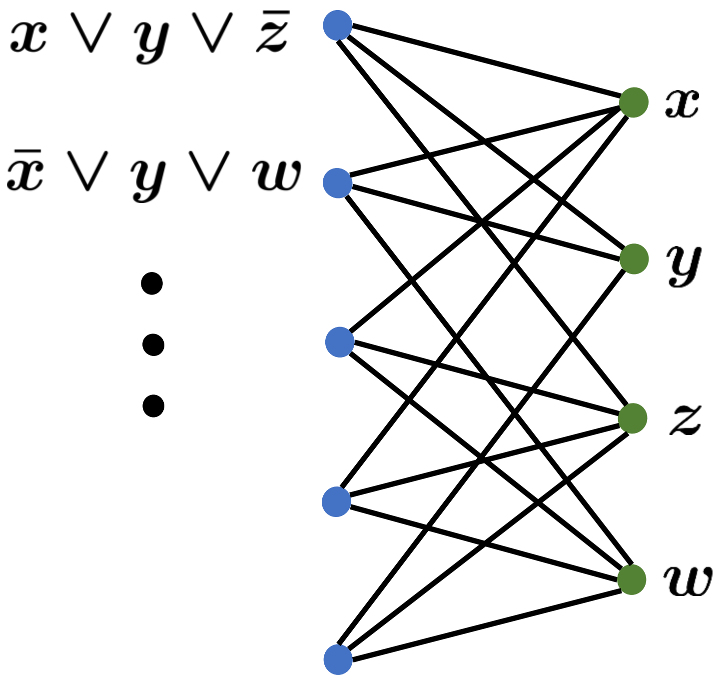
- Random k-SAT
  - Choose each of the $2^k\binom{n}{k}$ possible clauses with probability p
  - Choose m clauses of $2^k\binom{n}{k}$ possible clauses uniformly at random
- Clause-variable incidence graph (here, no signs)



- Related easy formula: random k-XOR
  - Only $2\binom{n}{k}$ possible XOR constraints
  - $x \oplus y \oplus z = 0$ or $x \oplus y \oplus z = 1$

Other candidate hard formulas

- Random k-SAT
  - Choose each of the $2^k\binom{n}{k}$ possible clauses with probability p
  - Choose m clauses of $2^k\binom{n}{k}$ possible clauses uniformly at random
- Clause-variable incidence graph (here, no signs)



- Related easy formula: random k-XOR
  - Only $2\binom{n}{k}$ possible XOR constraints
  - $x \oplus y \oplus z = 0$ or $x \oplus y \oplus z = 1$
- Rewrite constraint in CNF
  - $x \oplus y \oplus z = 0$ becomes 4 clauses:

$$(\bar{x} \vee \bar{y} \vee \bar{z}) \wedge (x \vee y \vee \bar{z}) \wedge (x \vee \bar{y} \vee z) \wedge (\bar{x} \vee y \vee z)$$

k-coloring formula

- Quite different from k-clique
  - NP-hard for $k = 3$

k-coloring formula

- Quite different from k-clique
  - NP-hard for $k = 3$

Formula $\mathrm{Color}(G, k)$:

Variable x_{vi} = [vertex v is colored i]

Every vertex has a color

$$\bigvee_{i \in [k]} x_{vi} \qquad \forall v \in V$$

A vertex has only one color

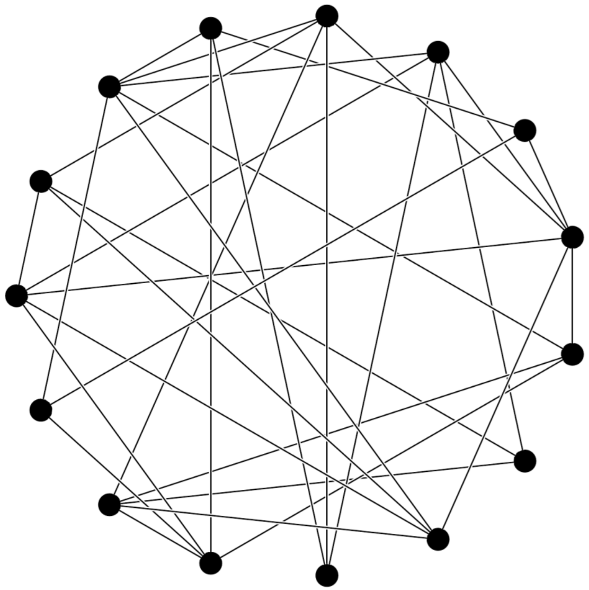
$$\overline{x}_{vi} \vee \overline{x}_{vi'} \qquad \begin{array}{c} \forall v \in V \\ \forall i \neq i' \in [k] \end{array}$$

Neighbors don't have same color

$$\overline{x}_{ui} \vee \overline{x}_{vi} \qquad \begin{array}{c} \forall (u, v) \in E \\ \forall i \in [k] \end{array}$$

k-coloring formula

- Quite different from k-clique
  - NP-hard for $k = 3$
- Natural hard candidate:
  - Erdős-Rényi random graph $\boldsymbol{G} \sim \mathcal{G}(n, p)$
    p close to k-colorable threshold



Formula $\mathrm{Color}(G, k)$:

Variable x_{vi} = [vertex v is colored i]

Every vertex has a color

$$\bigvee_{i \in [k]} x_{vi} \qquad \forall v \in V$$

A vertex has only one color

$$\overline{x}_{vi} \vee \overline{x}_{vi'} \qquad \begin{array}{c} \forall v \in V \\ \forall i \neq i' \in [k] \end{array}$$

Neighbors don't have same color

$$\overline{x}_{ui} \vee \overline{x}_{vi} \qquad \begin{array}{c} \forall (u, v) \in E \\ \forall i \in [k] \end{array}$$

Reasons for studying average-case proof complexity

- Stronger statement than worst-case: almost all graphs are hard

Reasons for studying average-case proof complexity

- Stronger statement than worst-case: almost all graphs are hard

- There are not many natural hard candidates (to prove lower bounds for stronger proof systems)
  - Many lower bounds are for “easy formulas”: pigeonhole principle, Tseitin, clique-coloring principle “a (k-1)-colorable graph does not contain a k-clique”

Reasons for studying average-case proof complexity

- Stronger statement than worst-case: almost all graphs are hard
- There are not many natural hard candidates (to prove lower bounds for stronger proof systems)
  - Many lower bounds are for "easy formulas": pigeonhole principle, Tseitin, clique-coloring principle "a (k-1)-colorable graph does not contain a k-clique"
  - This talk: focus on average-case complexity of three NP-hard problems

Reasons for studying average-case proof complexity

- Stronger statement than worst-case: almost all graphs are hard
- There are not many natural hard candidates (to prove lower bounds for stronger proof systems)
  - Many lower bounds are for "easy formulas": pigeonhole principle, Tseitin, clique-coloring principle "a (k-1)-colorable graph does not contain a k-clique"
  - This talk: focus on average-case complexity of three NP-hard problems
- Many combinatorial formulas are of independent interest

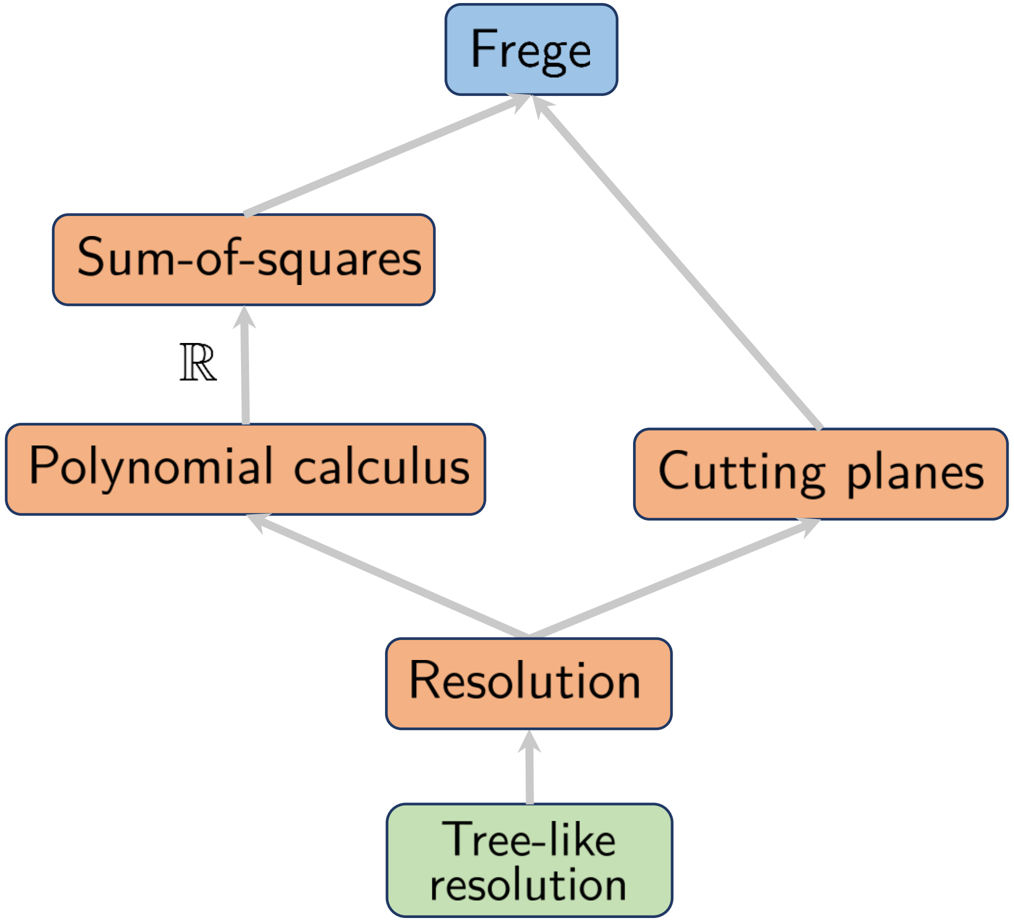
Proof Systems and Lower Bounds

Proof Systems

- Given unsat CNF formula, how can we refute it?

Proof Systems

- Given unsat CNF formula, how can we refute it?
- Define some proof systems

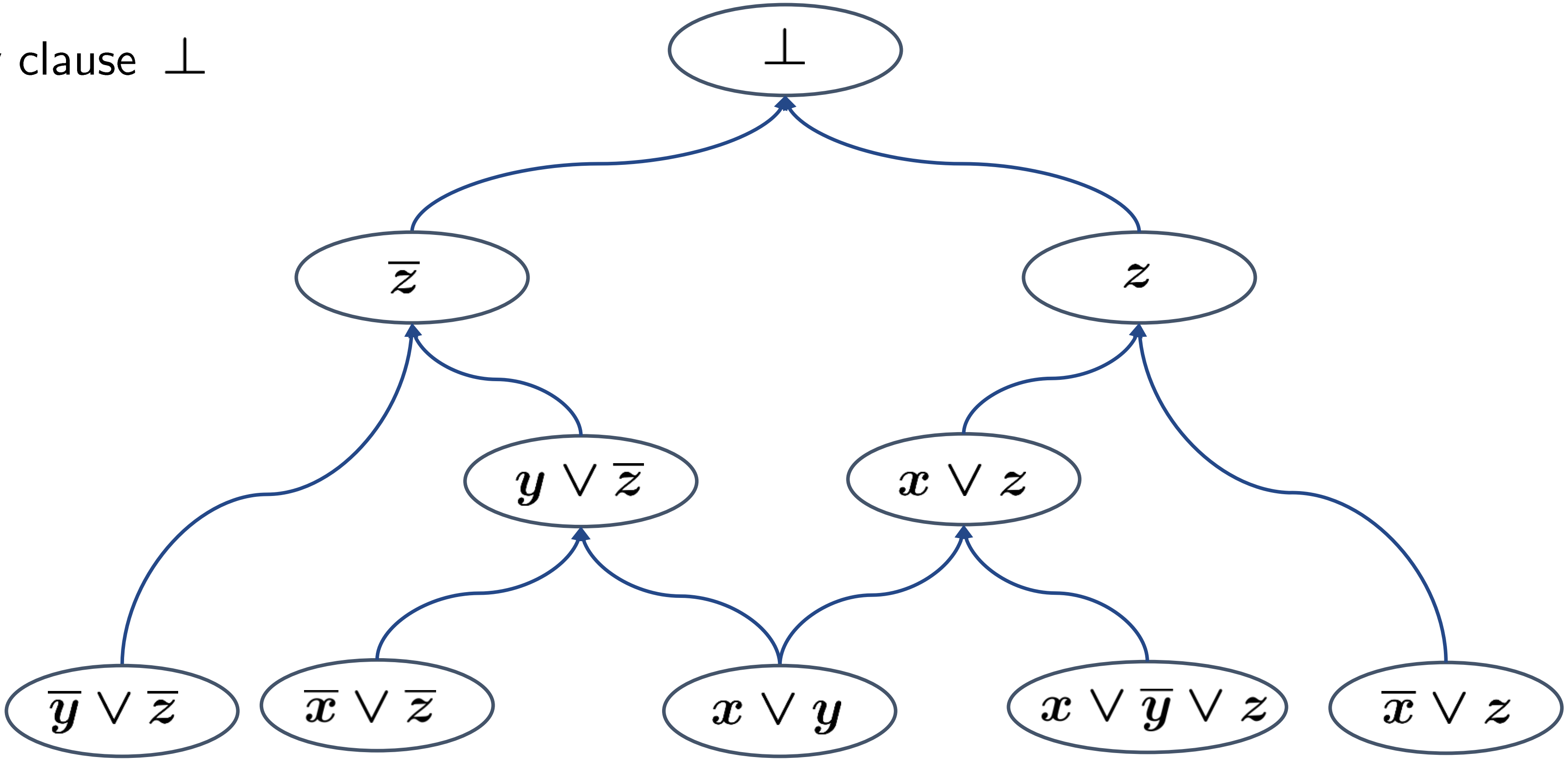


Resolution

UNSAT k-CNF formula F: $(\overline{y} \vee \overline{z}) \wedge (\overline{x} \vee \overline{z}) \wedge (x \vee y) \wedge (x \vee \overline{y} \vee z) \wedge (\overline{x} \vee z)$

Resolution rule: $\frac{C \vee x \quad D \vee \overline{x}}{C \vee D}$

Refutation: Derivation of empty clause $\bot$



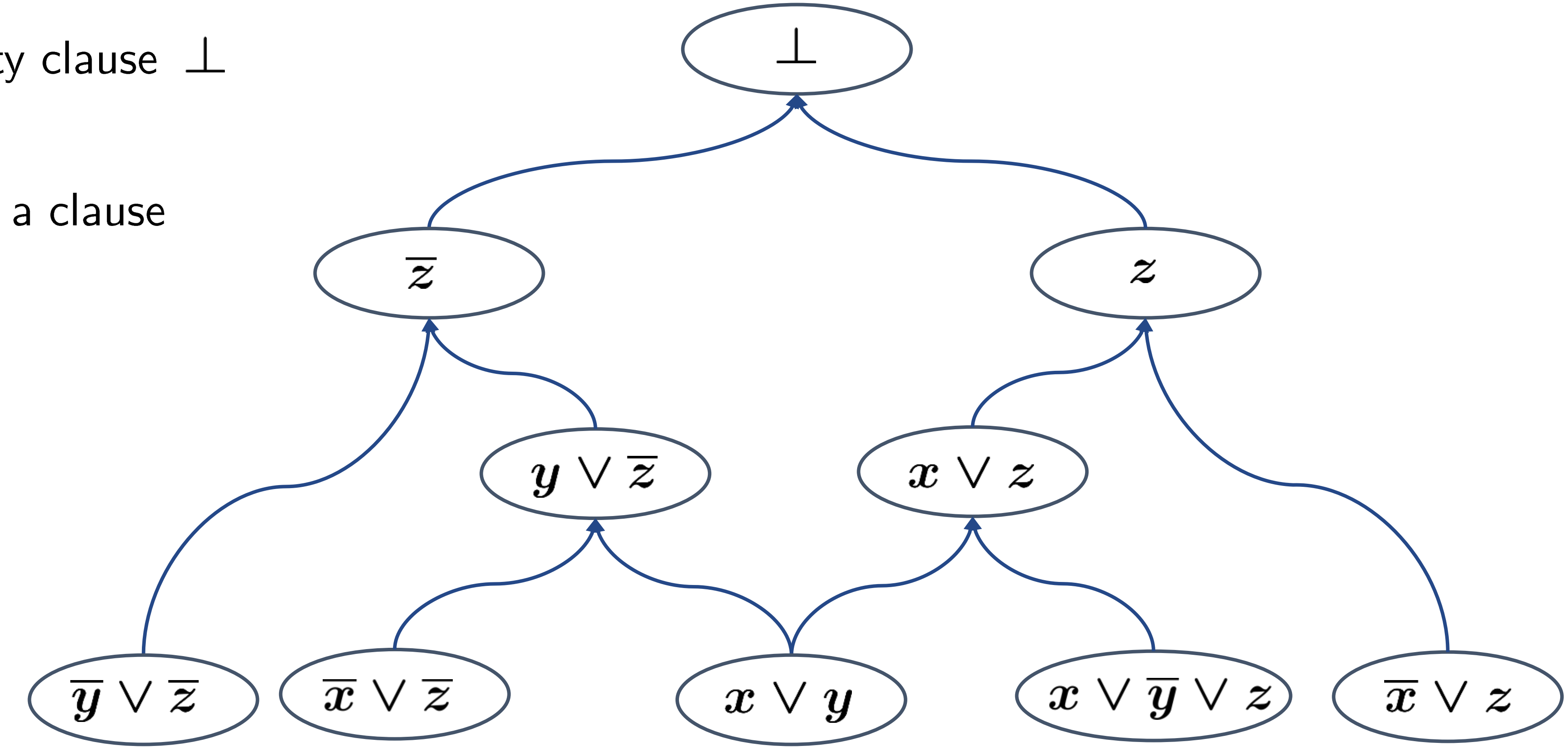
Resolution

UNSAT k-CNF formula F: $(\overline{y} \vee \overline{z}) \wedge (\overline{x} \vee \overline{z}) \wedge (x \vee y) \wedge (x \vee \overline{y} \vee z) \wedge (\overline{x} \vee z)$

Resolution rule: $\frac{C \vee x \quad D \vee \overline{x}}{C \vee D}$

Refutation: Derivation of empty clause $\perp$

Proof size: # clauses in proof
Proof width: max # literals in a clause



Resolution

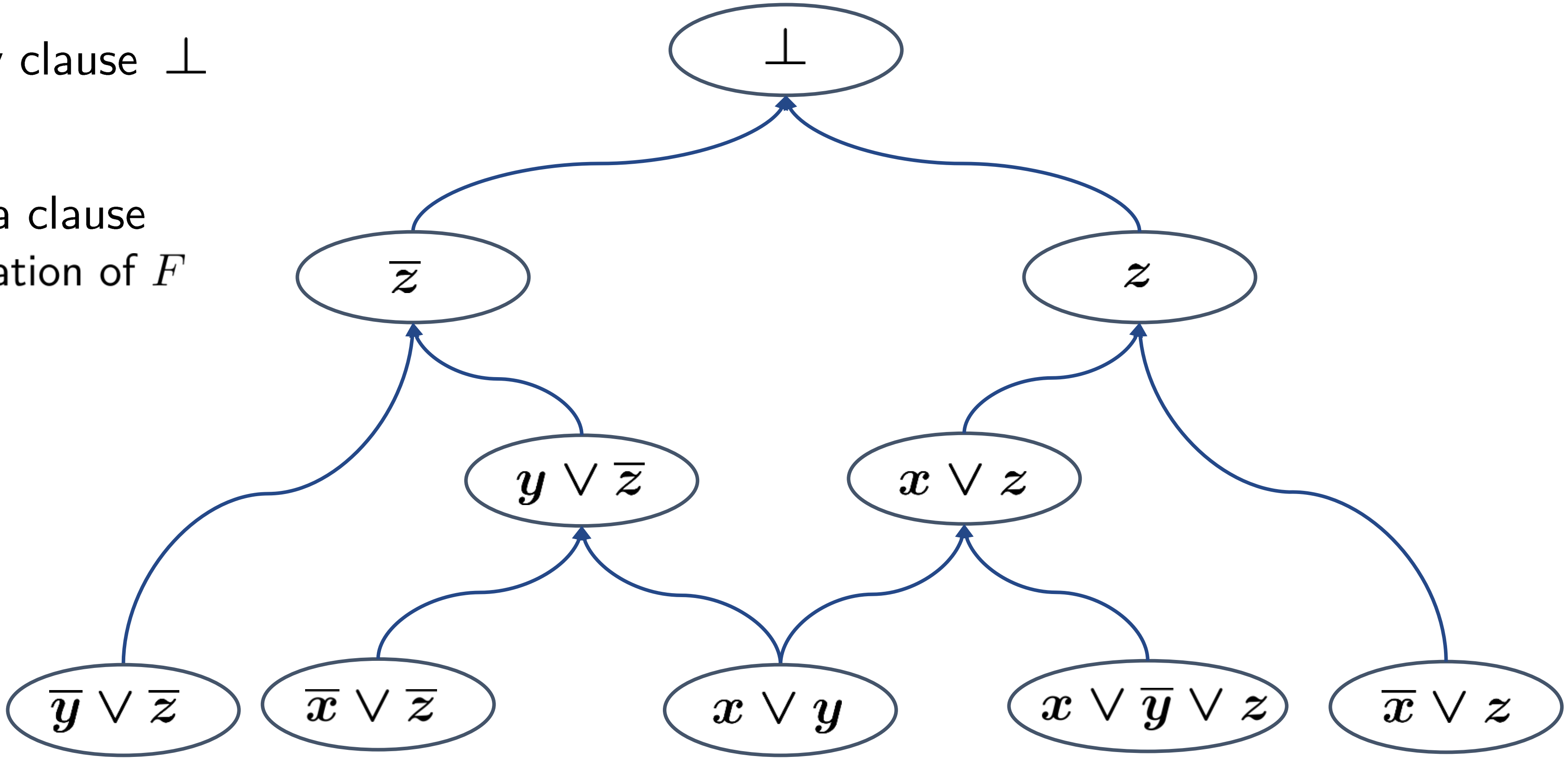
UNSAT k-CNF formula F: $(\overline{y} \vee \overline{z}) \wedge (\overline{x} \vee \overline{z}) \wedge (x \vee y) \wedge (x \vee \overline{y} \vee z) \wedge (\overline{x} \vee z)$

Resolution rule: $\dfrac{C \vee x \quad D \vee \overline{x}}{C \vee D}$

Refutation: Derivation of empty clause $\perp$

Proof size: # clauses in proof
Proof width: max # literals in a clause
$w =$ smallest width of any refutation of F

Algorithm in time $\approx n^w$



Resolution

UNSAT k-CNF formula F: $(\overline{y} \vee \overline{z}) \wedge (\overline{x} \vee \overline{z}) \wedge (x \vee y) \wedge (x \vee \overline{y} \vee z) \wedge (\overline{x} \vee z)$

Resolution rule: $\dfrac{C \vee x \quad D \vee \overline{x}}{C \vee D}$

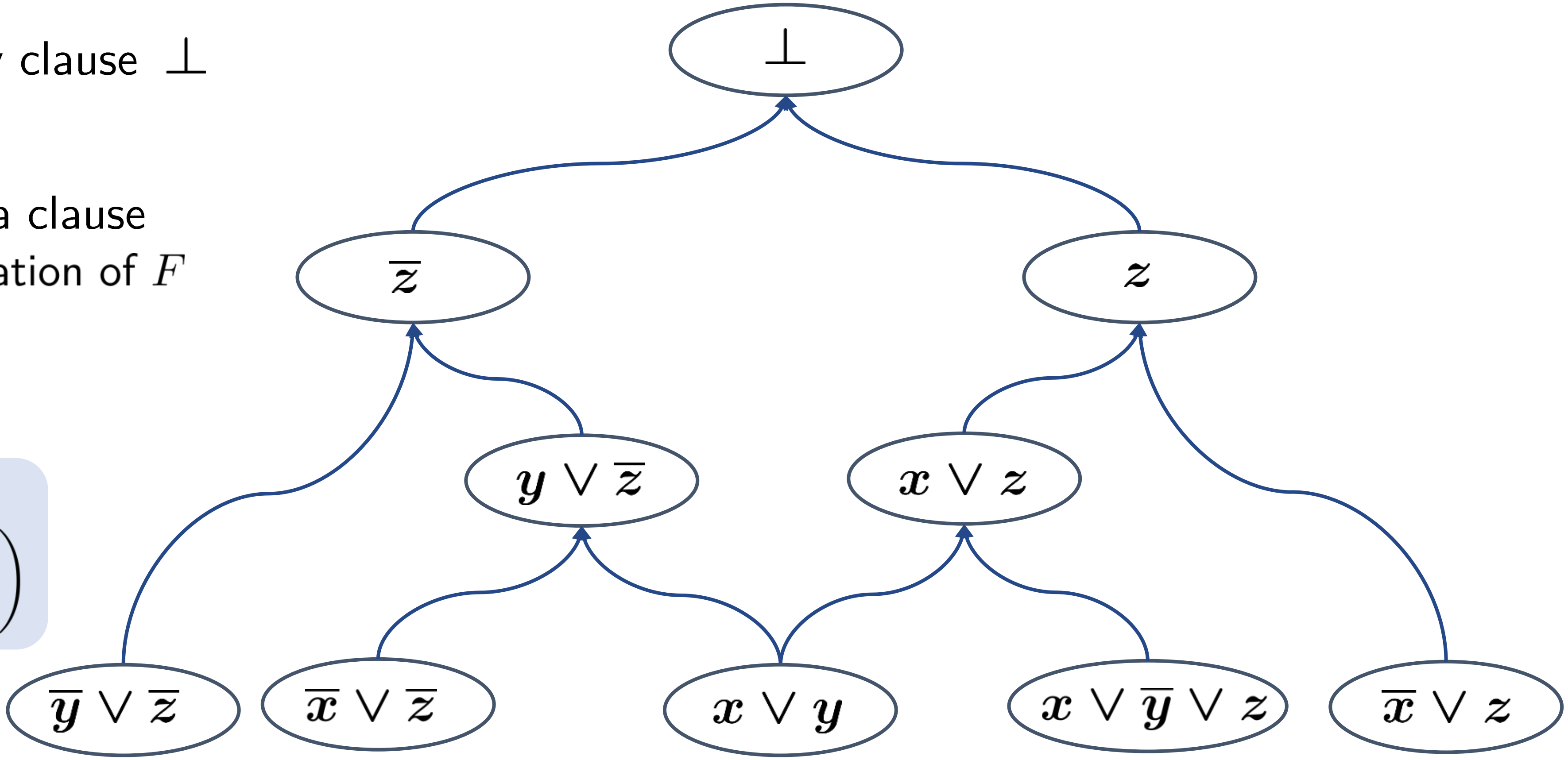
Refutation: Derivation of empty clause $\perp$

Proof size: # clauses in proof
Proof width: max # literals in a clause
$w =$ smallest width of any refutation of F

Algorithm in time $\approx n^w$

Theorem [BW01]

Proof size $\geq \exp\left(\Omega\left(\frac{(w-k)^2}{n}\right)\right)$



Resolution

UNSAT k-CNF formula F: $(\overline{y} \vee \overline{z}) \wedge (\overline{x} \vee \overline{z}) \wedge (x \vee y) \wedge (x \vee \overline{y} \vee z) \wedge (\overline{x} \vee z)$

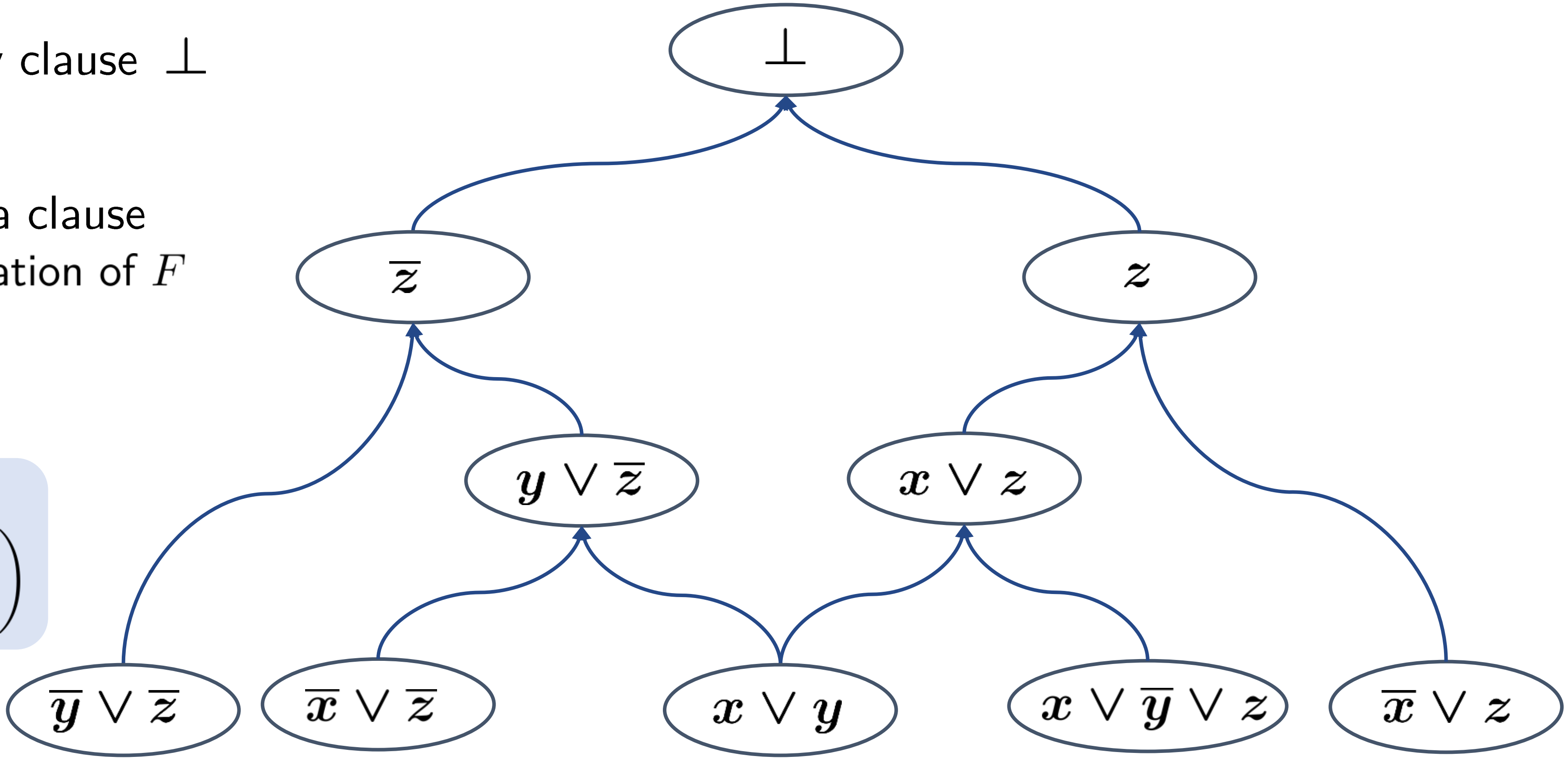
Resolution rule: $\dfrac{C \vee x \quad D \vee \overline{x}}{C \vee D}$

Refutation: Derivation of empty clause $\perp$

Proof size: $\#$ clauses in proof
Proof width: max $\#$ literals in a clause
$w =$ smallest width of any refutation of F

Algorithm in time $\approx n^w$

Theorem [BW01]
Proof size $\geq \exp\left(\Omega\left(\frac{(w-k)^2}{n}\right)\right)$



Tree-like: proof DAG is a tree

Resolution

UNSAT k-CNF formula F: $(\overline{y} \vee \overline{z}) \wedge (\overline{x} \vee \overline{z}) \wedge (x \vee y) \wedge (x \vee \overline{y} \vee z) \wedge (\overline{x} \vee z)$

Resolution rule: $\dfrac{C \vee x \quad D \vee \overline{x}}{C \vee D}$

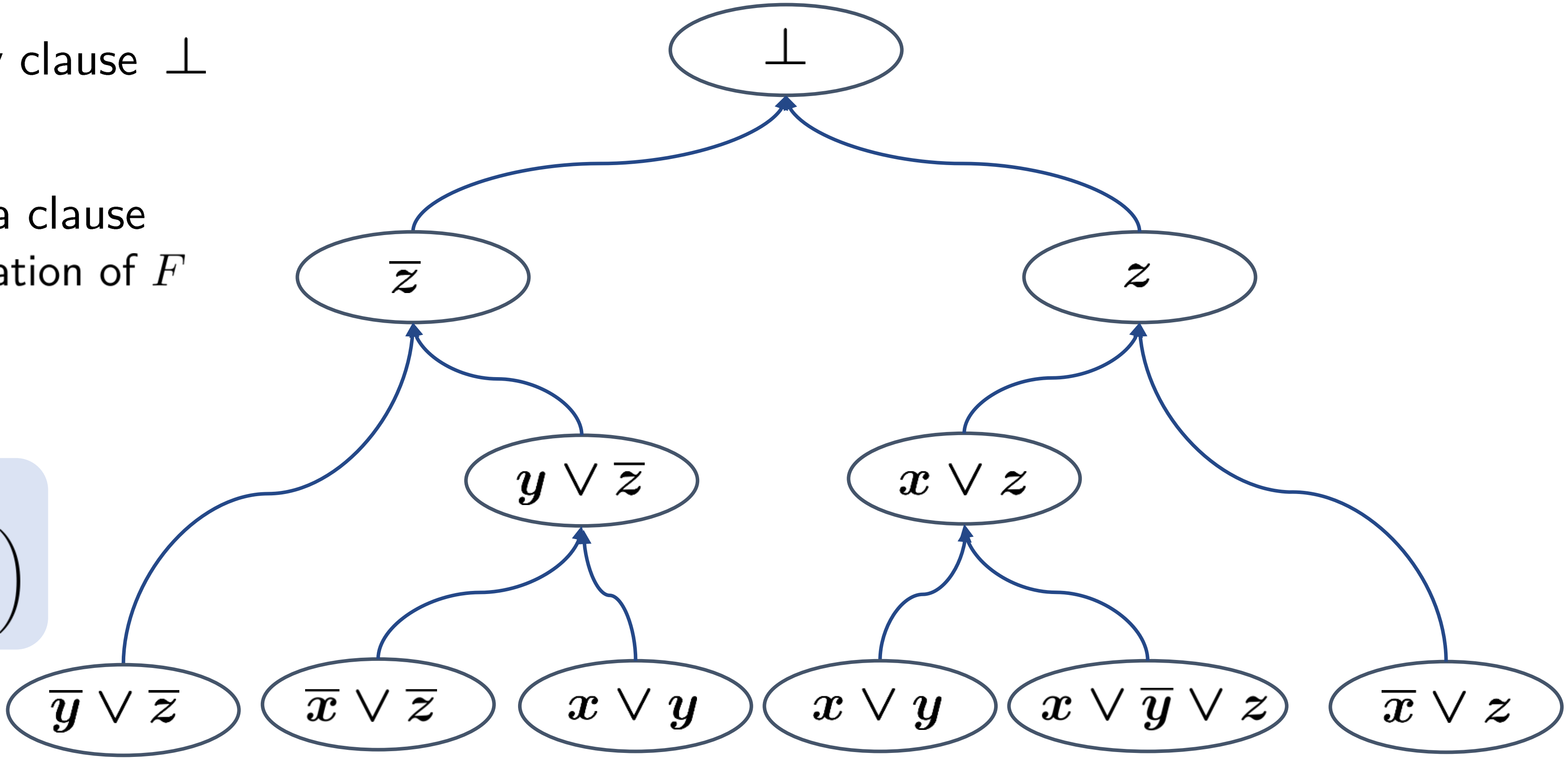
Refutation: Derivation of empty clause $\perp$

Proof size: # clauses in proof
Proof width: max # literals in a clause
$w =$ smallest width of any refutation of F

Algorithm in time $\approx n^w$

Theorem [BW01]
Proof size $\geq \exp\left(\Omega\left(\dfrac{(w-k)^2}{n}\right)\right)$

Tree-like: proof DAG is a tree



Polynomial calculus

Cutting planes

UNSAT k-CNF formula F: $(\overline{y} \vee \overline{z}) \wedge (\overline{x} \vee \overline{z}) \wedge (x \vee y) \wedge (x \vee \overline{y} \vee z) \wedge (\overline{x} \vee z)$

Polynomial calculus

Cutting planes

UNSAT k-CNF formula F: $(\overline{y} \vee \overline{z}) \wedge (\overline{x} \vee \overline{z}) \wedge (x \vee y) \wedge (x \vee \overline{y} \vee z) \wedge (\overline{x} \vee z)$

$(x \vee \overline{y} \vee z) \rightsquigarrow (1-x)y(1-z) = 0$

$(x \vee \overline{y} \vee z) \rightsquigarrow x + (1-y) + z \geq 1$

Polynomial calculus

Cutting planes

UNSAT k-CNF formula F: $(\overline{y} \vee \overline{z}) \wedge (\overline{x} \vee \overline{z}) \wedge (x \vee y) \wedge (x \vee \overline{y} \vee z) \wedge (\overline{x} \vee z)$

$(x \vee \overline{y} \vee z) \rightsquigarrow (1-x)y(1-z) = 0$

$(x \vee \overline{y} \vee z) \rightsquigarrow x + (1-y) + z \geq 1$

Boolean axioms: $x^2 - x = 0$

Linear combination: over $\mathbb{F}$ $\dfrac{p = 0 \quad q = 0}{\alpha p + \beta q = 0}$

Multiply by variable: $\dfrac{p = 0}{xp = 0}$

Refutation: Derivation of $1 = 0$

Polynomial calculus

Cutting planes

UNSAT k-CNF formula F: $(\overline{y} \vee \overline{z}) \wedge (\overline{x} \vee \overline{z}) \wedge (x \vee y) \wedge (x \vee \overline{y} \vee z) \wedge (\overline{x} \vee z)$

$(x \vee \overline{y} \vee z) \rightsquigarrow (1-x)y(1-z) = 0$

$(x \vee \overline{y} \vee z) \rightsquigarrow x + (1-y) + z \geq 1$

Boolean axioms: $x^2 - x = 0$

Linear combination: over $\mathbb{F}$ $\dfrac{p = 0 \quad q = 0}{\alpha p + \beta q = 0}$

Multiply by variable: $\dfrac{p = 0}{xp = 0}$

Refutation: Derivation of $1 = 0$

Proof size: # monomials in proof
Proof degree: max degree of any polynomial

Polynomial calculus

Cutting planes

UNSAT k-CNF formula F: $(\overline{y} \vee \overline{z}) \wedge (\overline{x} \vee \overline{z}) \wedge (x \vee y) \wedge (x \vee \overline{y} \vee z) \wedge (\overline{x} \vee z)$

Polynomial calculus

$(x \vee \overline{y} \vee z) \rightsquigarrow (1-x)y(1-z) = 0$

Boolean axioms: $x^2 - x = 0$

Linear combination over $\mathbb{F}$: $\dfrac{p = 0 \quad q = 0}{\alpha p + \beta q = 0}$

Multiply by variable: $\dfrac{p = 0}{xp = 0}$

Refutation: Derivation of $1 = 0$

Proof size: # monomials in proof
Proof degree: max degree of any polynomial

Cutting planes

$(x \vee \overline{y} \vee z) \rightsquigarrow x + (1-y) + z \geq 1$

Boolean axioms: $0 \leq x \leq 1$

Linear combination: $\dfrac{p \geq A \quad q \geq B}{\alpha p + \beta q \geq \alpha A + \beta B}$

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

Refutation: Derivation of $-1 \geq 0$

Polynomial calculus

Cutting planes

UNSAT k-CNF formula F: $(\overline{y} \vee \overline{z}) \wedge (\overline{x} \vee \overline{z}) \wedge (x \vee y) \wedge (x \vee \overline{y} \vee z) \wedge (\overline{x} \vee z)$

Polynomial calculus

$(x \vee \overline{y} \vee z) \rightsquigarrow (1-x)y(1-z) = 0$

Boolean axioms: $x^2 - x = 0$

Linear combination over $\mathbb{F}$: $\dfrac{p = 0 \quad q = 0}{\alpha p + \beta q = 0}$

Multiply by variable: $\dfrac{p = 0}{xp = 0}$

Refutation: Derivation of $1 = 0$

Proof size: # monomials in proof
Proof degree: max degree of any polynomial

Cutting planes

$(x \vee \overline{y} \vee z) \rightsquigarrow x + (1-y) + z \geq 1$

Boolean axioms: $0 \leq x \leq 1$

Linear combination: $\dfrac{p \geq A \quad q \geq B}{\alpha p + \beta q \geq \alpha A + \beta B}$

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

Refutation: Derivation of $-1 \geq 0$

Proof size: # inequalities in proof

Sum of squares

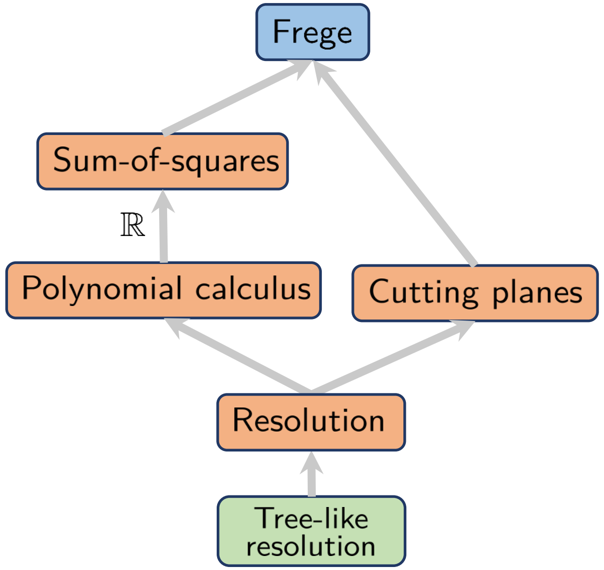
UNSAT k-CNF formula F: $(\overline{y} \vee \overline{z}) \wedge (\overline{x} \vee \overline{z}) \wedge (x \vee y) \wedge (x \vee \overline{y} \vee z) \wedge (\overline{x} \vee z)$

$(x \vee \overline{y} \vee z) \rightsquigarrow (1-x)y(1-z) = 0$

Boolean axioms: $x^2 - x = 0$

$(x \vee \overline{y} \vee z) \rightsquigarrow x + (1-y) + z \geq 1$

Boolean axioms: $0 \leq x \leq 1$



Sum of squares

UNSAT k-CNF formula F: $(\overline{y} \vee \overline{z}) \wedge (\overline{x} \vee \overline{z}) \wedge (x \vee y) \wedge (x \vee \overline{y} \vee z) \wedge (\overline{x} \vee z)$

$(x \vee \overline{y} \vee z) \rightsquigarrow (1-x)y(1-z) = 0$

Boolean axioms: $x^2 - x = 0$

$(x \vee \overline{y} \vee z) \rightsquigarrow x + (1-y) + z \geq 1$

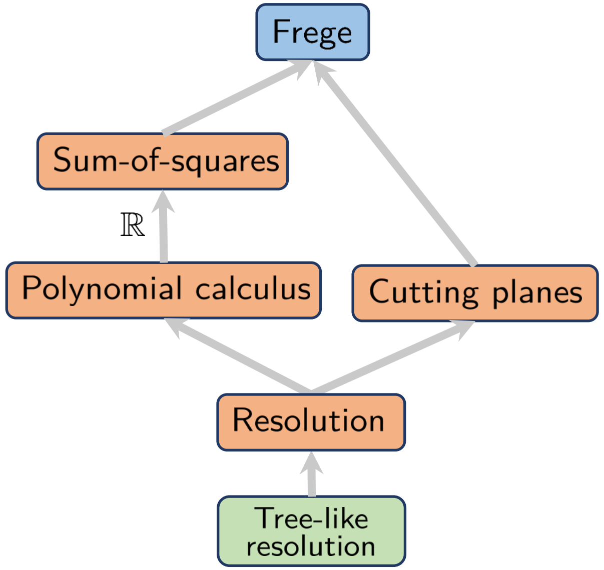
Boolean axioms: $0 \leq x \leq 1$

Polynomials $\mathcal{P} = \{P_1 = 0, P_2 = 0, \ldots, P_m = 0; Q_1 \geq 0, Q_2 \geq 0, \ldots, Q_\ell \geq 0\}$

SoS refutation of $\mathcal{P}$: $R_1, R_2, \ldots, R_m; S_1, S_2, \ldots, S_\ell$ s.t.

$$\sum_{i \in [m]} R_i P_i + \sum_{i \in [\ell]} S_i Q_i = -1$$

where each S_i is a sum of squares



Sum of squares

UNSAT k-CNF formula F: $(\overline{y} \vee \overline{z}) \wedge (\overline{x} \vee \overline{z}) \wedge (x \vee y) \wedge (x \vee \overline{y} \vee z) \wedge (\overline{x} \vee z)$

$(x \vee \overline{y} \vee z) \rightsquigarrow (1-x)y(1-z) = 0$

Boolean axioms: $x^2 - x = 0$

$(x \vee \overline{y} \vee z) \rightsquigarrow x + (1-y) + z \geq 1$

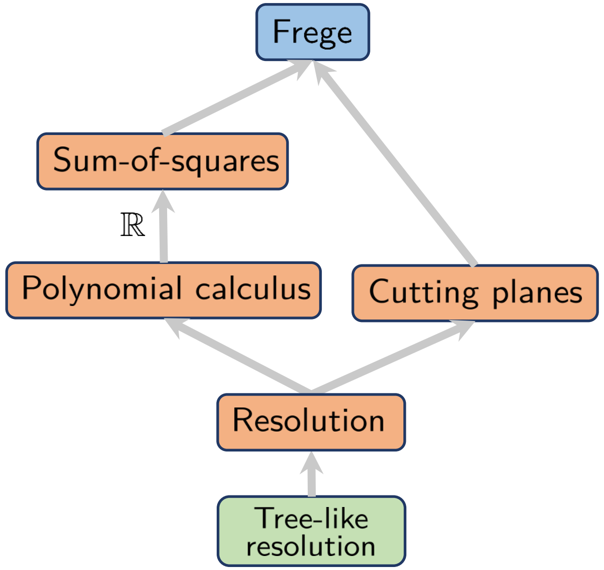
Boolean axioms: $0 \leq x \leq 1$

Polynomials $\mathcal{P} = \{P_1 = 0, P_2 = 0, \ldots, P_m = 0; Q_1 \geq 0, Q_2 \geq 0, \ldots, Q_\ell \geq 0\}$

SoS refutation of $\mathcal{P}$: $R_1, R_2, \ldots, R_m; S_1, S_2, \ldots, S_\ell$ s.t.

$$\sum_{i \in [m]} R_i P_i + \sum_{i \in [\ell]} S_i Q_i = -1$$

where each S_i is a sum of squares



Proof size: # monomials when we expand proof
Proof degree: max degree of any polynomial

Average-case hardness results

	k-clique	k-coloring	3-SAT	3-XOR
Tree-like Resolution			HARD [Chvátal, Szemerédi '88] Improved [Ben-Sasson, Galesi '01] (size $\exp(n/\Delta^{1+\epsilon})$) $\Delta = m/n$	
Resolution			HARD [Chvátal, Szemerédi '88] $\exp(n/\Delta^{2+\epsilon})$ Improved [Beame, Karp, Pitassi, Saks '98], [Ben-Sasson '01]	
Polynomial Calculus				
Sum of Squares				
Cutting Planes				

Average-case hardness results

	k-clique	k-coloring	3-SAT	3-XOR
Tree-like Resolution			HARD [Chvátal, Szemerédi '88] Improved [Ben-Sasson, Galesi '01] (size $\exp(n/\Delta^{1+\epsilon})$) $\Delta = m/n$	
Resolution			HARD [Chvátal, Szemerédi '88] $\exp(n/\Delta^{2+\epsilon})$ Improved [Beame, Karp, Pitassi, Saks '98], [Ben-Sasson '01]	
Polynomial Calculus			$\mathbb{F} \neq 2$: HARD [Ben-Sasson, Impagliazzo '99]	
			$\mathbb{F} = 2$: HARD [Alekhnovich, Razborov '01]	EASY
Sum of Squares			HARD [Grigoriev '01, Schoenebeck '08]	
Cutting Planes				

Average-case hardness results

	k-clique	k-coloring	3-SAT	3-XOR
Tree-like Resolution			HARD [Chvátal, Szemerédi '88] Improved [Ben-Sasson, Galesi '01] (size $\exp(n/\Delta^{1+\epsilon})$) $\Delta = m/n$	
Resolution			HARD [Chvátal, Szemerédi '88] $\exp(n/\Delta^{2+\epsilon})$ Improved [Beame, Karp, Pitassi, Saks '98], [Ben-Sasson '01]	
Polynomial Calculus			$\mathbb{F} \neq 2$: HARD [Ben-Sasson, Impagliazzo '99]	
			$\mathbb{F} = 2$: HARD [Alekhnovich, Razborov '01]	EASY
Sum of Squares			HARD [Grigoriev '01, Schoenebeck '08]	
Cutting Planes			OPEN $\Theta(\log n)$-SAT [Fleming, Pankratov, Pitassi, Robere '17] [Hrubeš, Pudlák '17]	Quasi-poly EASY [Fleming, Göös, Impagliazzo, Pitassi, Robere, Tan, Wigderson '21] [Dadush, Tiwari '20]

Average-case hardness results

<table>
<tr><th></th><th>k-clique</th><th>k-coloring</th><th></th><th>3-SAT</th><th>3-XOR</th></tr>
<tr><td>Tree-like Resolution</td><td rowspan="8"></td><td rowspan="2">HARD
[Beame, Culberson, Mitchell, Moore ’05]</td><td colspan="3">HARD [Chvátal, Szemerédi ’88]
Improved [Ben-Sasson, Galesi ’01] (size $\exp(n/\Delta^{1+\epsilon})$) $\Delta = m/n$</td></tr>
<tr><td>Resolution</td><td colspan="3">HARD [Chvátal, Szemerédi ’88] $\exp(n/\Delta^{2+\epsilon})$
Improved [Beame, Karp, Pitassi, Saks ’98], [Ben-Sasson ’01]</td></tr>
<tr><td rowspan="2">Polynomial Calculus</td><td rowspan="2">OPEN</td><td>$\mathbb{F} \neq 2$</td><td colspan="2">HARD [Ben-Sasson, Impagliazzo ’99]</td></tr>
<tr><td>$\mathbb{F} = 2$</td><td>HARD [Alekhnovich, Razborov ’01]</td><td>EASY</td></tr>
<tr><td>Sum of Squares</td><td>OPEN
[Kothari, Manohar ‘21]
$\mathcal{G}(n, 1/2)$: $d \geq \Omega(\log n)$</td><td colspan="3">HARD
[Grigoriev ’01, Schoenebeck ’08]</td></tr>
<tr><td>Cutting Planes</td><td>OPEN</td><td colspan="2">OPEN
$\Theta(\log n)$-SAT
[Fleming, Pankratov, Pitassi, Robere ’17] [Hrubeš, Pudlák ’17]</td><td>Quasi-poly EASY
[Fleming, Göös, Impagliazzo, Pitassi, Robere, Tan, Wigderson ’21]
[Dadush, Tiwari ‘20]</td></tr>
</table>

Average-case hardness results

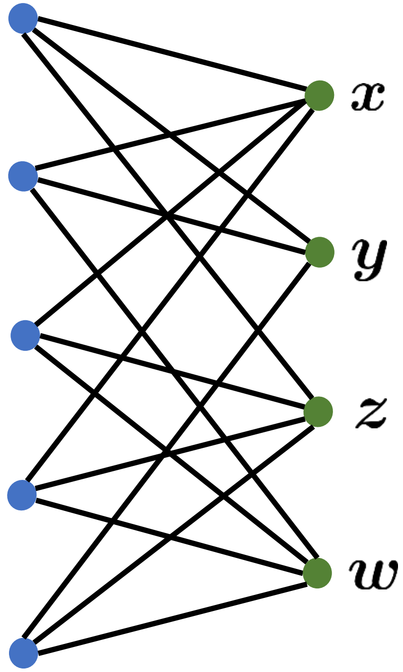
<table>
<tr><th></th><th>k-clique</th><th>k-coloring</th><th colspan="2">3-SAT</th><th>3-XOR</th></tr>
<tr><td>Tree-like Resolution</td><td>HARD (size $n^{\Omega(k)}$)
[Beyersdorff, Galesi, Lauria ‘11]</td><td rowspan="2">HARD
[Beame, Culberson, Mitchell, Moore ’05]</td><td colspan="3">HARD [Chvátal, Szemerédi ’88]
Improved [Ben-Sasson, Galesi '01] (size $\exp(n/\Delta^{1+\epsilon})$) $\Delta = m/n$</td></tr>
<tr><td>Resolution</td><td>OPEN
Some partial results*</td><td colspan="3">HARD [Chvátal, Szemerédi ’88] $\exp(n/\Delta^{2+\epsilon})$
Improved [Beame, Karp, Pitassi, Saks ’98], [Ben-Sasson ’01]</td></tr>
<tr><td rowspan="2">Polynomial Calculus</td><td rowspan="2">OPEN</td><td rowspan="2">OPEN</td><td>$\mathbb{F} \neq 2$</td><td colspan="2">HARD [Ben-Sasson, Impagliazzo ’99]</td></tr>
<tr><td>$\mathbb{F} = 2$</td><td>HARD [Alekhnovich, Razborov ’01]</td><td>EASY</td></tr>
<tr><td>Sum of Squares</td><td>OPEN
Some partial results**
$\mathcal{G}(n, 1/2)$: degree $= \Theta(\log n)$</td><td>OPEN
[Kothari, Manohar ‘21]
$\mathcal{G}(n, 1/2)$: $d \geq \Omega(\log n)$</td><td colspan="3">HARD
[Grigoriev ’01, Schoenebeck ’08]</td></tr>
<tr><td>Cutting Planes</td><td>OPEN</td><td>OPEN</td><td colspan="2">OPEN
$\Theta(\log n)$-SAT
[Fleming, Pankratov, Pitassi, Robere ’17] [Hrubeš, Pudlák ’17]</td><td>Quasi-poly EASY
[Fleming, Göös, Impagliazzo, Pitassi, Robere, Tan, Wigderson ’21]
[Dadush, Tiwari ‘20]</td></tr>
</table>

* [Beame, Impagliazzo, Sabharwal ‘01], [Pang ’21], [Atserias, Bonacina, **dR**, Lauria, Nordström, Razborov ‘18], [Lauria, Pudlák, Rödl, Thapen ‘13]
** [Meka, Potechin and Wigderson ‘15], ..., [Barak, Hopkins, Kelner, Kothari, Moitra, Potechin ’16]

Finding Structure in Randomness

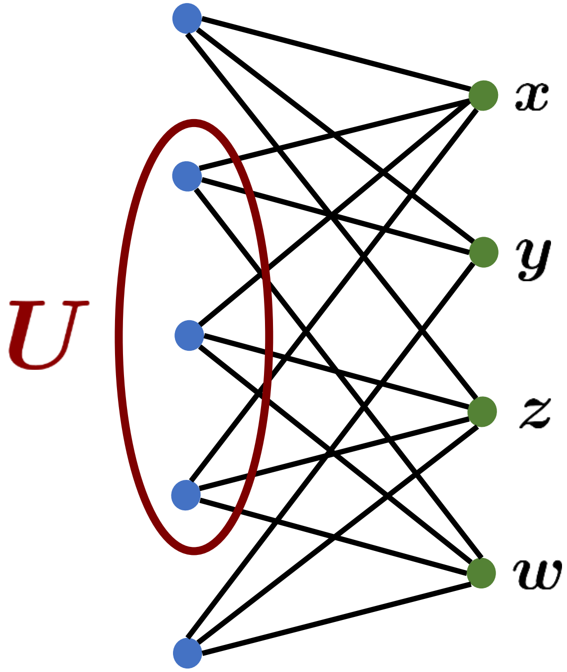
Random k-SAT (and k-XOR)

- Expansion G is (s, ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph



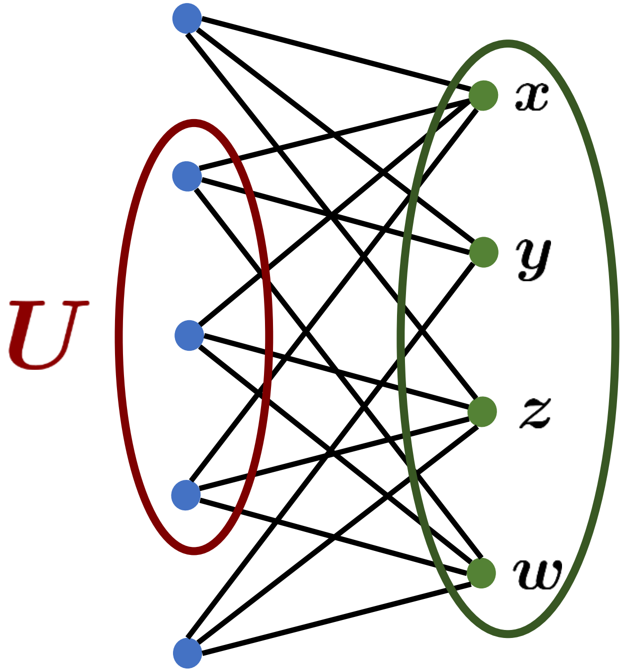
Random k-SAT (and k-XOR)

- Expansion G is (s, ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph



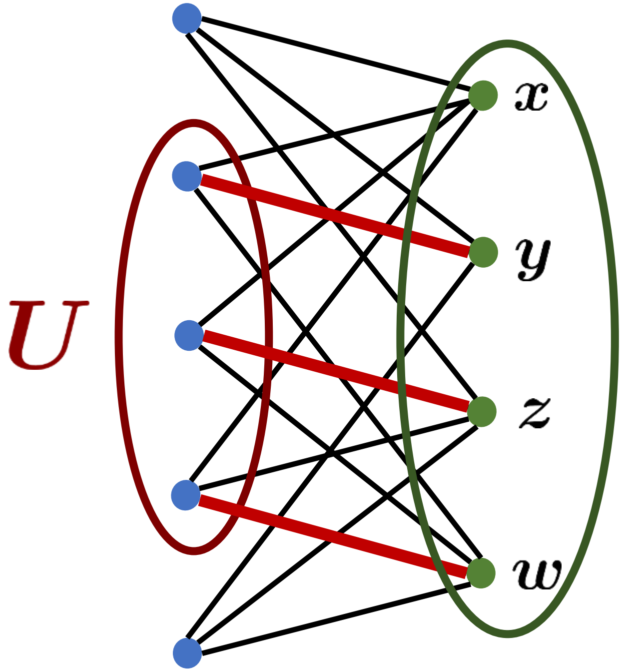
Random k-SAT (and k-XOR)

- Expansion G is (s, ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph



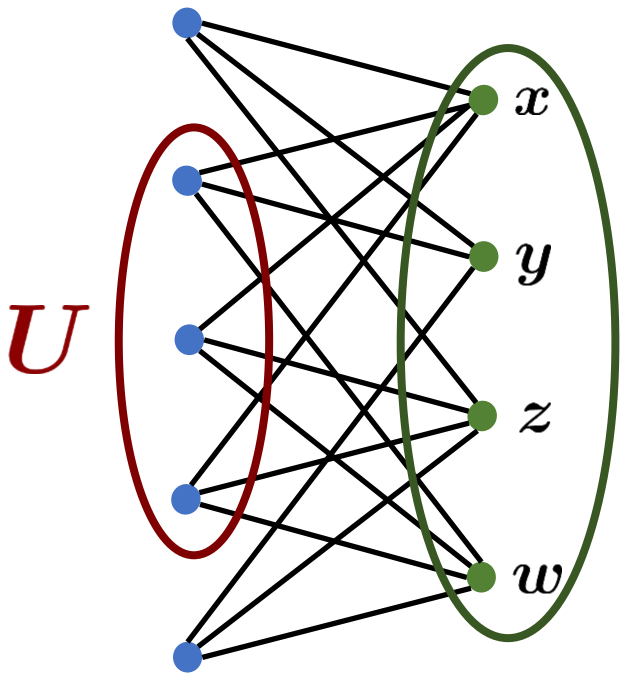
Random k-SAT (and k-XOR)

- Expansion G is (s, ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph



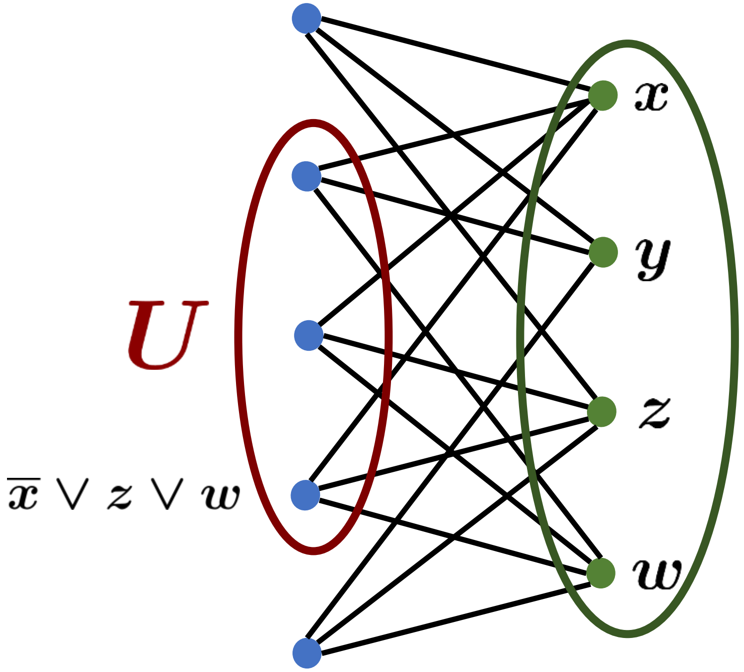
Random k-SAT (and k-XOR)

- Expansion G is (s, ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph



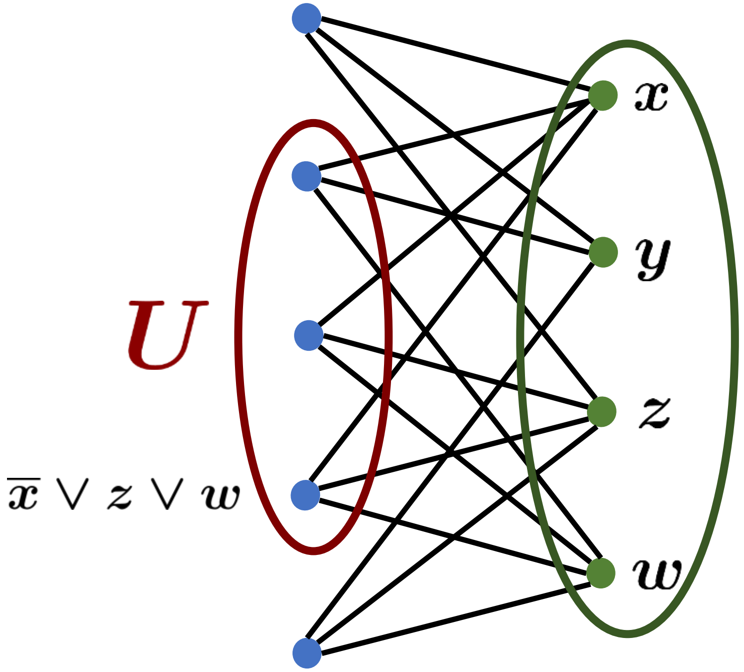
Random k-SAT (and k-XOR)

- Expansion G is (s,ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph



Random k-SAT (and k-XOR)

- Expansion G is (s, ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph

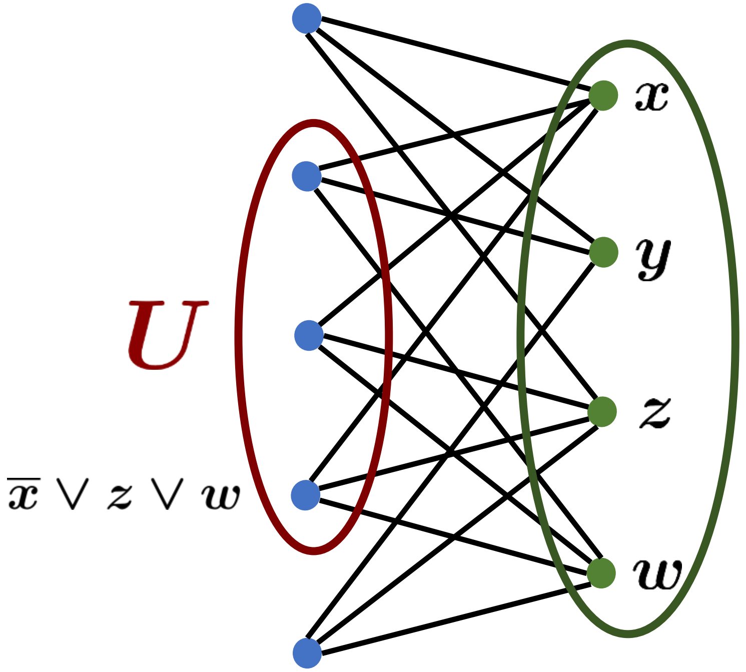


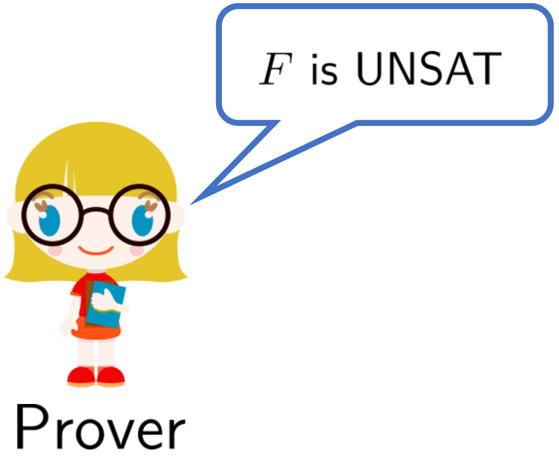
Prover

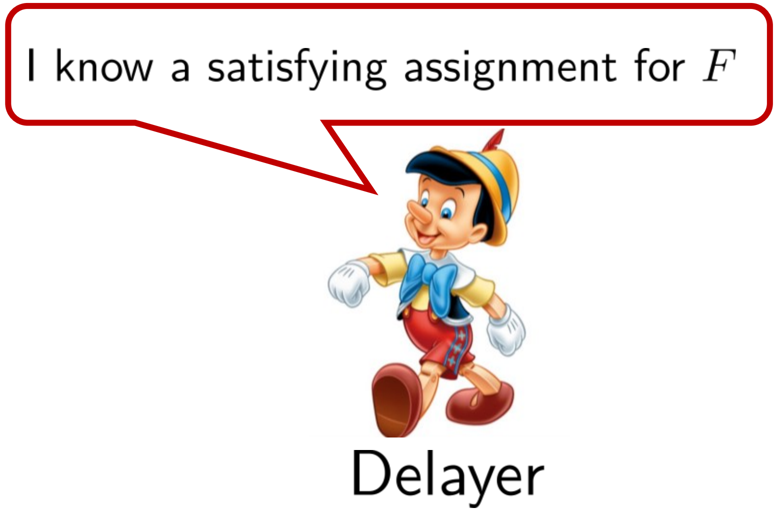
Delayer

Random k-SAT (and k-XOR)

- Expansion G is (s, ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph

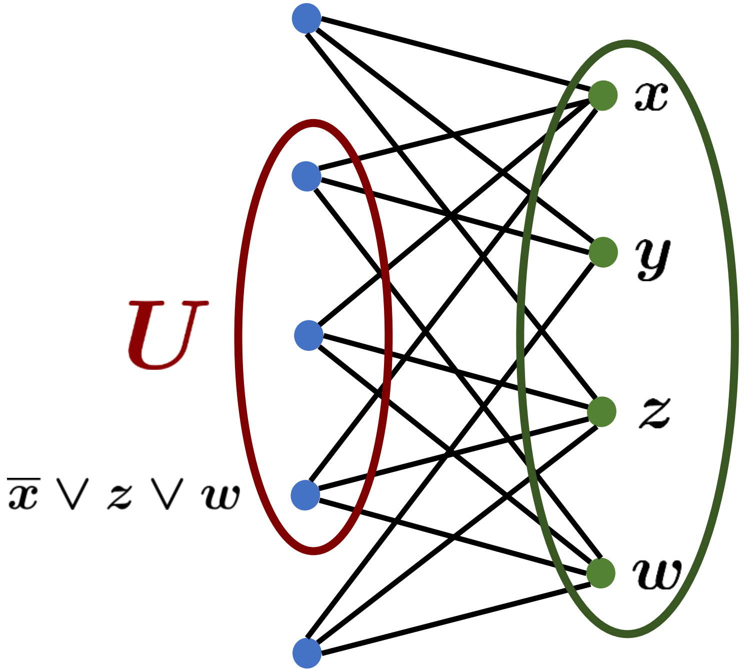


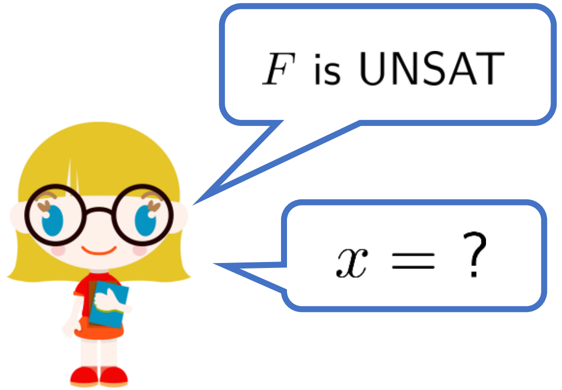




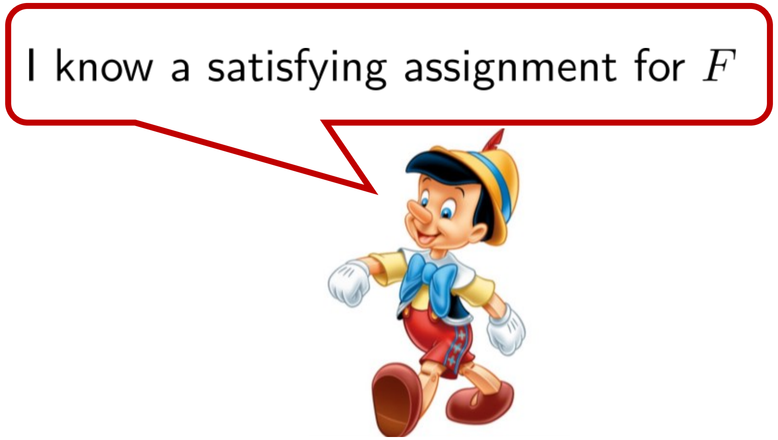
Random k-SAT (and k-XOR)

- Expansion G is (s, ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph





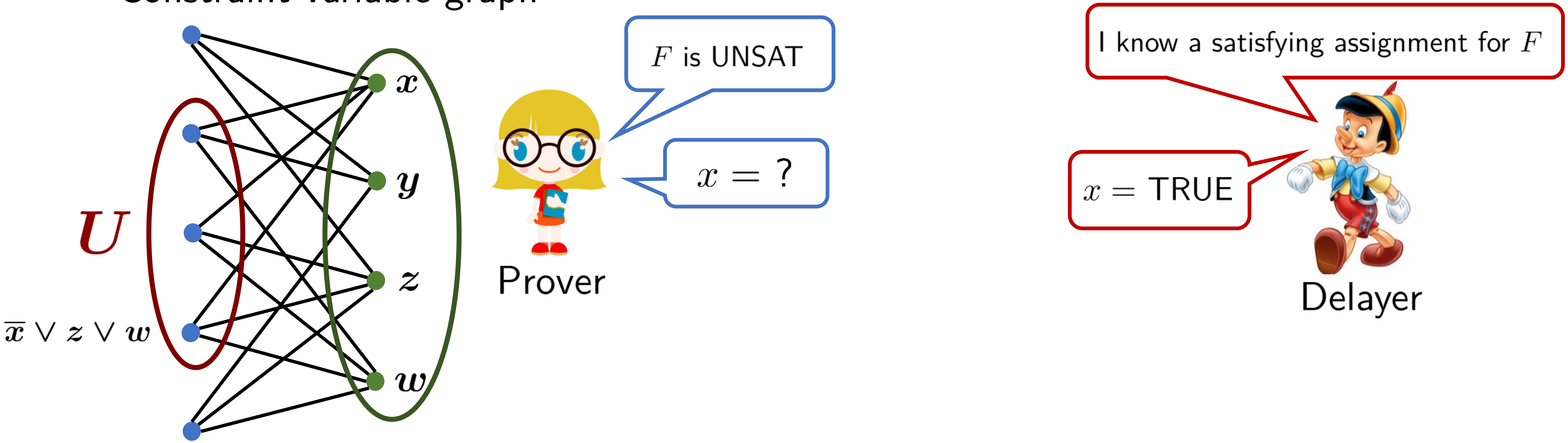
Prover



Delayer

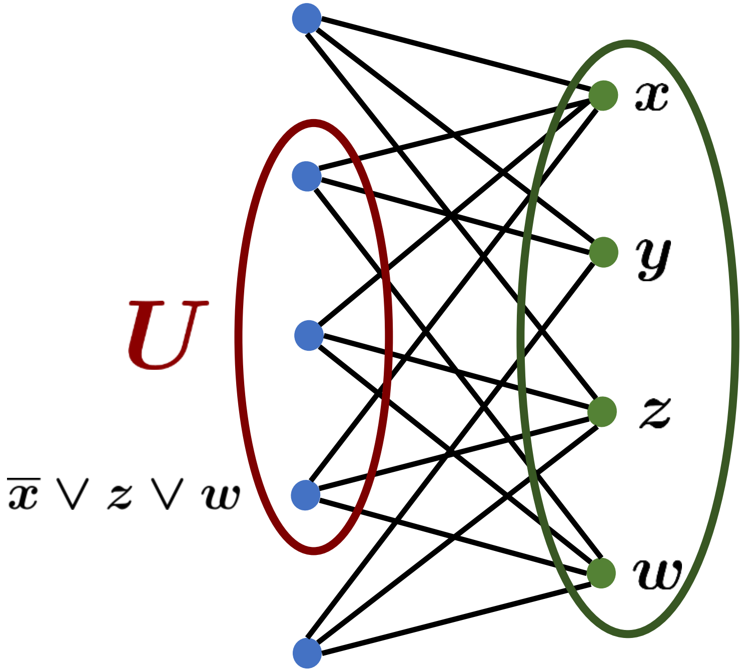
Random k-SAT (and k-XOR)

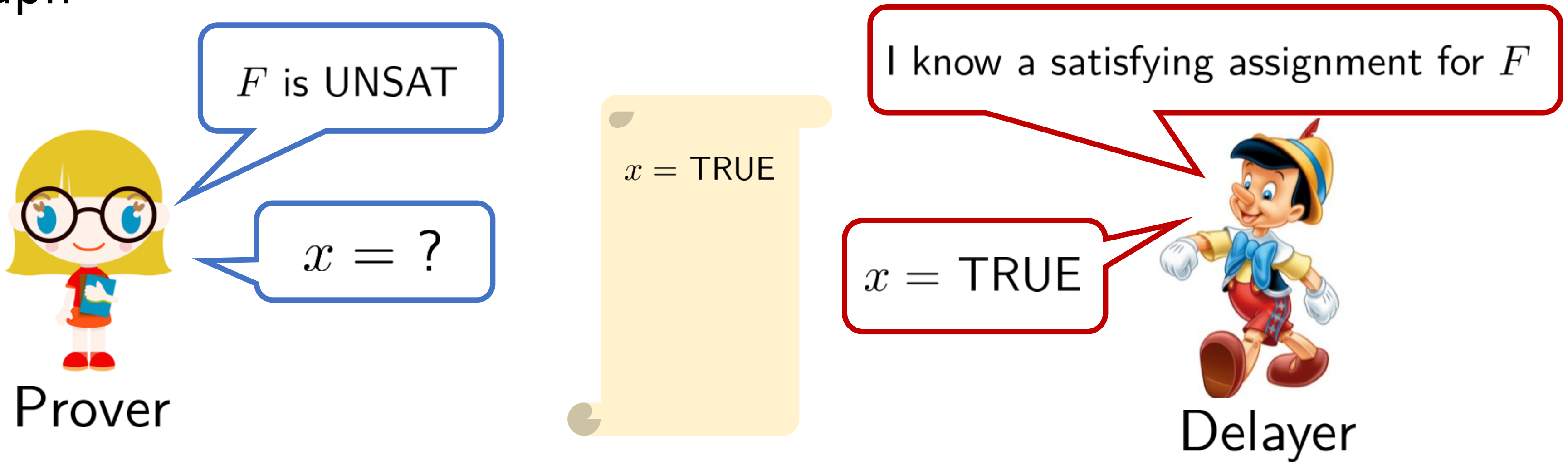
- Expansion G is (s, ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph



Random k-SAT (and k-XOR)

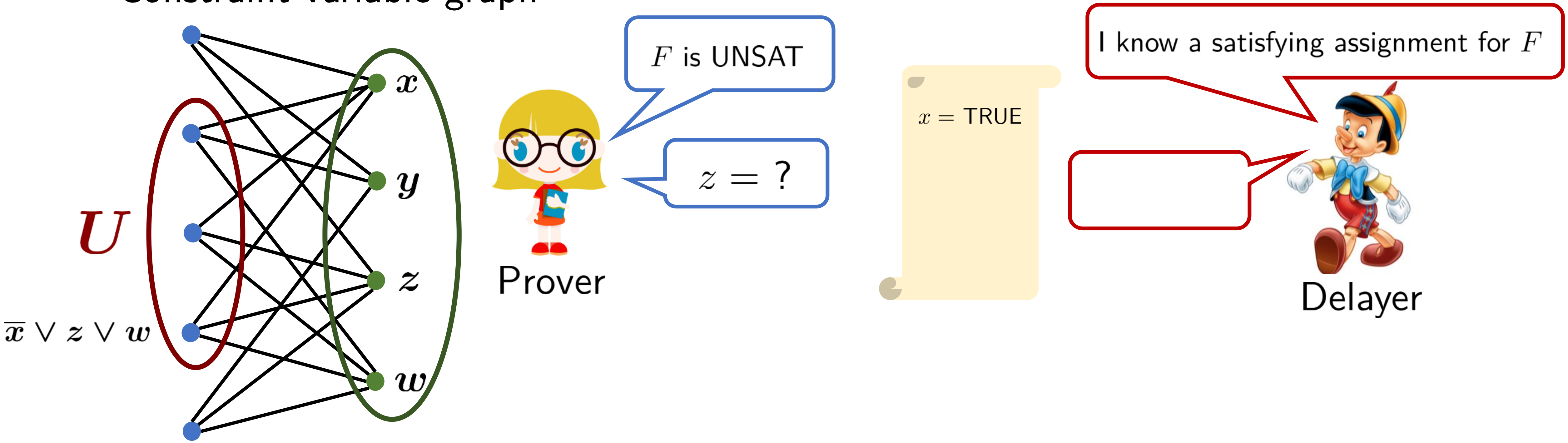
- Expansion G is (s,ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph





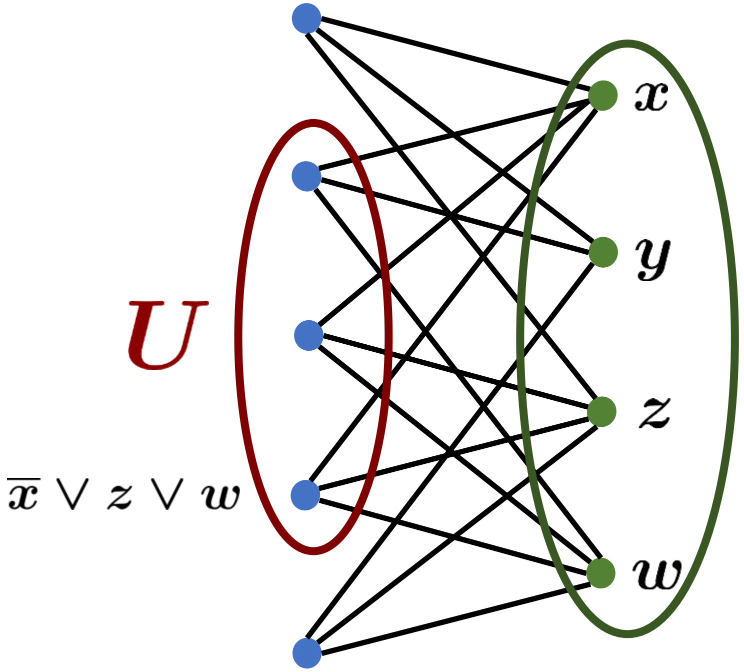
Random k-SAT (and k-XOR)

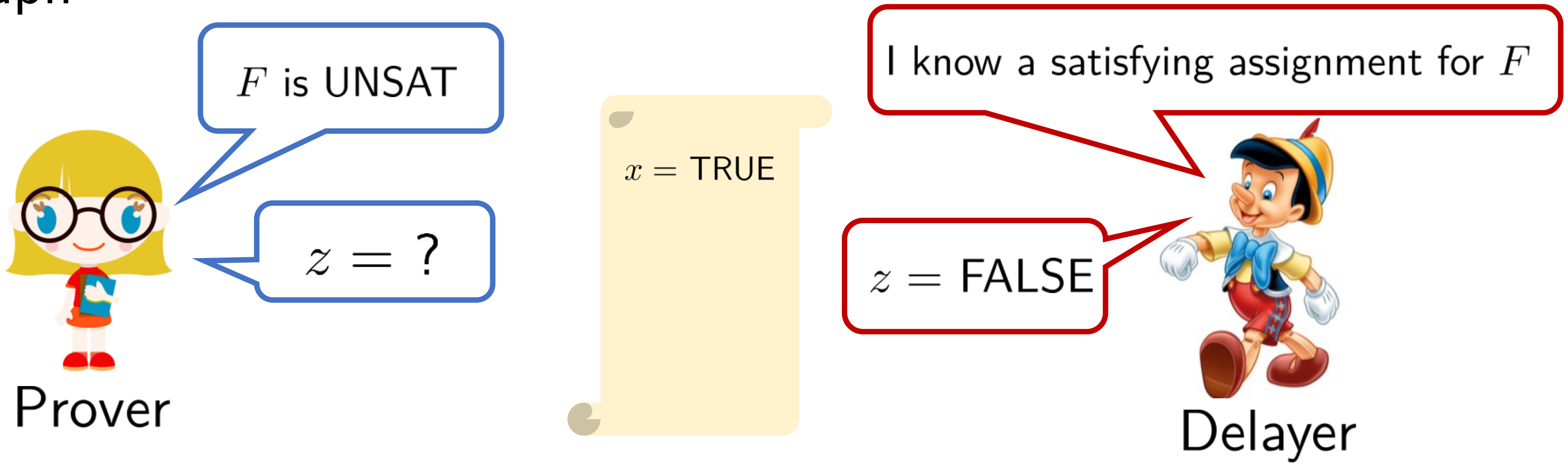
- Expansion G is (s,ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph



Random k-SAT (and k-XOR)

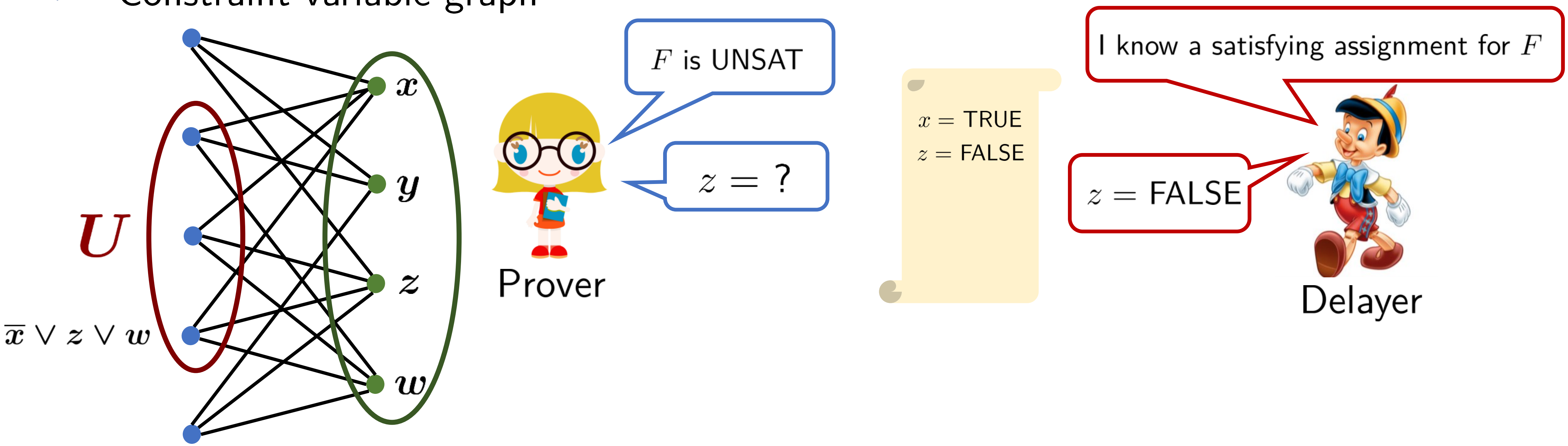
- Expansion G is (s,ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph





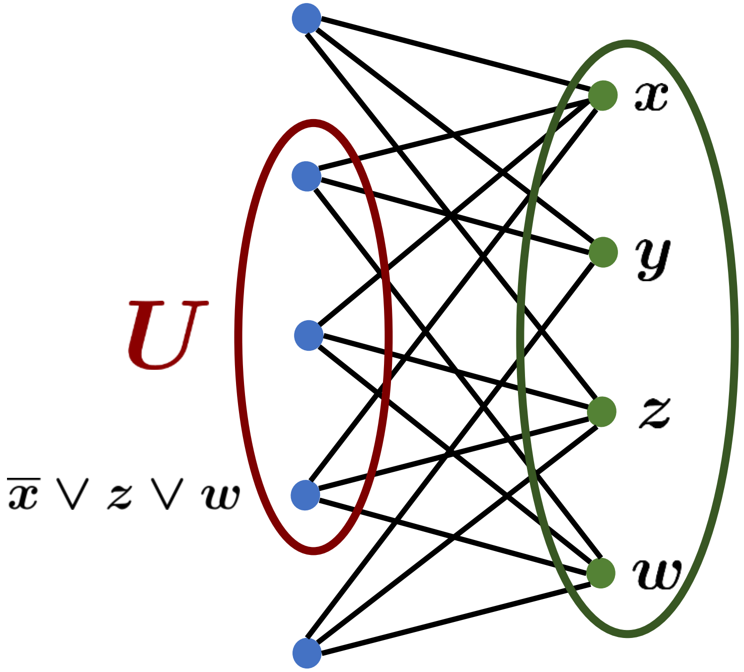
Random k-SAT (and k-XOR)

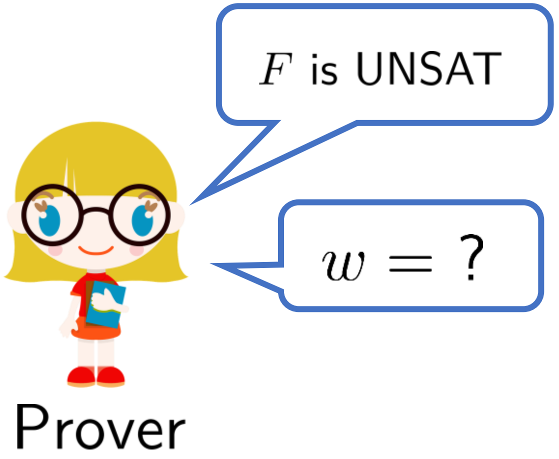
- Expansion G is (s,ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph

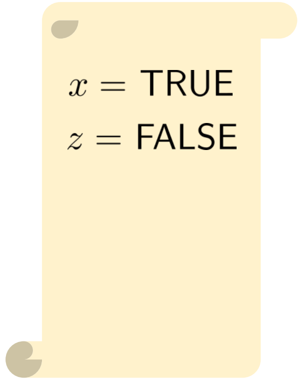


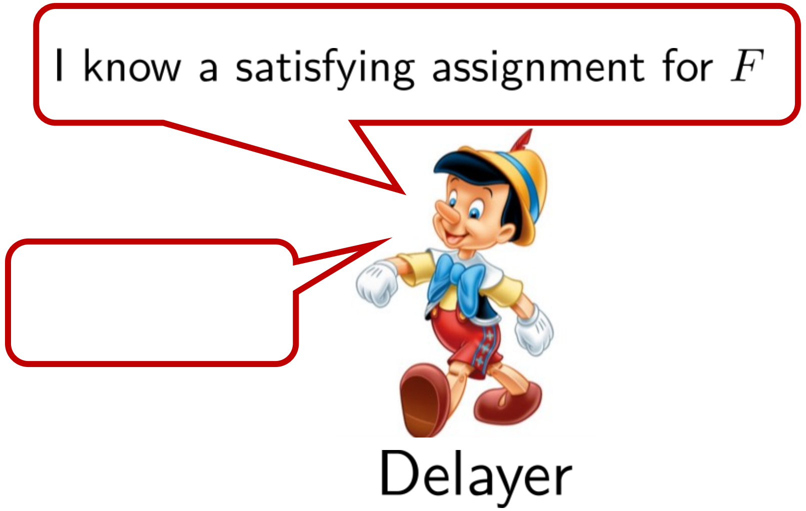
Random k-SAT (and k-XOR)

- Expansion G is (s, ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph



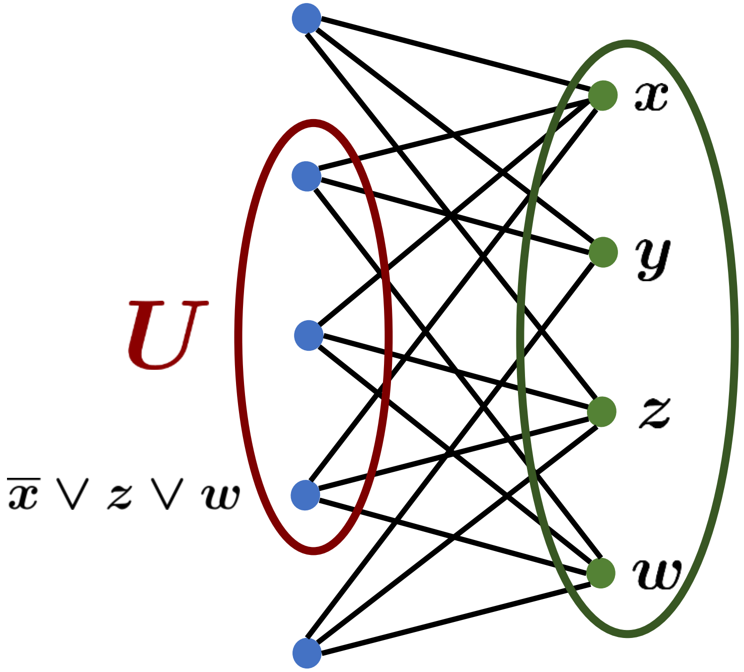


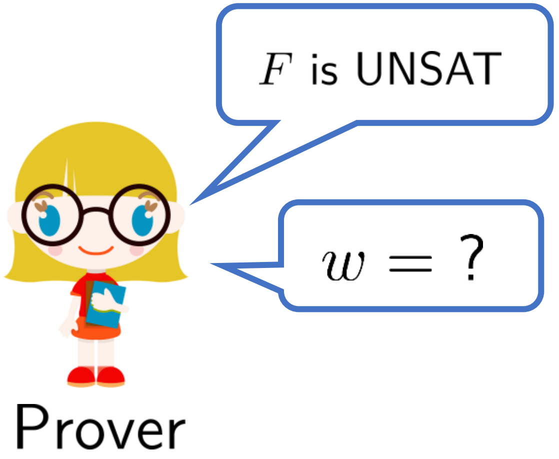


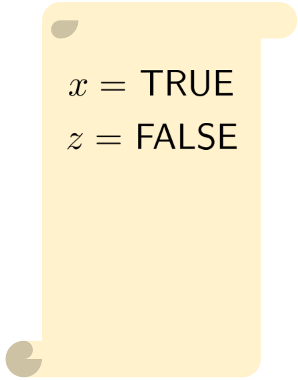


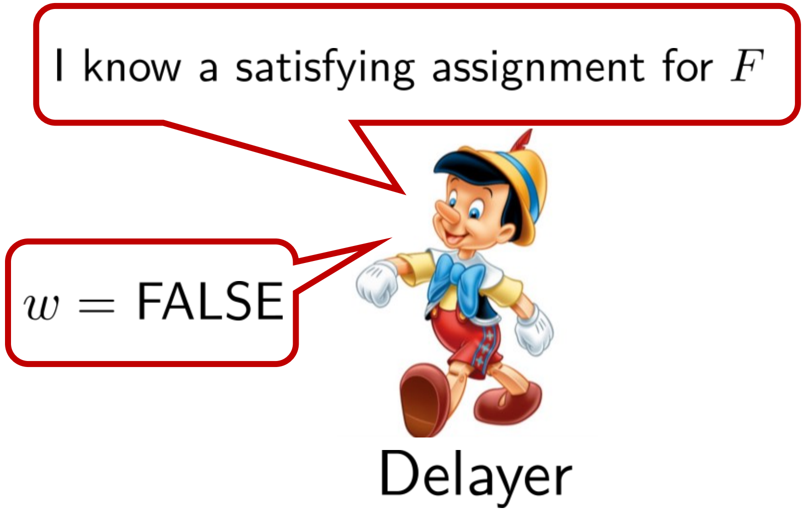
Random k-SAT (and k-XOR)

- Expansion G is (s,ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph



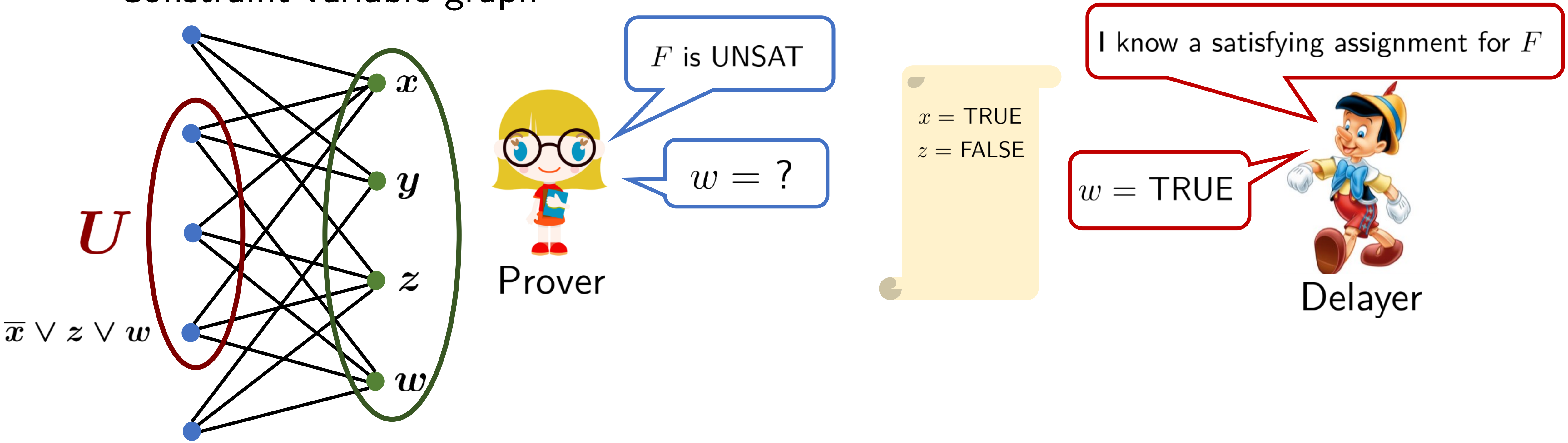






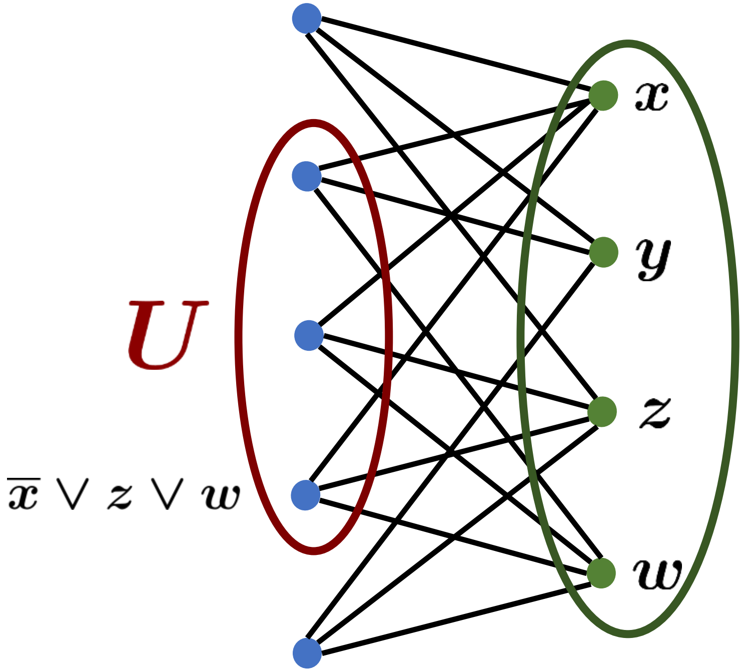
Random k-SAT (and k-XOR)

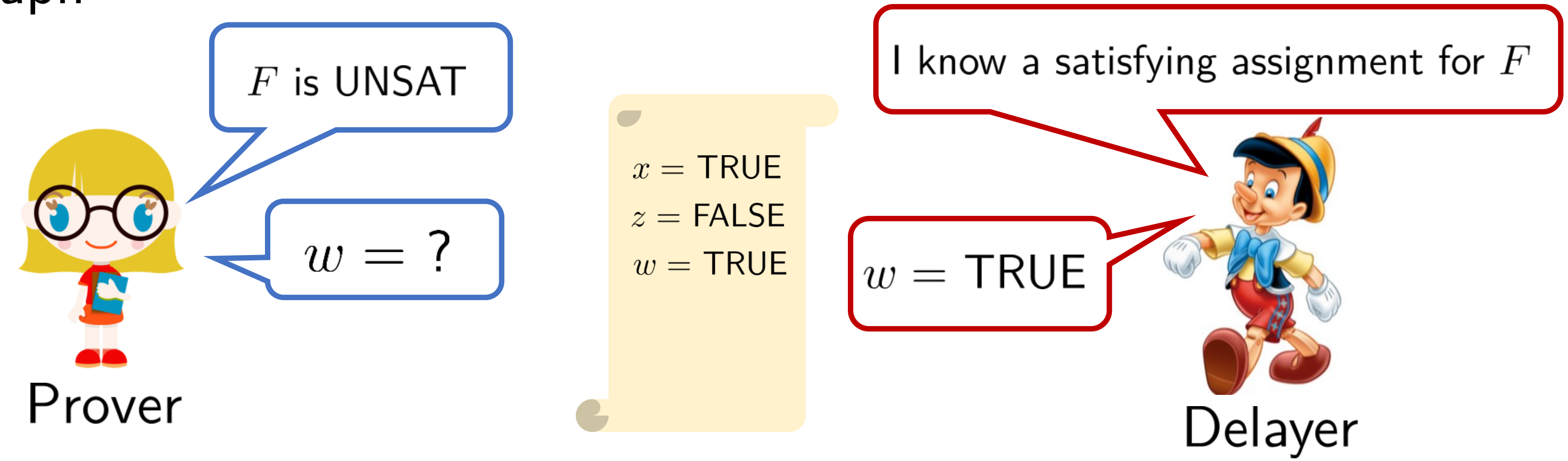
- Expansion G is (s, ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph



Random k-SAT (and k-XOR)

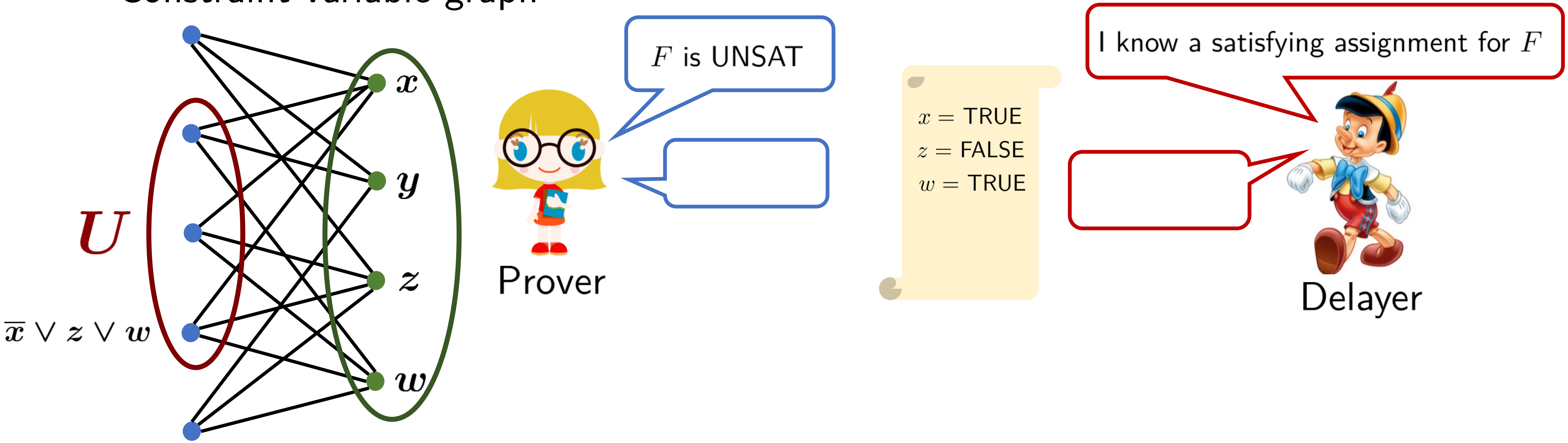
- Expansion G is (s, ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph





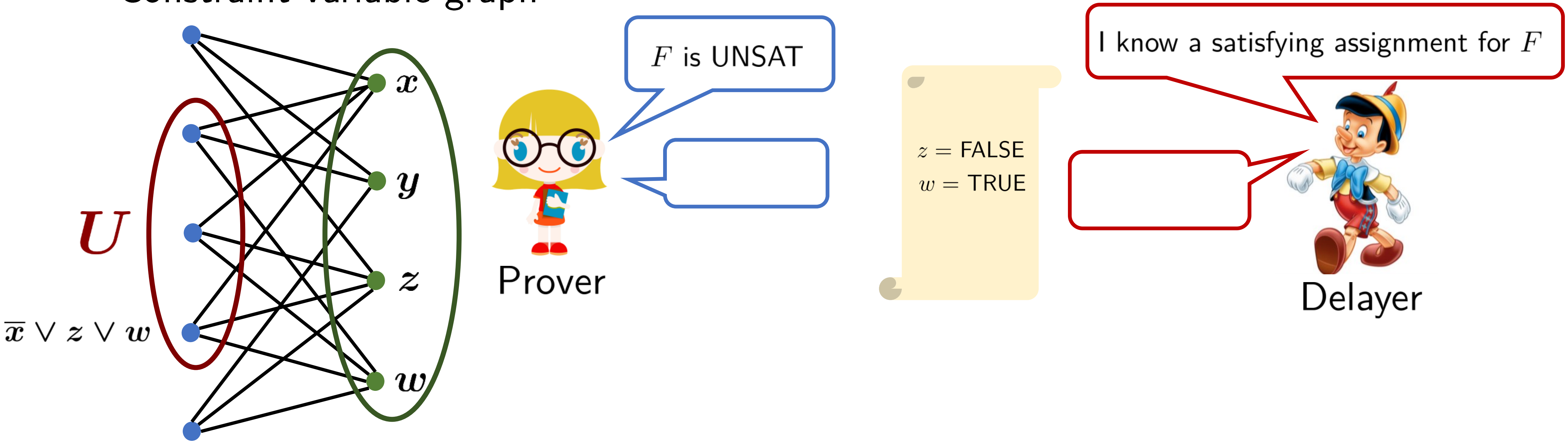
Random k-SAT (and k-XOR)

- Expansion G is (s,ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph



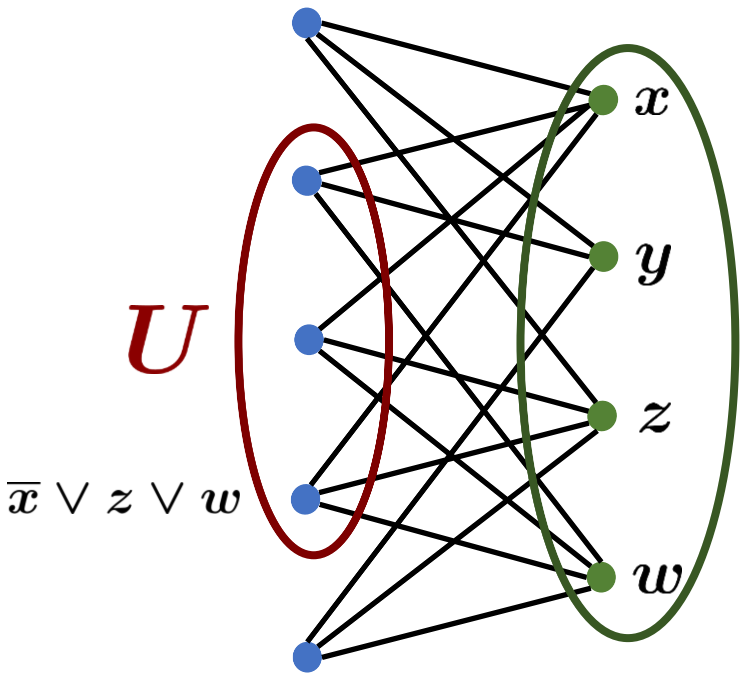
Random k-SAT (and k-XOR)

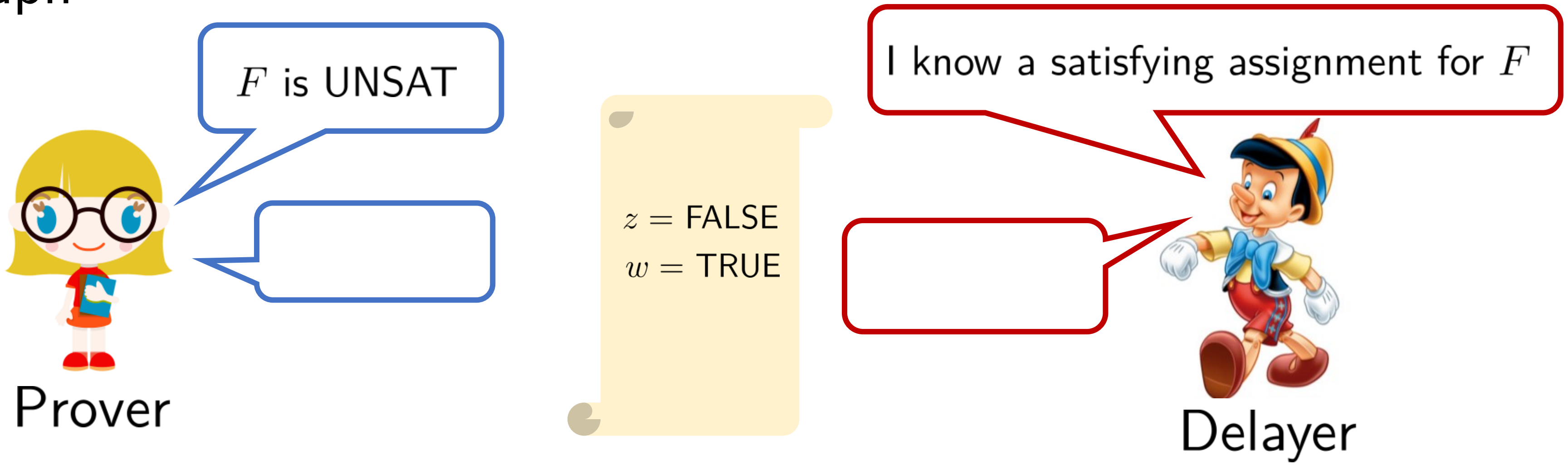
- Expansion G is (s, ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph



Random k-SAT (and k-XOR)

- Expansion G is (s,ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph

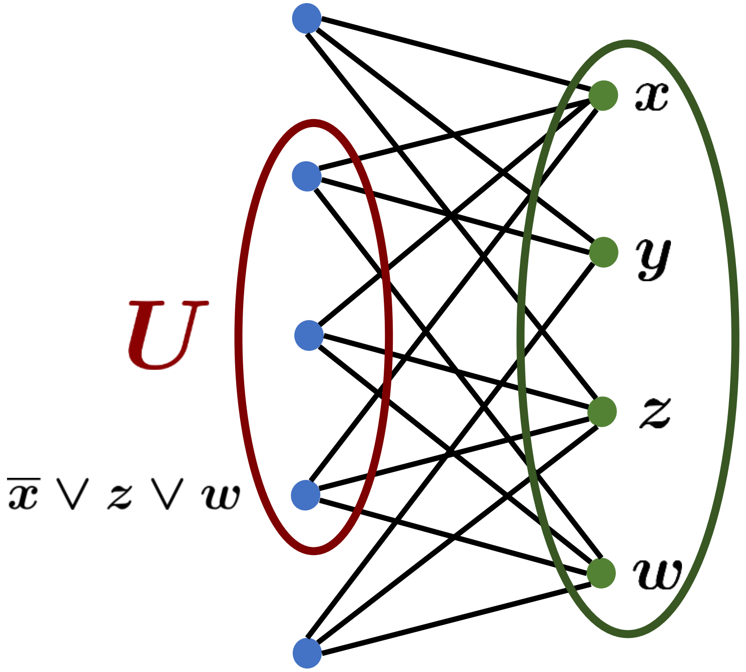


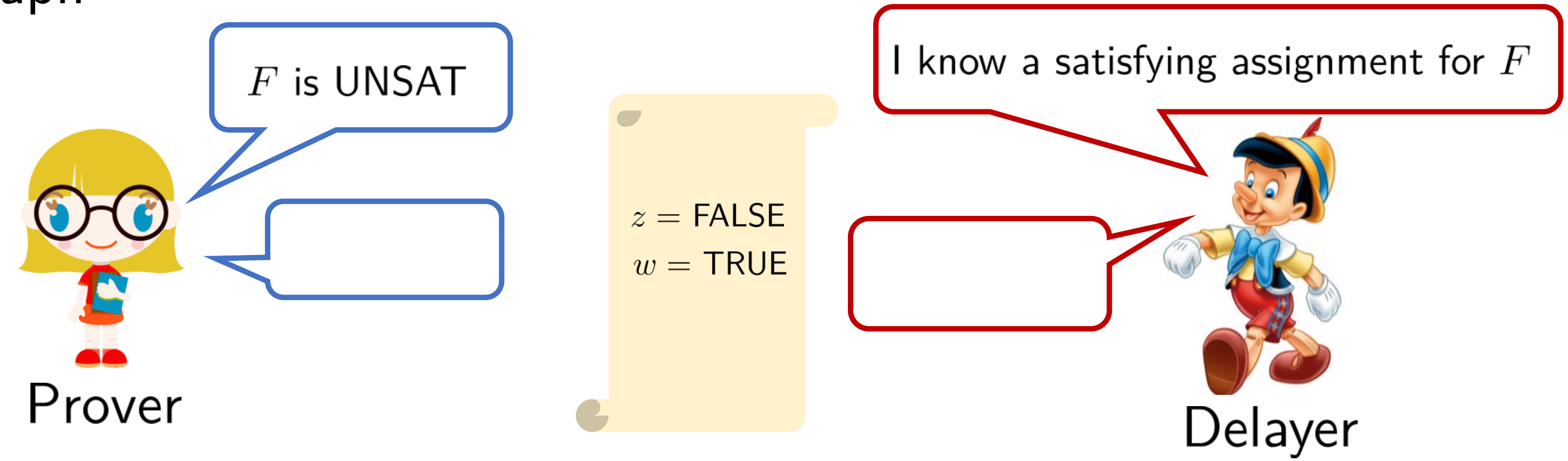


Delayer wins the r-game on F:
if with $\leq r$ lines in scroll, Prover cannot exhibit falsified clause

Random k-SAT (and k-XOR)

- Expansion G is (s, ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph



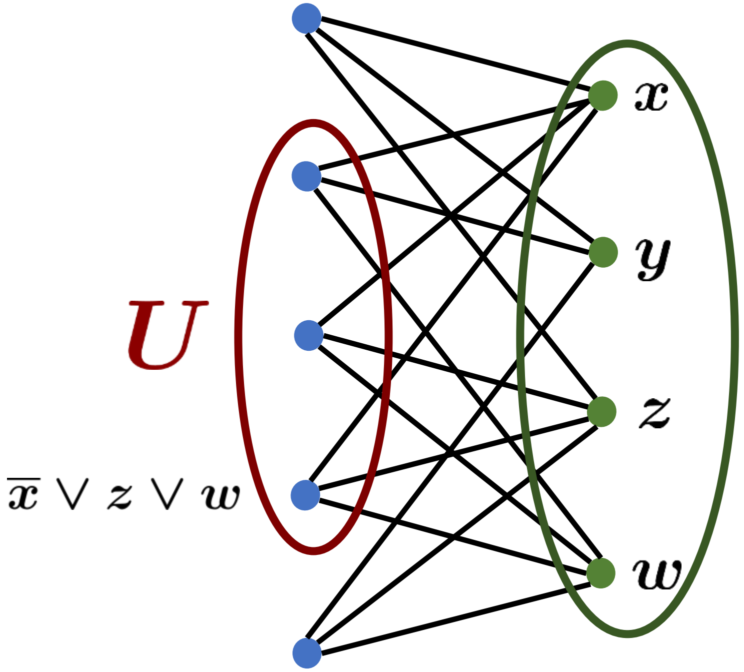


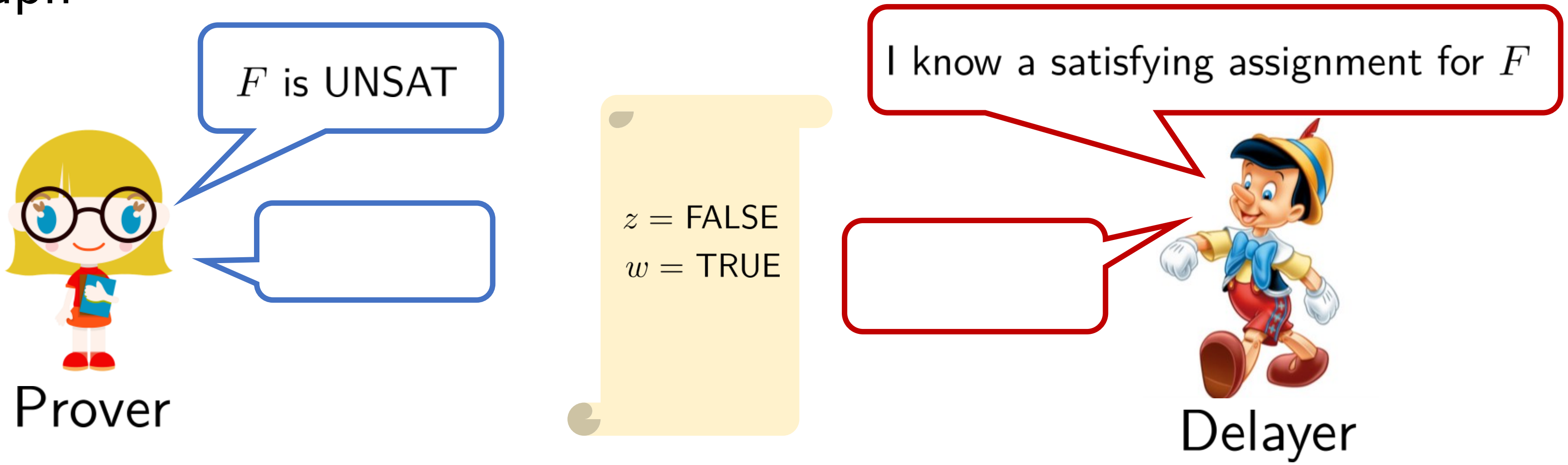
Delayer wins the r-game on F:
if with $\leq r$ lines in scroll, Prover cannot exhibit falsified clause

If Delayer wins the r-game on F, then resolution requires width r to refute F

Random k-SAT (and k-XOR)

- Expansion G is (s,ϵ)-bipartite expander if $\forall U \subseteq V$: $|U| \leq s \Rightarrow |N(U)| \geq (1+\epsilon)|U|$
- Constraint-variable graph





Delayer wins the r-game on F:
if with $\leq r$ lines in scroll, Prover cannot exhibit falsified clause

If Delayer wins the r-game on F, then resolution requires width r to refute F

Lemma 1. If G is (s,ϵ)-bipartite expander Delayer wins if $r \leq \epsilon s/(d+\epsilon)$

Lemma 2. W.h.p. constraint-variable graph of random 3-CNF is a good expander

Coloring hard for resolution [Beame, Culberson, Mitchell, Moore '05]

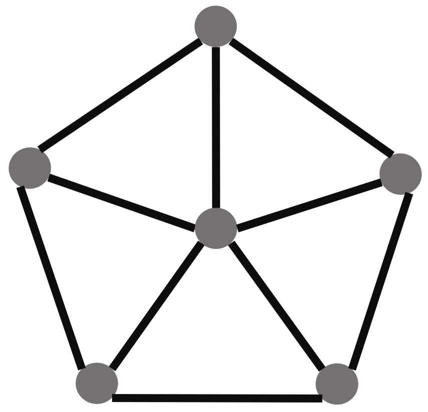
Coloring hard for resolution [Beame, Culberson, Mitchell, Moore '05]

$s =$ maximum number s.t. any s-vertex $H \subseteq G$ is k-colorable

Coloring hard for resolution [Beame, Culberson, Mitchell, Moore '05]

$s =$ maximum number s.t. any s-vertex $H \subseteq G$ is k-colorable

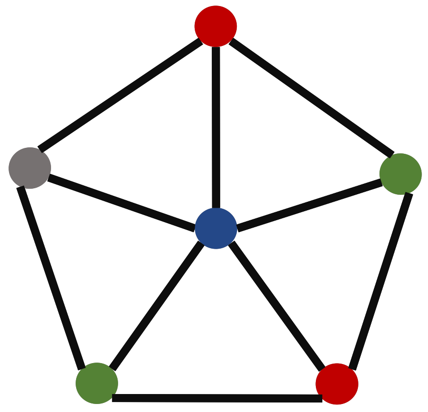
$\boldsymbol{k = 3}$



Coloring hard for resolution [Beame, Culberson, Mitchell, Moore '05]

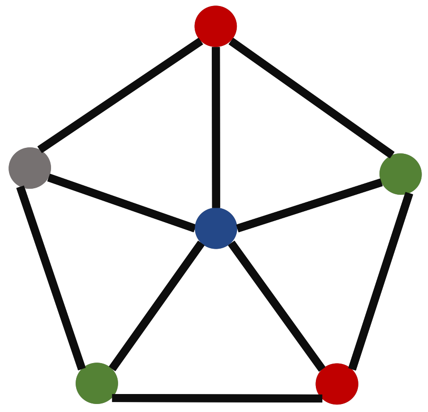
$s =$ maximum number s.t. any s-vertex $H \subseteq G$ is k-colorable

$\boldsymbol{k = 3}$



Coloring hard for resolution [Beame, Culberson, Mitchell, Moore '05]

$s =$ maximum number s.t. any s-vertex $H \subseteq G$ is k-colorable

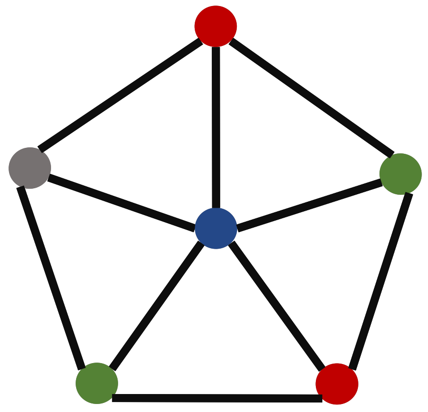


Coloring hard for resolution [Beame, Culberson, Mitchell, Moore '05]

$s =$ maximum number s.t. any s-vertex $H \subseteq G$ is k-colorable

$\beta_k(H) = \#$ of vertices in H of degree between 1 and $k-1$

$\boldsymbol{k = 3}$

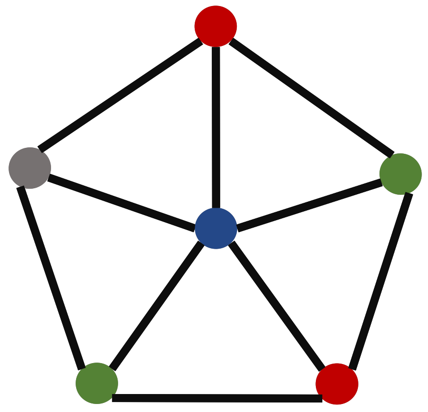


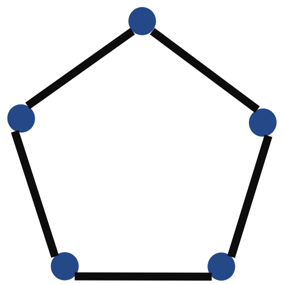
$\boldsymbol{s = 5}$

Coloring hard for resolution [Beame, Culberson, Mitchell, Moore '05]

$s =$ maximum number s.t. any s-vertex $H \subseteq G$ is k-colorable

$\beta_k(H) = \#$ of vertices in H of degree between 1 and $k-1$



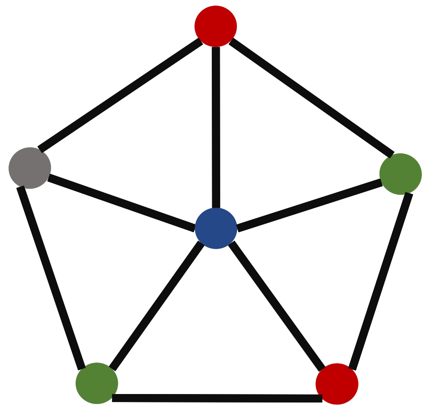


Coloring hard for resolution [Beame, Culberson, Mitchell, Moore '05]

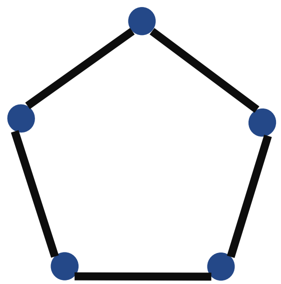
$s =$ maximum number s.t. any s-vertex $H \subseteq G$ is k-colorable

$\beta_k(H) = \#$ of vertices in H of degree between 1 and $k-1$

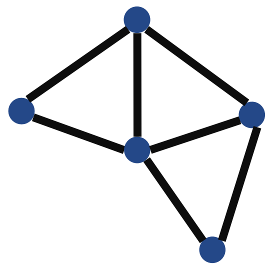
$\boldsymbol{k = 3}$



$\boldsymbol{s = 5}$



$\beta_k(H) = 5$



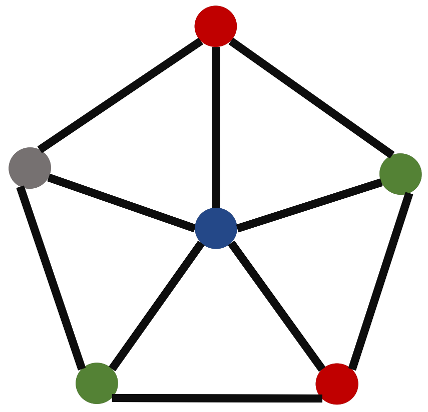
$\beta_k(H) = 2$

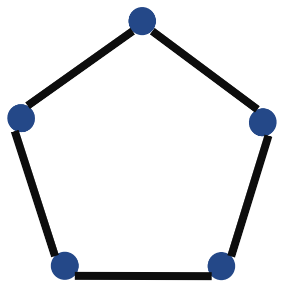
Coloring hard for resolution [Beame, Culberson, Mitchell, Moore '05]

$s =$ maximum number s.t. any s-vertex $H \subseteq G$ is k-colorable

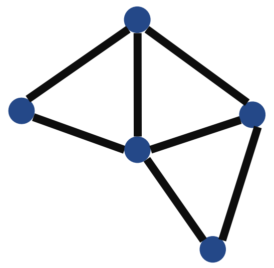
$\beta_k(H) = \#$ of vertices in H of degree between 1 and $k-1$

Subcritical k-expansion $e_k(G) = \max_{2 \le t \le s} \min_{\substack{H \subseteq G \\ H \text{ connected} \\ t/2 \le V(H) \le t}} \beta_k(H)$





$\beta_k(H) = 5$



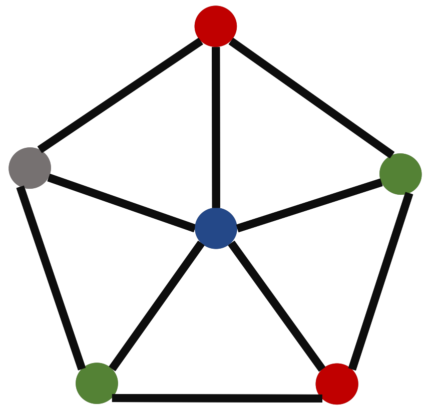
$\beta_k(H) = 2$

Coloring hard for resolution [Beame, Culberson, Mitchell, Moore '05]

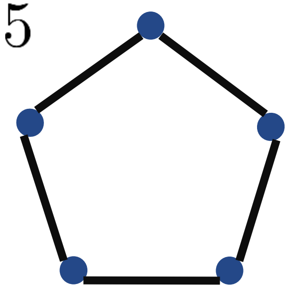
$s =$ maximum number s.t. any s-vertex $H \subseteq G$ is k-colorable

$\beta_k(H) = \#$ of vertices in H of degree between 1 and $k-1$

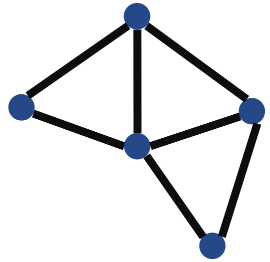
Subcritical k-expansion $e_k(G) = \max_{2 \le t \le s} \min_{\substack{H \subseteq G \\ H \text{ connected} \\ t/2 \le V(H) \le t}} \beta_k(H)$



$t = 5$



$\beta_k(H) = 5$



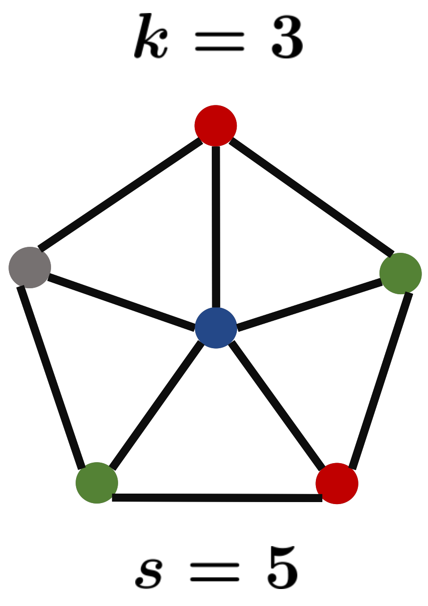
$\beta_k(H) = 2$

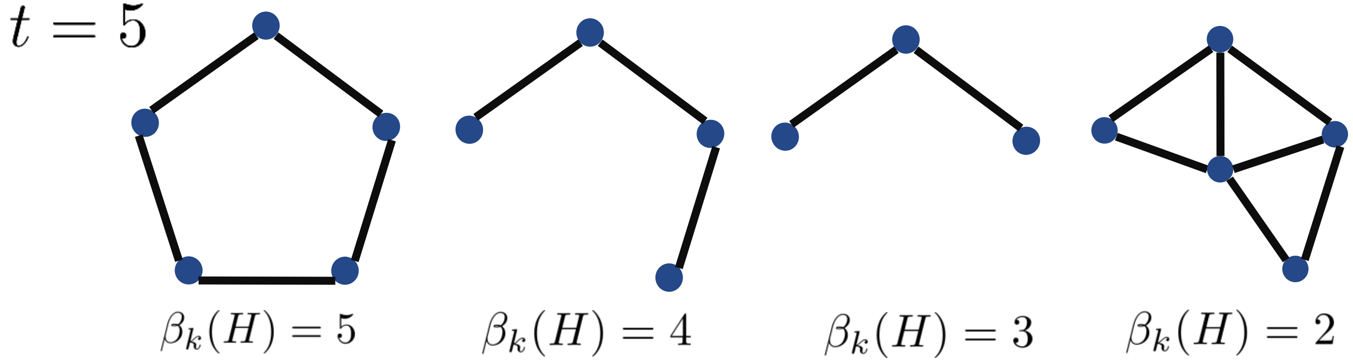
Coloring hard for resolution [Beame, Culberson, Mitchell, Moore '05]

$s =$ maximum number s.t. any s-vertex $H \subseteq G$ is k-colorable

$\beta_k(H) = \#$ of vertices in H of degree between 1 and $k-1$

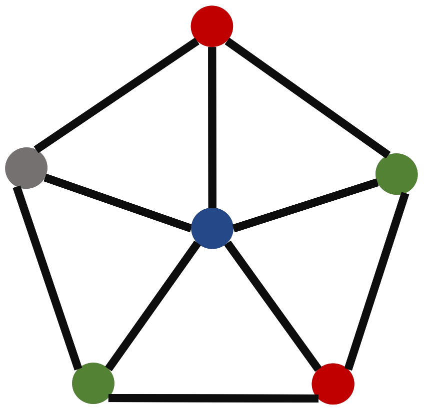
Subcritical k-expansion $$e_k(G) = \max_{2 \le t \le s} \min_{\substack{H \subseteq G \\ H \text{ connected} \\ t/2 \le V(H) \le t}} \beta_k(H)$$





Coloring hard for resolution [Beame, Culberson, Mitchell, Moore '05]

$\boldsymbol{k = 3}$



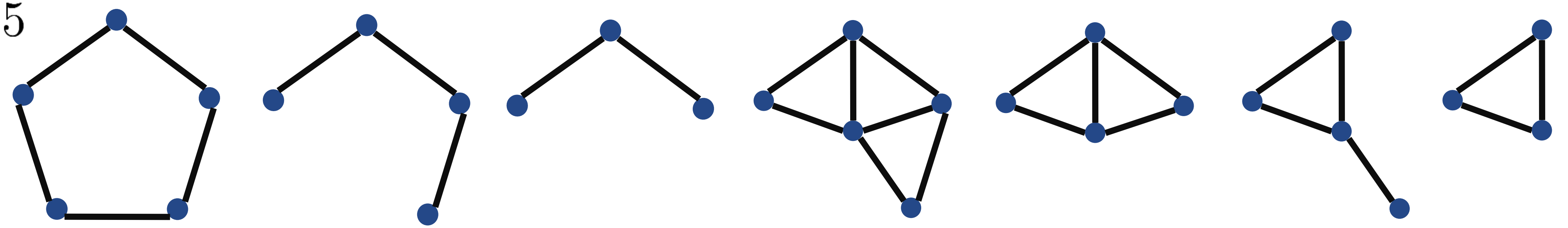
$\boldsymbol{s = 5}$

$s =$ maximum number s.t. any s-vertex $H \subseteq G$ is k-colorable

$\beta_k(H) = \#$ of vertices in H of degree between 1 and $k-1$

Subcritical k-expansion $\displaystyle e_k(G) = \max_{2 \le t \le s} \min_{\substack{H \subseteq G \\ H \text{ connected} \\ t/2 \le V(H) \le t}} \beta_k(H)$

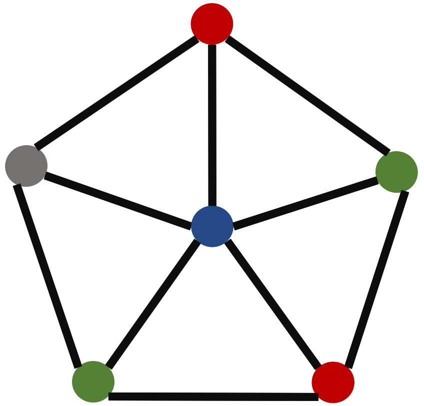
$t = 5$



$\beta_k(H) = 5$ $\beta_k(H) = 4$ $\beta_k(H) = 3$ $\beta_k(H) = 2$ $\beta_k(H) = 2$ $\beta_k(H) = 3$ $\beta_k(H) = 3$

Coloring hard for resolution [Beame, Culberson, Mitchell, Moore '05]

$\boldsymbol{k = 3}$



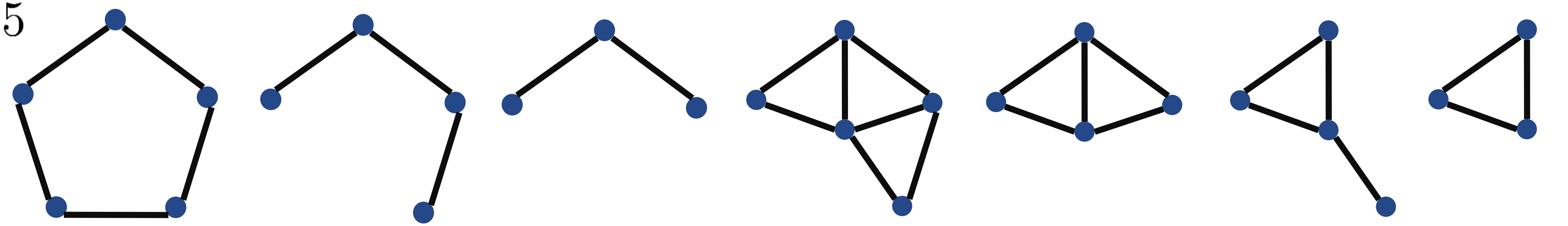
$\boldsymbol{s = 5}$

$s =$ maximum number s.t. any s-vertex $H \subseteq G$ is k-colorable

$\beta_k(H) = \#$ of vertices in H of degree between 1 and $k-1$

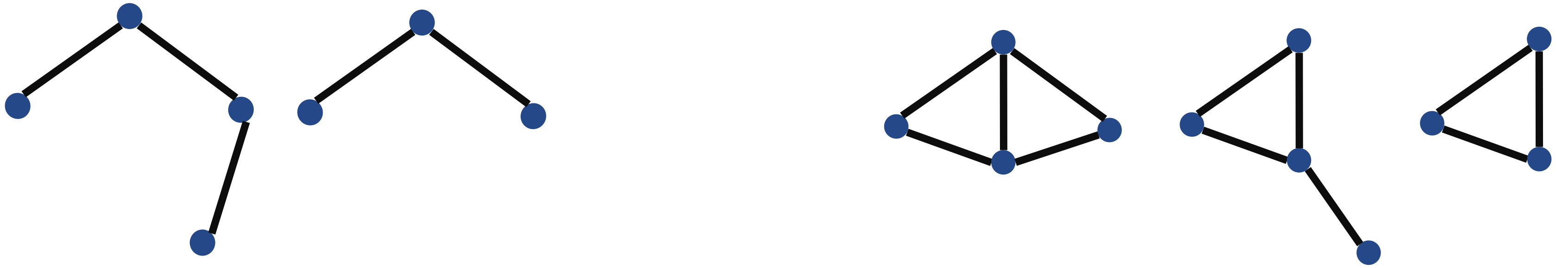
Subcritical k-expansion $\displaystyle e_k(G) = \max_{2 \leq t \leq s} \min_{\substack{H \subseteq G \\ H \text{ connected} \\ t/2 \leq V(H) \leq t}} \beta_k(H)$

$t = 5$



$\beta_k(H) = 5$ $\beta_k(H) = 4$ $\beta_k(H) = 3$ $\beta_k(H) = 2$ $\beta_k(H) = 2$ $\beta_k(H) = 3$ $\beta_k(H) = 3$

$t = 4$



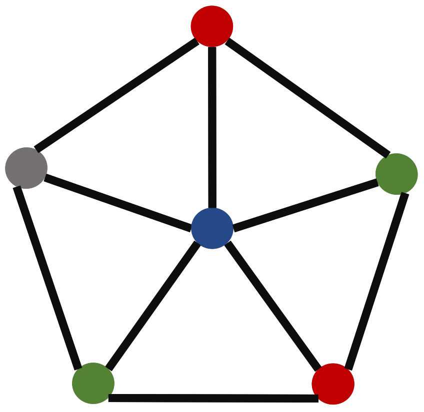
Coloring hard for resolution [Beame, Culberson, Mitchell, Moore '05]

$s =$ maximum number s.t. any s-vertex $H \subseteq G$ is k-colorable

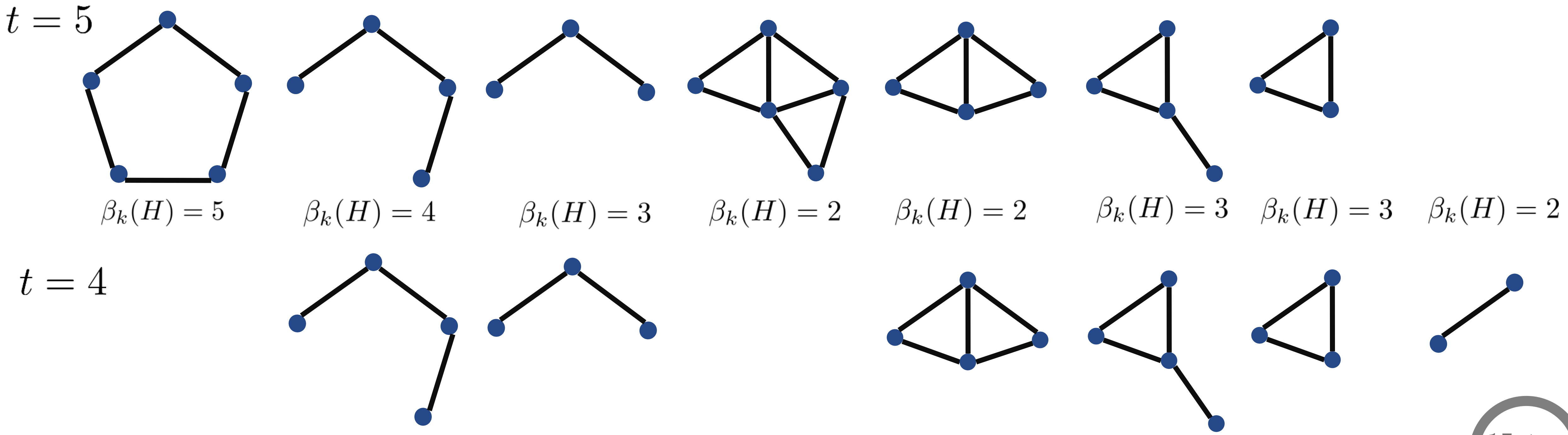
$\beta_k(H) = \#$ of vertices in H of degree between 1 and $k-1$

Subcritical k-expansion $e_k(G) = \max_{2 \le t \le s} \min_{\substack{H \subseteq G \\ H \text{ connected} \\ t/2 \le V(H) \le t}} \beta_k(H)$

$\boldsymbol{k = 3}$

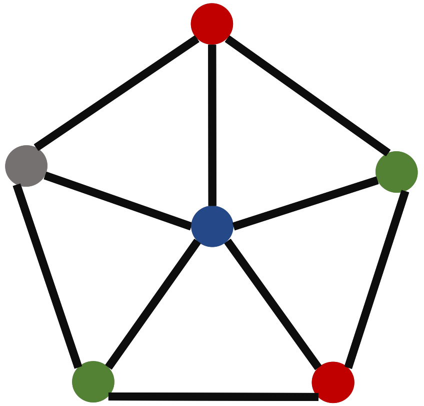


$\boldsymbol{s = 5}$



Coloring hard for resolution [Beame, Culberson, Mitchell, Moore '05]

$\boldsymbol{k = 3}$



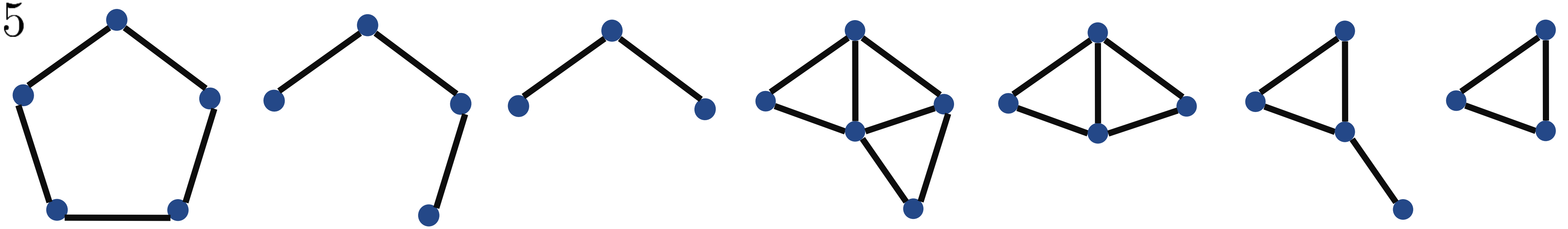
$\boldsymbol{s = 5}$

$s =$ maximum number s.t. any s-vertex $H \subseteq G$ is k-colorable

$\beta_k(H) = \#$ of vertices in H of degree between 1 and $k-1$

Subcritical k-expansion $\displaystyle e_k(G) = \max_{2 \leq t \leq s} \min_{\substack{H \subseteq G \\ H \text{ connected} \\ t/2 \leq V(H) \leq t}} \beta_k(H)$

$t = 5$



$\beta_k(H) = 5$ $\beta_k(H) = 4$ $\beta_k(H) = 3$ $\beta_k(H) = 2$ $\beta_k(H) = 2$ $\beta_k(H) = 3$ $\beta_k(H) = 3$ $\beta_k(H) = 2$

Lemma 1. Resolution width of refuting Color$(G, k) \geq e_k(G)$

Lemma 2. Let $G \sim \mathcal{G}_{n,m}$ for $m = \Delta n$. W.h.p. $e_k(G) \geq \epsilon_k n / \Delta^{1+2/(k-2)}$

Clique formula

Clique(G, k) for $G \sim \mathcal{G}(n, p)$ and p close to k-clique threshold

Clique formula

Clique(G, k) for $G \sim \mathcal{G}(n, p)$ and p close to k-clique threshold

- $2^{k/(n-\Delta)^6}$-hard for resolution for very dense graph and large $k \geq n^{5/6}$
  $\Delta =$ average degree
  [Beame, Impagliazzo, Sabharwal '01]

Clique formula

Clique(G, k) for $G \sim \mathcal{G}(n, p)$ and p close to k-clique threshold

- $2^{k/(n-\Delta)^6}$-hard for resolution for very dense graph and large $k \geq n^{5/6}$
  $\Delta =$ average degree [Beame, Impagliazzo, Sabharwal '01]
- $2^{\Omega(k^{1-\epsilon})}$-hard for resolution for $k \leq n^{1/3}$ [Pang '21]

Clique formula

Clique(G, k) for $G \sim \mathcal{G}(n, p)$ and p close to k-clique threshold

- $2^{k/(n-\Delta)^6}$-hard for resolution for very dense graph and large $k \geq n^{5/6}$
  $\Delta =$ average degree [Beame, Impagliazzo, Sabharwal '01]
- $2^{\Omega(k^{1-\epsilon})}$-hard for resolution for $k \leq n^{1/3}$ [Pang '21]
- $n^{\Omega(k)}$-hard for tree-like resolution [Lauria, Pudlák, Rödl, Thapen '13]

Clique formula

Clique(G, k) for $G \sim \mathcal{G}(n, p)$ and p close to k-clique threshold

- $2^{k/(n-\Delta)^6}$-hard for resolution for very dense graph and large $k \geq n^{5/6}$
  $\Delta =$ average degree [Beame, Impagliazzo, Sabharwal '01]
- $2^{\Omega(k^{1-\epsilon})}$-hard for resolution for $k \leq n^{1/3}$ [Pang '21]
- $n^{\Omega(k)}$-hard for tree-like resolution [Lauria, Pudlák, Rödl, Thapen '13]
- $n^{\Omega(k)}$-hard for regular resolution [Atserias, Bonacina, dR, Lauria, Nordström, Razborov '18]

Clique formula

Clique(G, k) for $G \sim \mathcal{G}(n, p)$ and p close to k-clique threshold

- $2^{k/(n-\Delta)^6}$-hard for resolution for very dense graph and large $k \geq n^{5/6}$
  $\Delta =$ average degree [Beame, Impagliazzo, Sabharwal '01]
- $2^{\Omega(k^{1-\epsilon})}$-hard for resolution for $k \leq n^{1/3}$ [Pang '21]
- $n^{\Omega(k)}$-hard for tree-like resolution [Lauria, Pudlák, Rödl, Thapen '13]
- $n^{\Omega(k)}$-hard for regular resolution [Atserias, Bonacina, dR, Lauria, Nordström, Razborov '18]
- $\Omega(\log n)$-degree in sum-of-squares for $p = 1/2$ [Meka, Potechin and Wigderson '15], ..., [Barak, Hopkins, Kelner, Kothari, Moitra, Potechin '16]

Clique formula

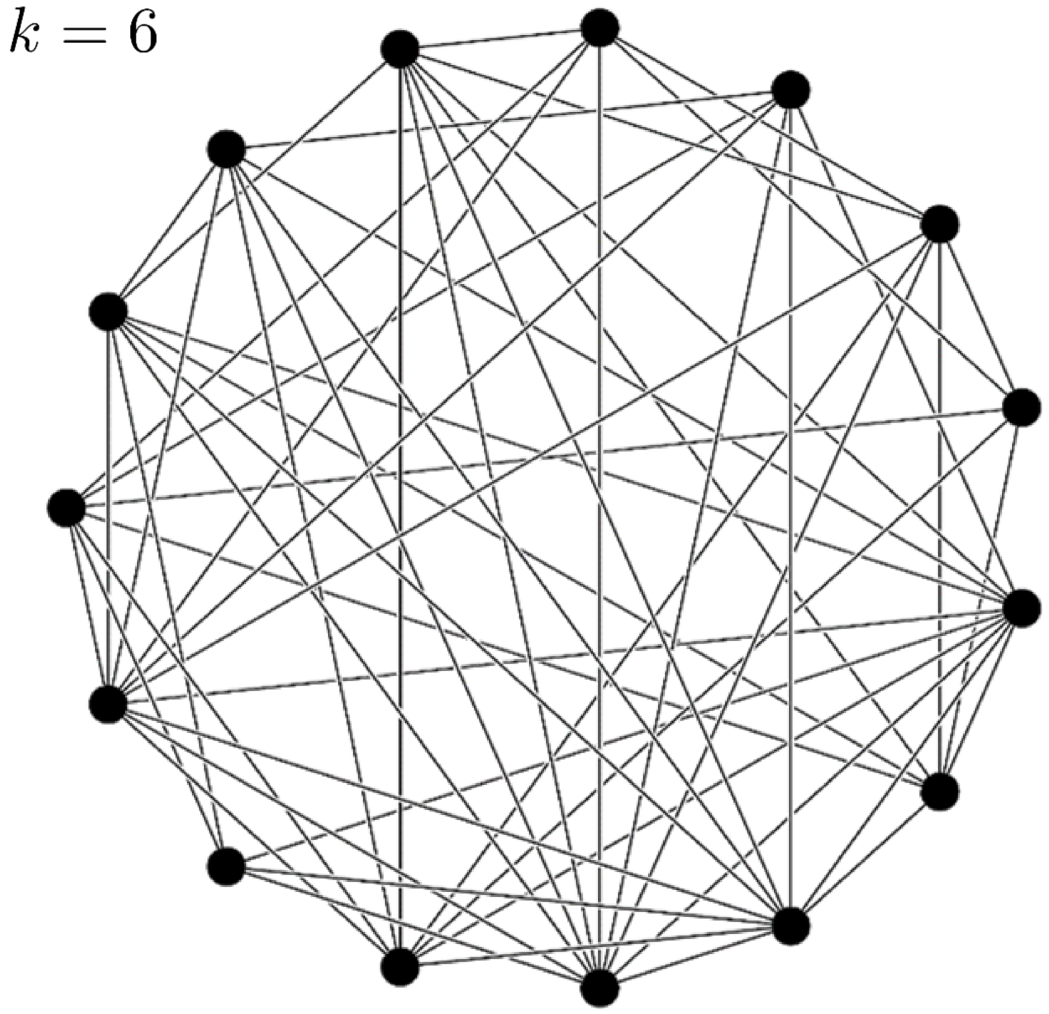
Clique(G, k) for $G \sim \mathcal{G}(n, p)$ and p close to k-clique threshold

- $2^{k/(n-\Delta)^6}$-hard for resolution for very dense graph and large $k \geq n^{5/6}$
  $\Delta =$ average degree [Beame, Impagliazzo, Sabharwal '01]
- $2^{\Omega(k^{1-\epsilon})}$-hard for resolution for $k \leq n^{1/3}$ [Pang '21]
- $n^{\Omega(k)}$-hard for tree-like resolution [Lauria, Pudlák, Rödl, Thapen '13]
- $n^{\Omega(k)}$-hard for regular resolution [Atserias, Bonacina, dR, Lauria, Nordström, Razborov '18]
- $\Omega(\log n)$-degree in sum-of-squares for $p = 1/2$ [Meka, Potechin and Wigderson '15], ..., [Barak, Hopkins, Kelner, Kothari, Moitra, Potechin '16]

Open: Show that resolution, polynomial calculus or sum of squares requires size $n^{\Omega(k)}$ to refute Clique(G, k)

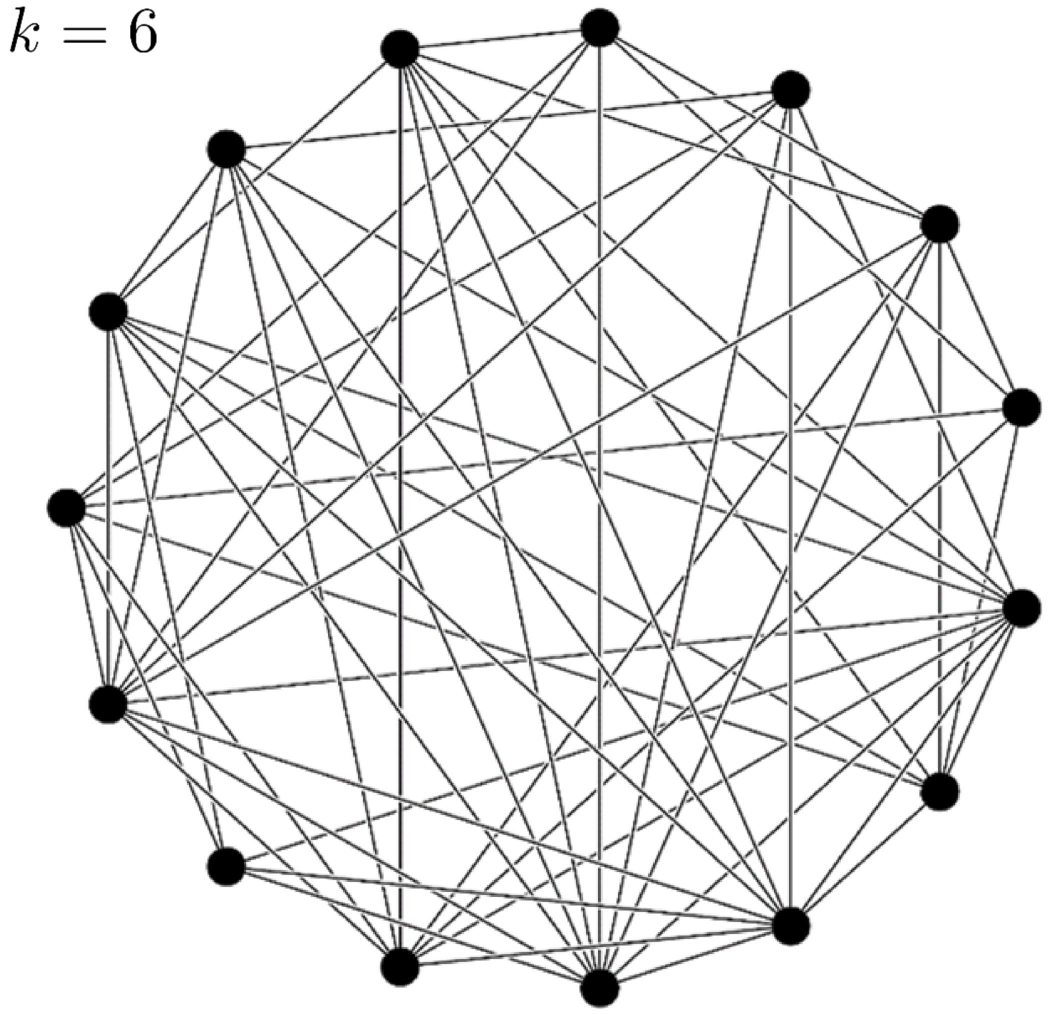
Clique formula hard for tree-like resolution

[Beyersdorff, Galesi, Lauria '11]



Clique formula hard for tree-like resolution

[Beyersdorff, Galesi, Lauria '11]

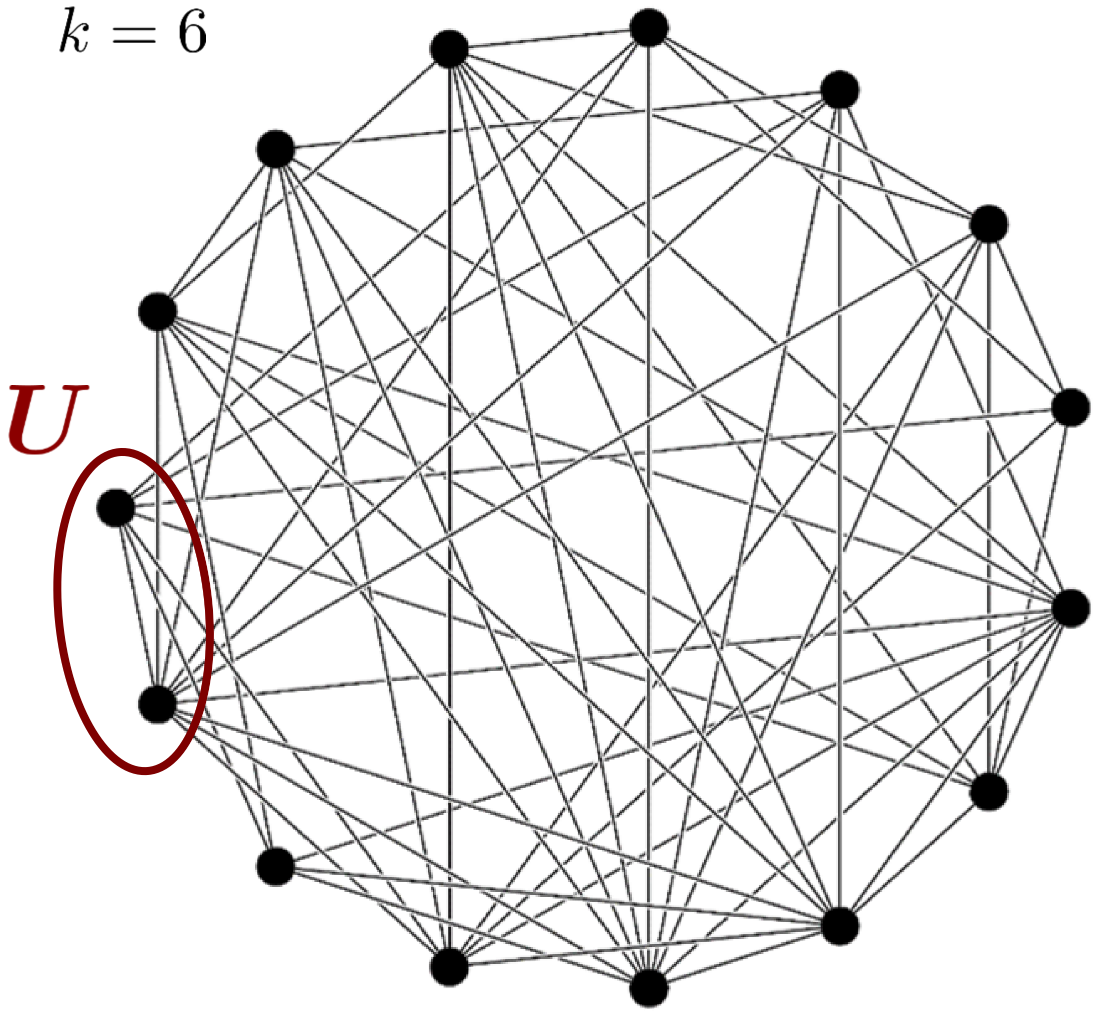


common neighbors of $\boldsymbol{U}$: $\widehat{\boldsymbol{N}}(\boldsymbol{U}) = \bigcap_{v \in U} \boldsymbol{N}(\boldsymbol{u})$

Clique formula hard for tree-like resolution

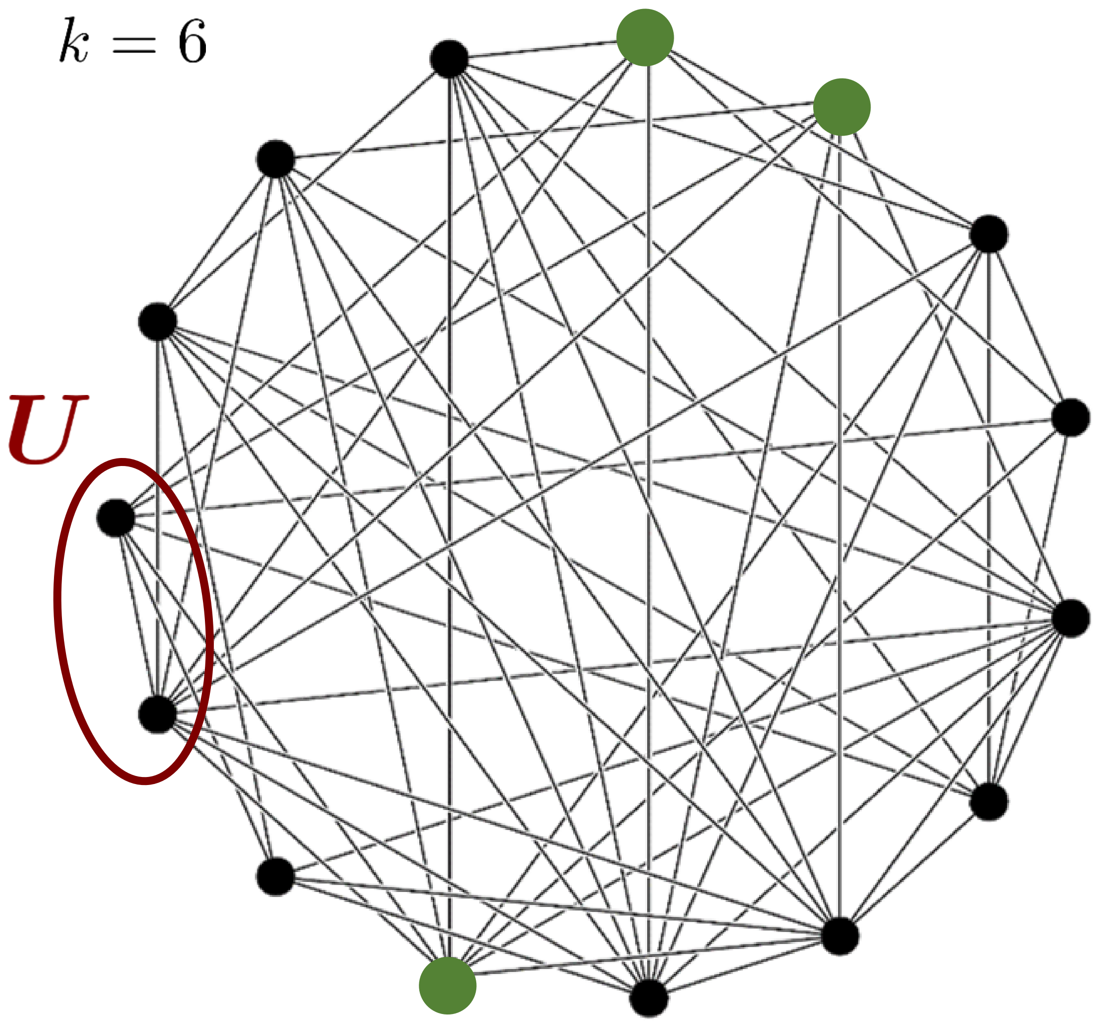
[Beyersdorff, Galesi, Lauria '11]

common neighbors of $\boldsymbol{U}$: $\widehat{\boldsymbol{N}}(\boldsymbol{U}) = \bigcap_{v \in U} \boldsymbol{N}(\boldsymbol{u})$



Clique formula hard for tree-like resolution

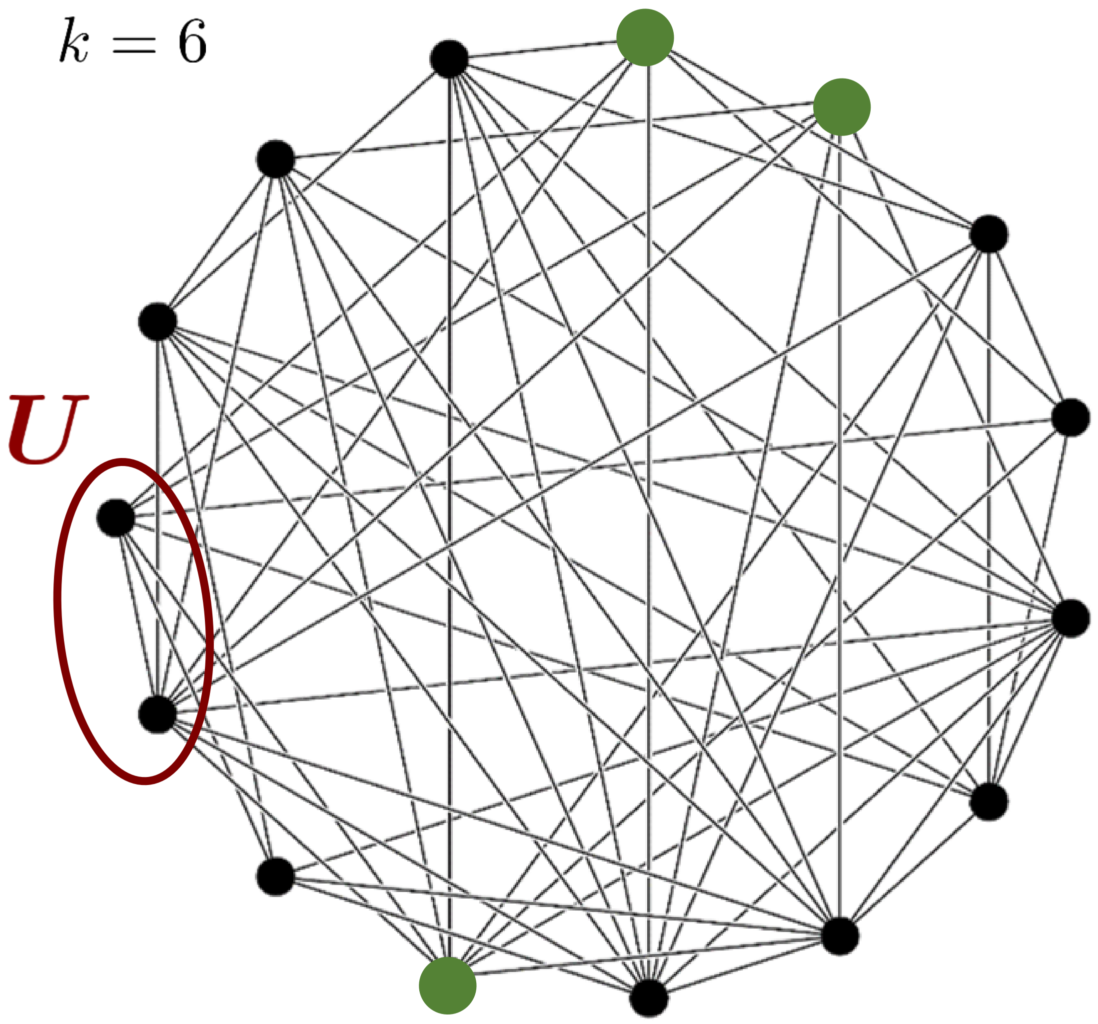
[Beyersdorff, Galesi, Lauria '11]



common neighbors of U: $\widehat{N}(U) = \bigcap_{v \in U} N(u)$

Clique formula hard for tree-like resolution

[Beyersdorff, Galesi, Lauria '11]



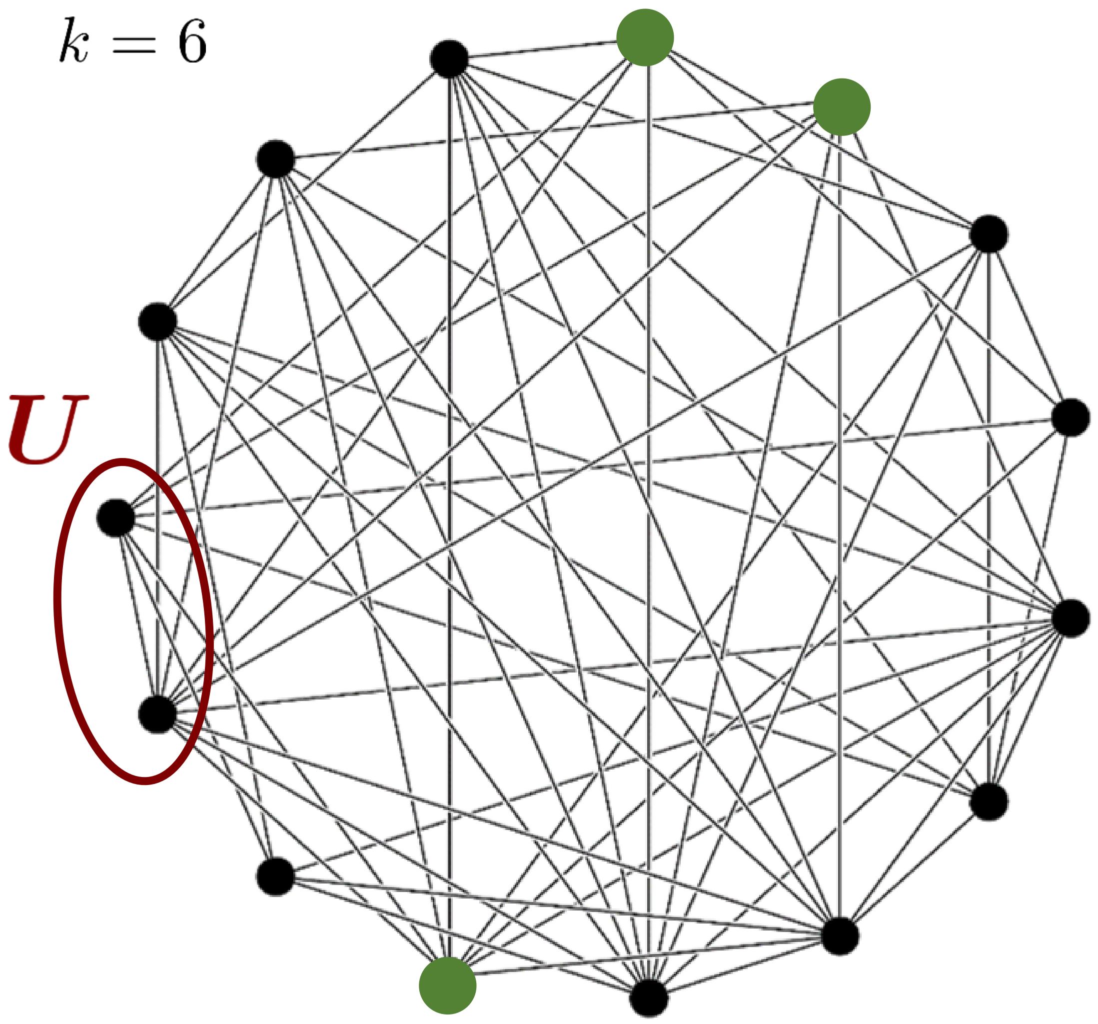
common neighbors of $\boldsymbol{U}$: $\widehat{\boldsymbol{N}}(\boldsymbol{U}) = \bigcap_{v \in U} \boldsymbol{N}(\boldsymbol{u})$

G is neighbor dense if:

can extend any $\boldsymbol{r}$-clique, $\boldsymbol{r} \leq \boldsymbol{k}/4$, in many ways, i.e.,

Clique formula hard for tree-like resolution

[Beyersdorff, Galesi, Lauria '11]



common neighbors of U: $\widehat{N}(U) = \bigcap_{v \in U} N(u)$

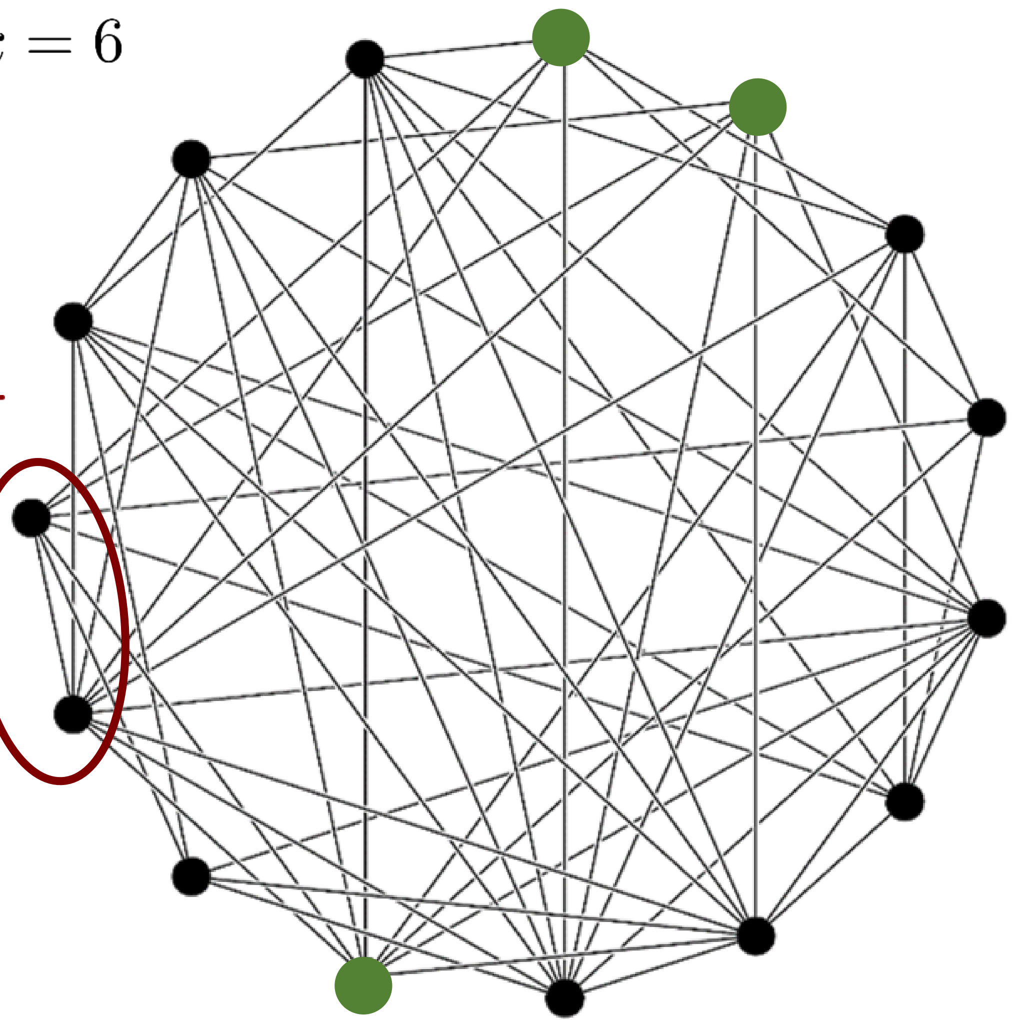
G is neighbor dense if:

can extend any r-clique, $r \leq k/4$, in many ways, i.e.,

$\forall U \subseteq V$: $|U| \leq k/4 \Rightarrow |\widehat{N}(U)| \gtrsim \sqrt{n}$

Clique formula hard for tree-like resolution

[Beyersdorff, Galesi, Lauria '11]



common neighbors of U: $\widehat{N}(U) = \bigcap_{v \in U} N(u)$

G is neighbor dense if:

can extend any r-clique, $r \le k/4$, in many ways, i.e.,

$\forall U \subseteq V$: $|U| \le k/4 \Rightarrow |\widehat{N}(U)| \gtrsim \sqrt{n}$

Lemma 1. W.h.p. $G \sim \mathcal{G}(n,p)$ is neighbor dense

(for p close to k-clique threshold)

Lemma 2. Tree-like refutation of neighbor dense G

must have size $\ge n^{\Omega(k)}$

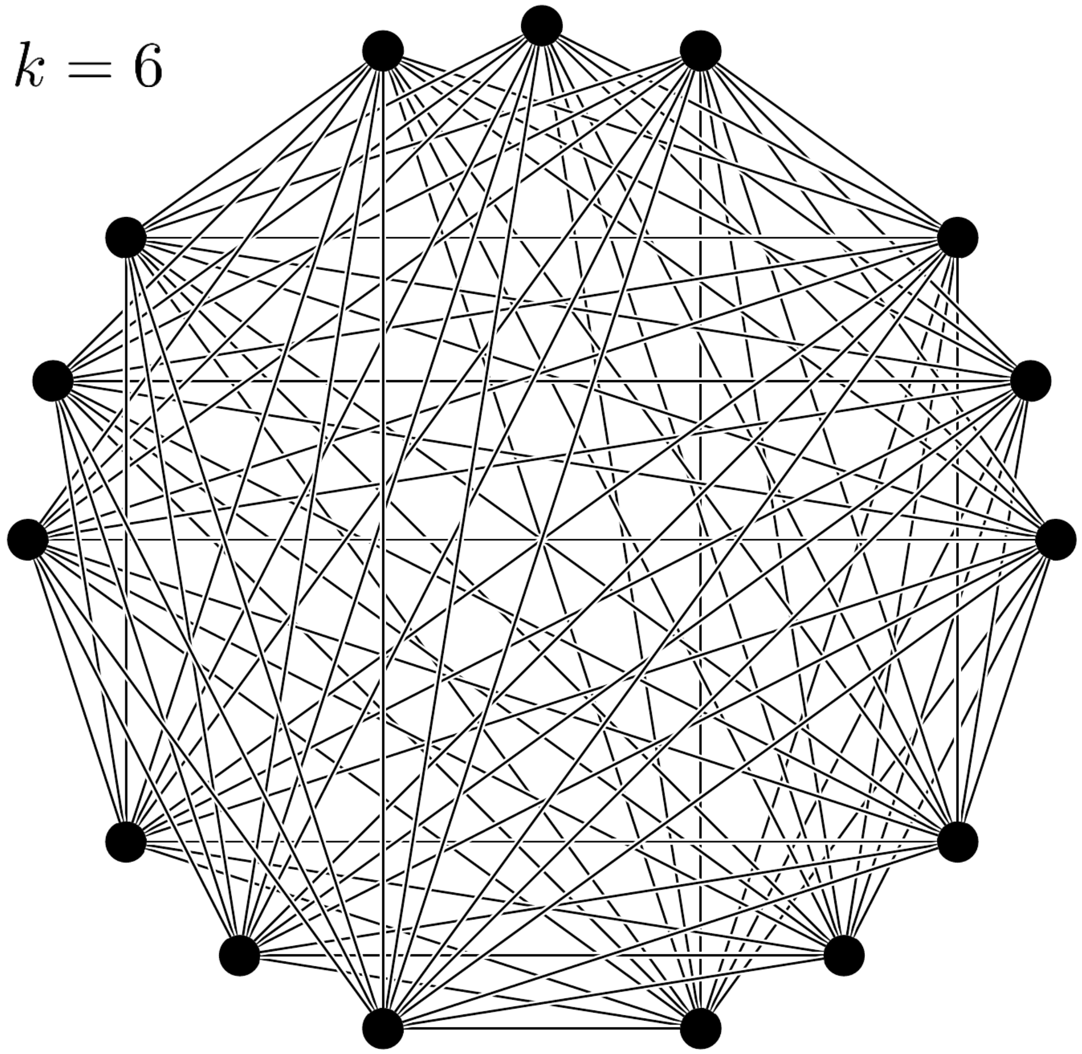
Property not enough for stronger proof systems

G is neighbor dense if:

can extend any r-clique, $r \leq k/4$, in many ways, i.e.,

$\forall U \subseteq V$: $|U| \leq k/4 \Rightarrow |\widehat{N}(U)| \gtrsim \sqrt{n}$

Property not enough for stronger proof systems



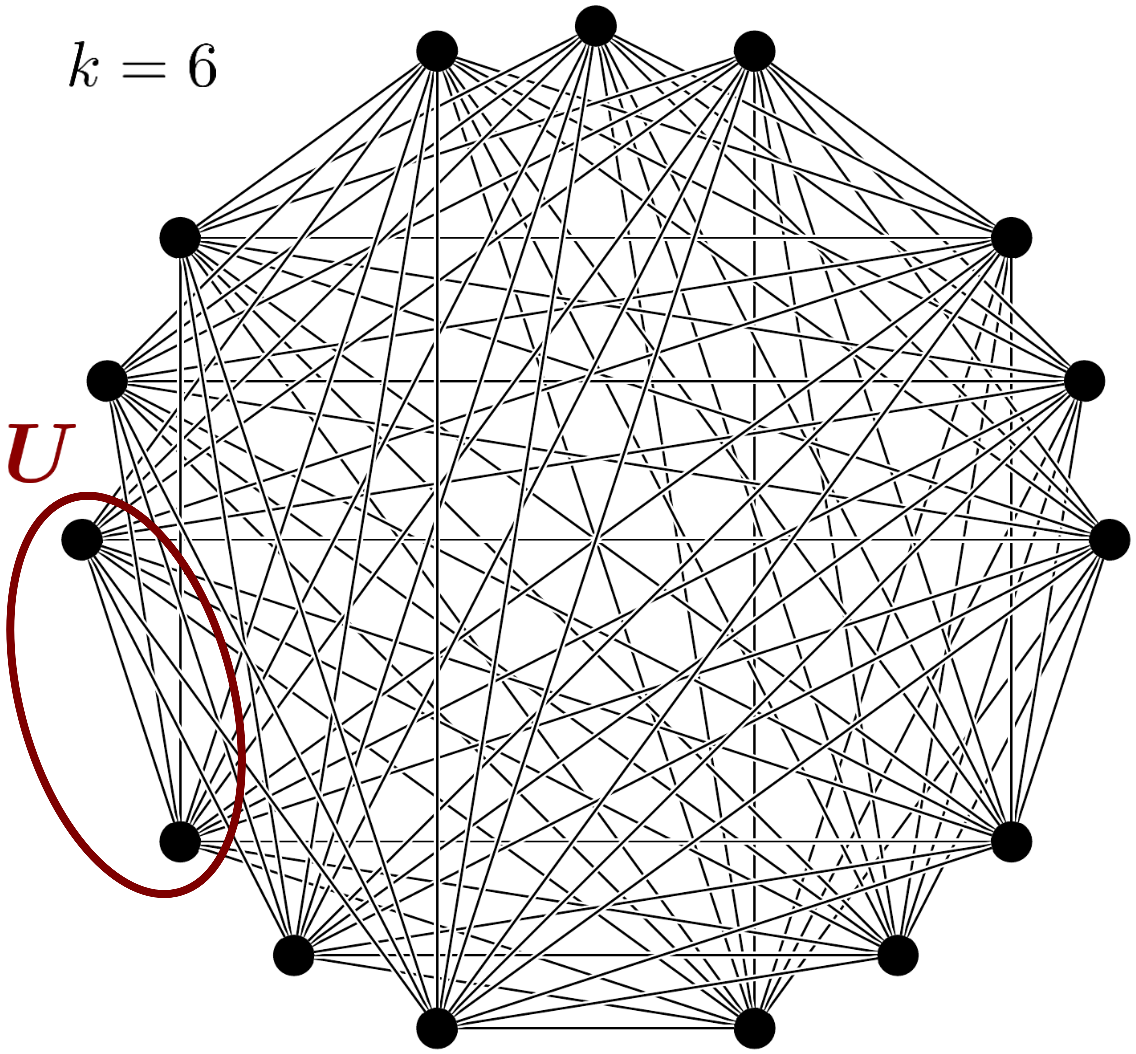
G is neighbor dense if:

can extend any r-clique, $r \leq k/4$, in many ways, i.e.,

$\forall U \subseteq V$: $|U| \leq k/4 \Rightarrow |\widehat{N}(U)| \gtrsim \sqrt{n}$

$(k-1)$-complete partite graph satisfies it!

Property not enough for stronger proof systems



G is neighbor dense if:

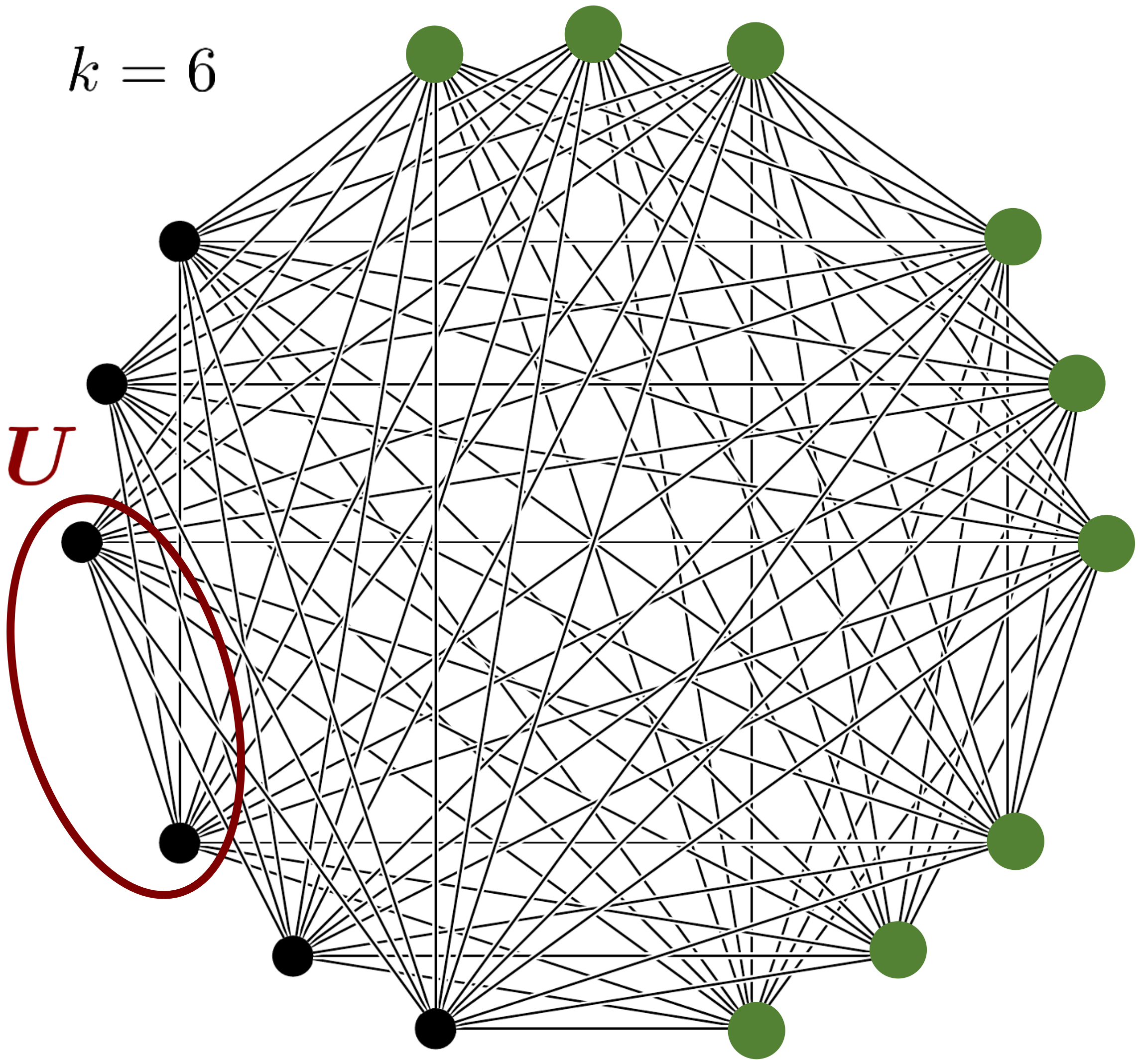
can extend any r-clique, $r \leq k/4$, in many ways, i.e.,

$\forall U \subseteq V$: $|U| \leq k/4 \Rightarrow |\widehat{N}(U)| \gtrsim \sqrt{n}$

$(k-1)$-complete partite graph satisfies it!

Property not enough for stronger proof systems

$k = 6$



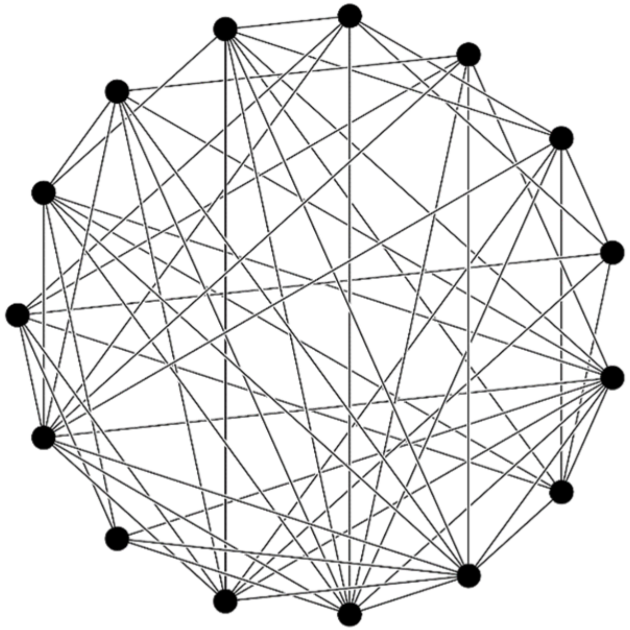
G is neighbor dense if:

can extend any r-clique, $r \leq k/4$, in many ways, i.e.,

$\forall U \subseteq V$: $|U| \leq k/4 \Rightarrow |\widehat{N}(U)| \gtrsim \sqrt{n}$

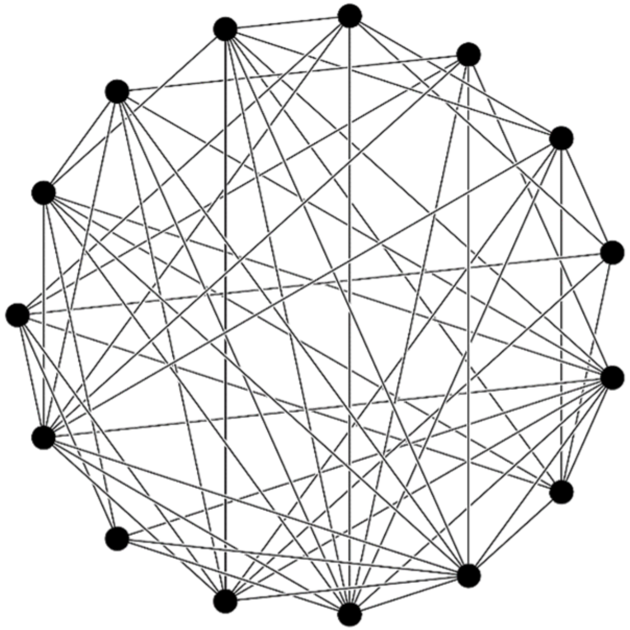
$(k-1)$-complete partite graph satisfies it!

Clique-dense property [Atserias, Bonacina, dR, Lauria, Nordström, Razborov '18]



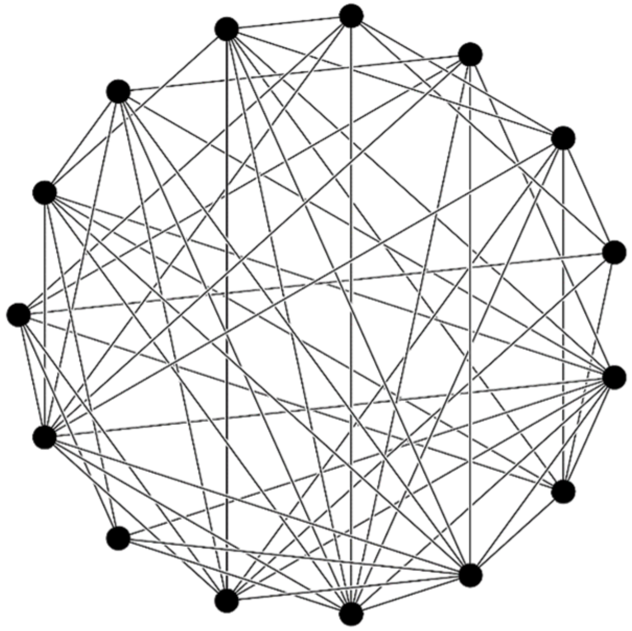
Clique-dense property [Atserias, Bonacina, dR, Lauria, Nordström, Razborov '18]

1. Can extend any $(k/20)$-clique in many ways



Clique-dense property [Atserias, Bonacina, dR, Lauria, Nordström, Razborov '18]

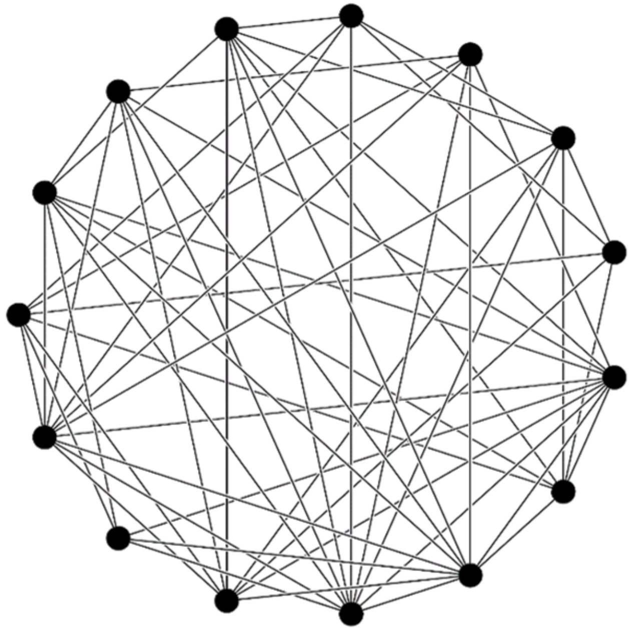
1. Can extend any $(k/20)$-clique in many ways
2. Any $W \subseteq V$ that can extend any $(k/100)$-clique in many ways can also extend *almost* any $(k/10)$-clique in many ways



Clique-dense property [Atserias, Bonacina, dR, Lauria, Nordström, Razborov '18]

1. Can extend any $(\boldsymbol{k/20})$-clique in many ways
2. Any $\boldsymbol{W} \subseteq \boldsymbol{V}$ that can extend any $(\boldsymbol{k/100})$-clique in many ways can also extend *almost* any $(\boldsymbol{k/10})$-clique in many ways

(somewhat) more formally:

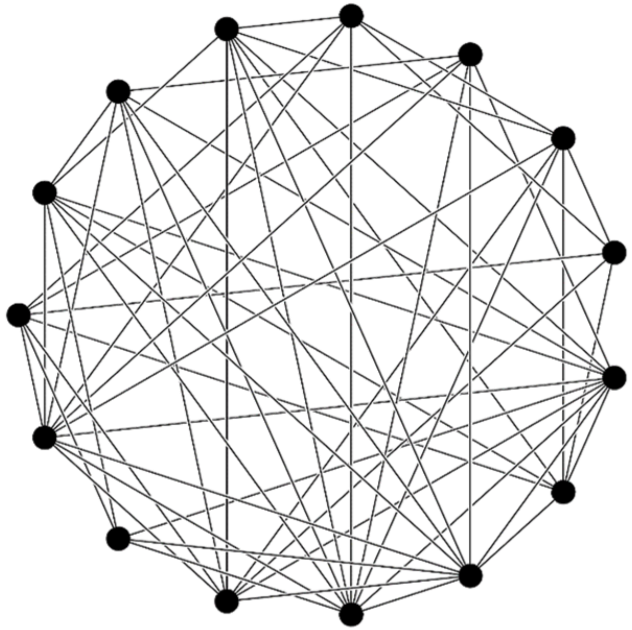


Clique-dense property [Atserias, Bonacina, dR, Lauria, Nordström, Razborov '18]

1. Can extend any $(k/20)$-clique in many ways
2. Any $W \subseteq V$ that can extend any $(k/100)$-clique in many ways can also extend *almost* any $(k/10)$-clique in many ways

(somewhat) more formally: $r = k/100$

$\forall W \subseteq V$ that can extend any r-clique in many ways:

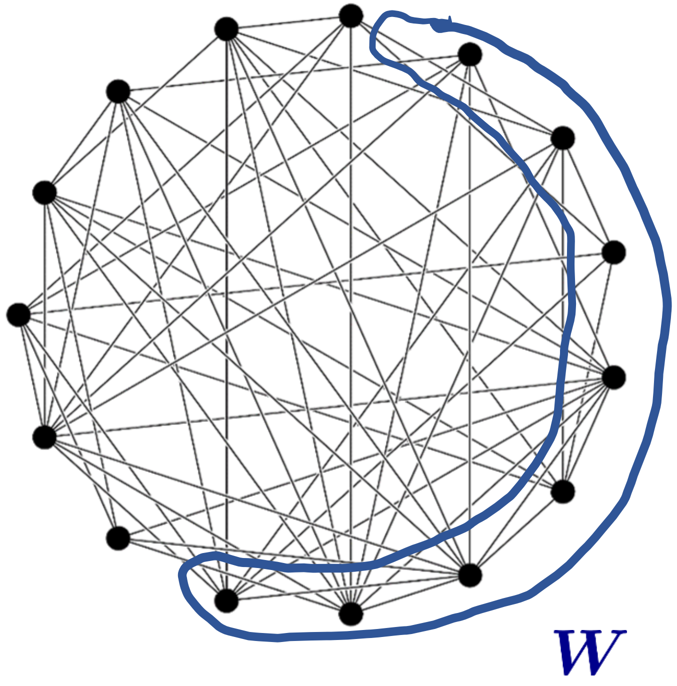


Clique-dense property [Atserias, Bonacina, dR, Lauria, Nordström, Razborov '18]

1. Can extend any $(k/20)$-clique in many ways
2. Any $W \subseteq V$ that can extend any $(k/100)$-clique in many ways can also extend *almost* any $(k/10)$-clique in many ways

(somewhat) more formally: $r = k/100$

$\forall W \subseteq V$ that can extend any r-clique in many ways:



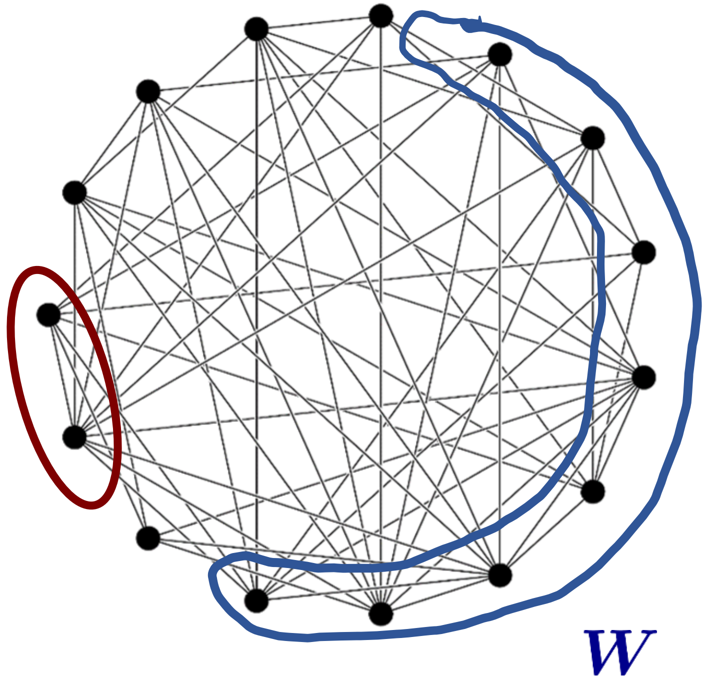
Clique-dense property [Atserias, Bonacina, dR, Lauria, Nordström, Razborov ‘18]

1. Can extend any $(k/20)$-clique in many ways
2. Any $W \subseteq V$ that can extend any $(k/100)$-clique in many ways

can also extend *almost* any $(k/10)$-clique in many ways

(somewhat) more formally: $r = k/100$

$\forall W \subseteq V$ that can extend any r-clique in many ways:

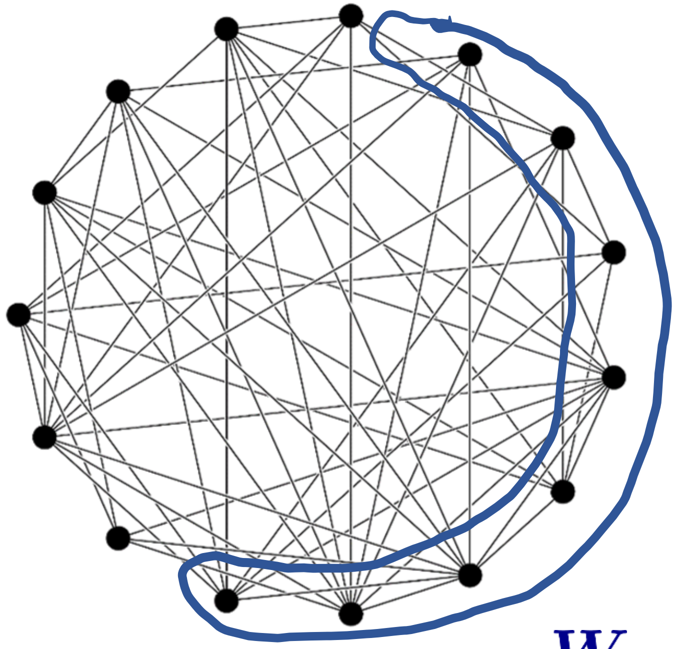


Clique-dense property [Atserias, Bonacina, dR, Lauria, Nordström, Razborov '18]

1. Can extend any $(k/20)$-clique in many ways
2. Any $W \subseteq V$ that can extend any $(k/100)$-clique in many ways can also extend *almost* any $(k/10)$-clique in many ways

(somewhat) more formally: $r = k/100$

$\forall W \subseteq V$ that can extend any r-clique in many ways:



Clique-dense property [Atserias, Bonacina, dR, Lauria, Nordström, Razborov '18]

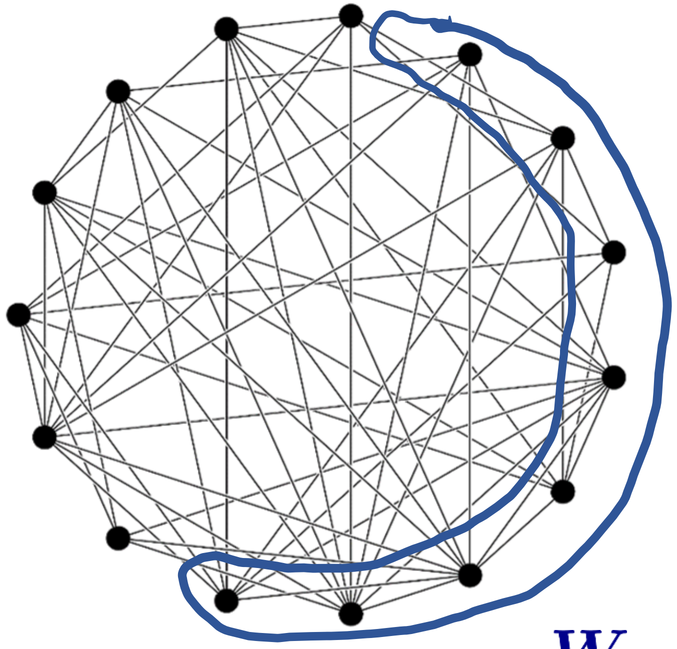
1. Can extend any $(k/20)$-clique in many ways
2. Any $W \subseteq V$ that can extend any $(k/100)$-clique in many ways

can also extend *almost* any $(k/10)$-clique in many ways

(somewhat) more formally: $r = k/100$

$\forall W \subseteq V$ that can extend any r-clique in many ways:

$\exists$ *small* set S s.t. any ℓ-clique $\ell \leq 10r$



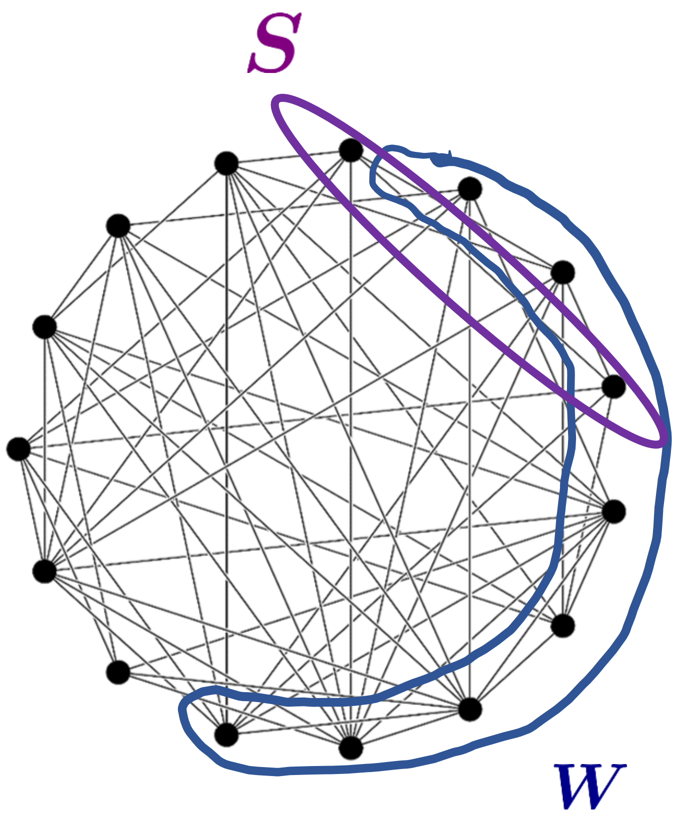
Clique-dense property [Atserias, Bonacina, dR, Lauria, Nordström, Razborov '18]

1. Can extend any $(k/20)$-clique in many ways
2. Any $W \subseteq V$ that can extend any $(k/100)$-clique in many ways can also extend *almost* any $(k/10)$-clique in many ways

(somewhat) more formally: $r = k/100$

$\forall W \subseteq V$ that can extend any r-clique in many ways:

$\exists$ *small* set S s.t. any ℓ-clique $\ell \leq 10r$



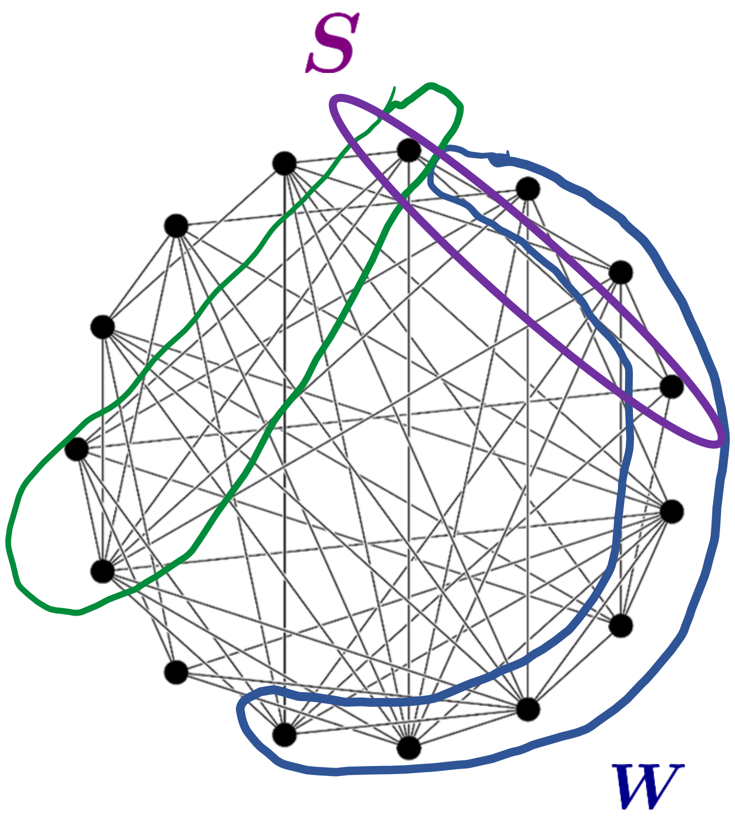
Clique-dense property [Atserias, Bonacina, dR, Lauria, Nordström, Razborov '18]

1. Can extend any $(k/20)$-clique in many ways
2. Any $W \subseteq V$ that can extend any $(k/100)$-clique in many ways can also extend *almost* any $(k/10)$-clique in many ways

(somewhat) more formally: $r = k/100$

$\forall W \subseteq V$ that can extend any r-clique in many ways:

$\exists$ *small* set S s.t. any ℓ-clique $\ell \leq 10r$



Clique-dense property [Atserias, Bonacina, dR, Lauria, Nordström, Razborov '18]

1. Can extend any $(k/20)$-clique in many ways
2. Any $W \subseteq V$ that can extend any $(k/100)$-clique in many ways can also extend *almost* any $(k/10)$-clique in many ways

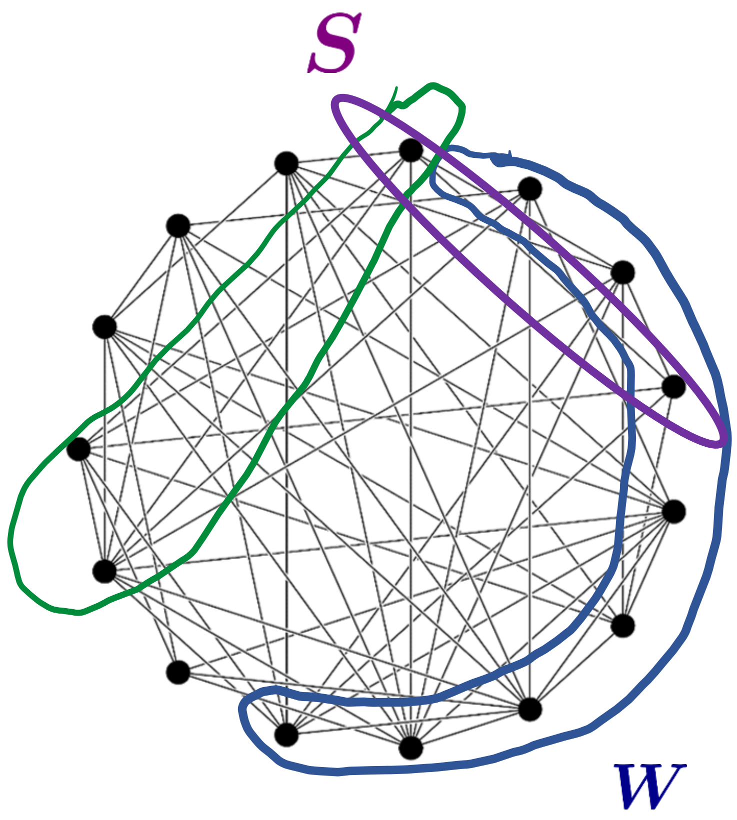
(somewhat) more formally: $r = k/100$

$\forall W \subseteq V$ that can extend any r-clique in many ways:

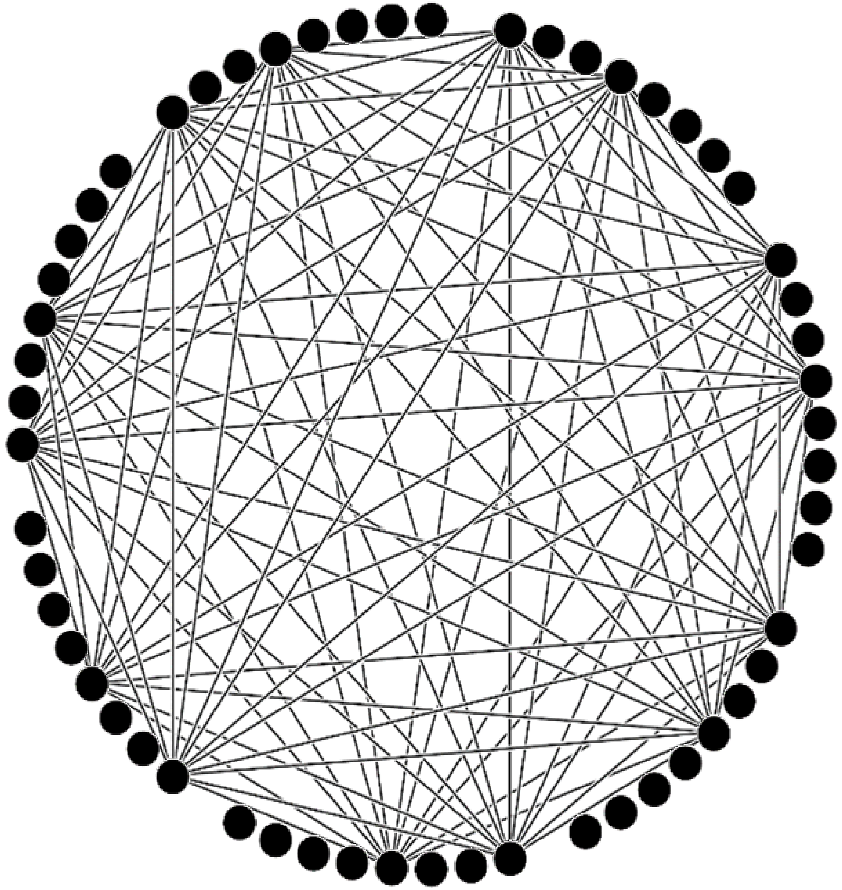
$\exists$ *small* set S s.t. any ℓ-clique $\ell \leq 10r$

that cannot be extended in W in many ways

must intersect S in many vertices



(k-1)-partite graph does not satisfy clique-denseness



$r = k/100$

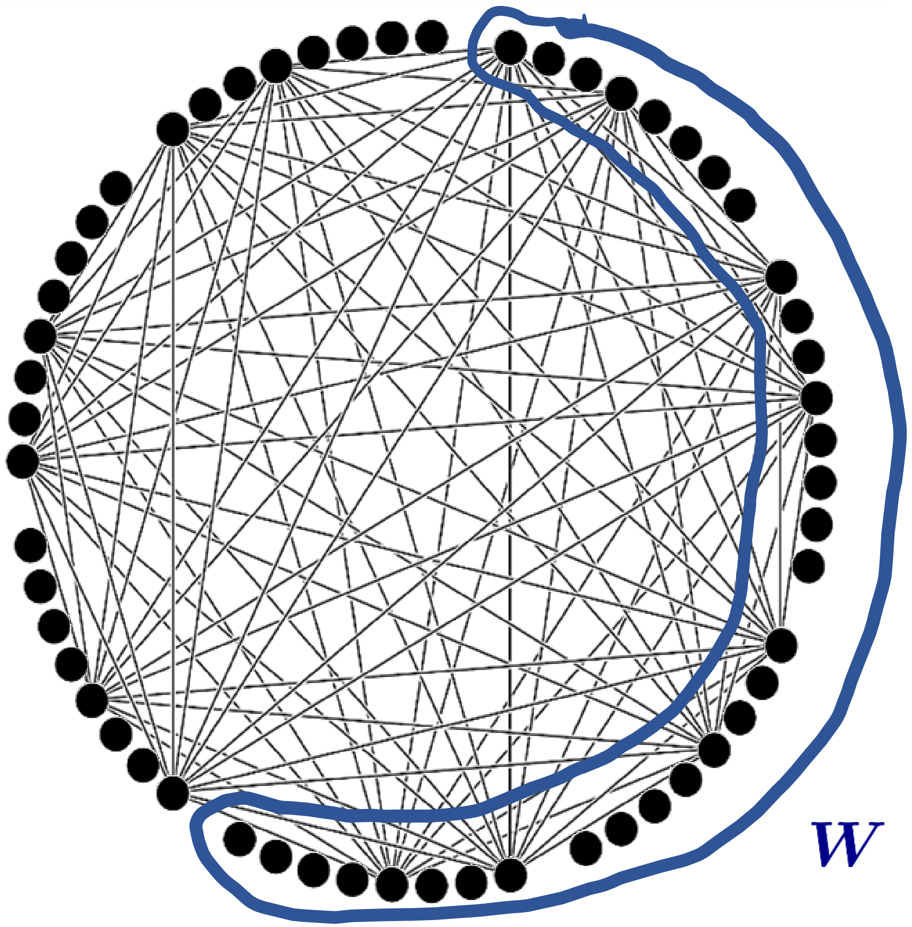
$\forall W \subseteq V$ that can extend any r-clique in many ways:

$\exists$ *small* set S s.t. any ℓ-clique $\ell \leq 10r$

that cannot be extended in W in many ways

must intersect S in many vertices

(k-1)-partite graph does not satisfy clique-denseness



$r = k/100$

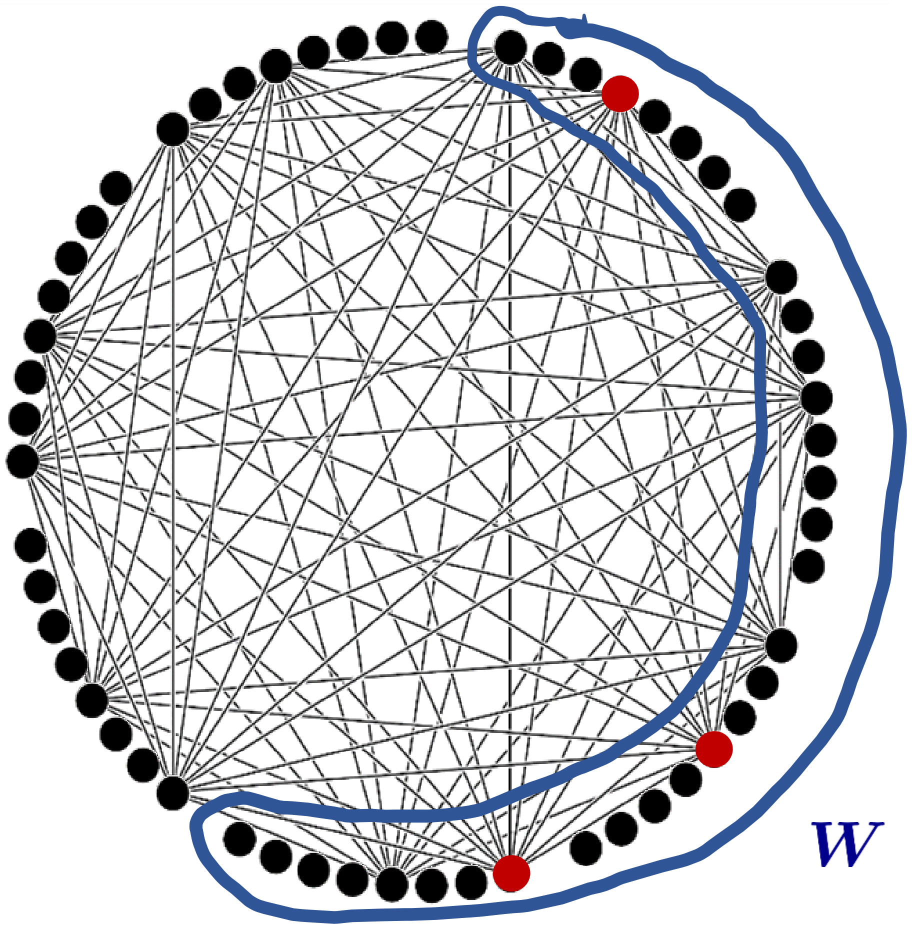
$\forall W \subseteq V$ that can extend any r-clique in many ways:

$\exists$ *small* set S s.t. any ℓ-clique $\ell \leq 10r$

that cannot be extended in W in many ways

must intersect S in many vertices

(k-1)-partite graph does not satisfy clique-denseness



$r = k/100$

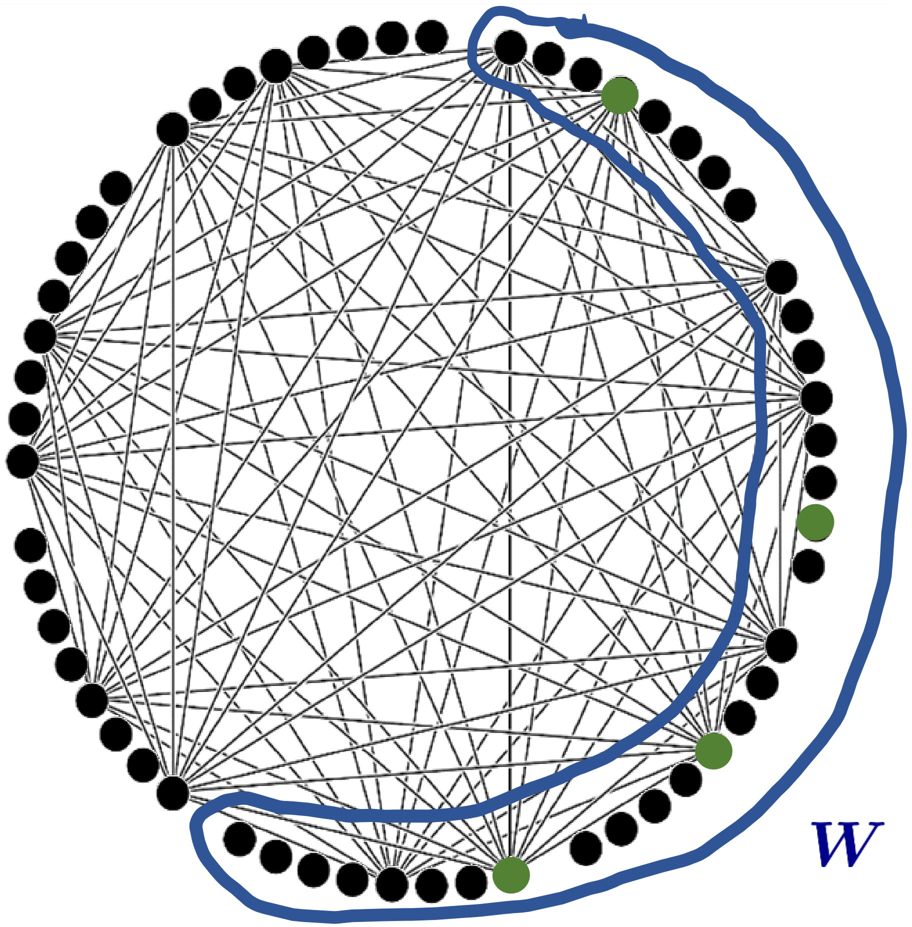
$\forall W \subseteq V$ that can extend any r-clique in many ways:

$\exists$ *small* set S s.t. any ℓ-clique $\ell \leq 10r$

that cannot be extended in W in many ways

must intersect S in many vertices

(k-1)-partite graph does not satisfy clique-denseness



$r = k/100$

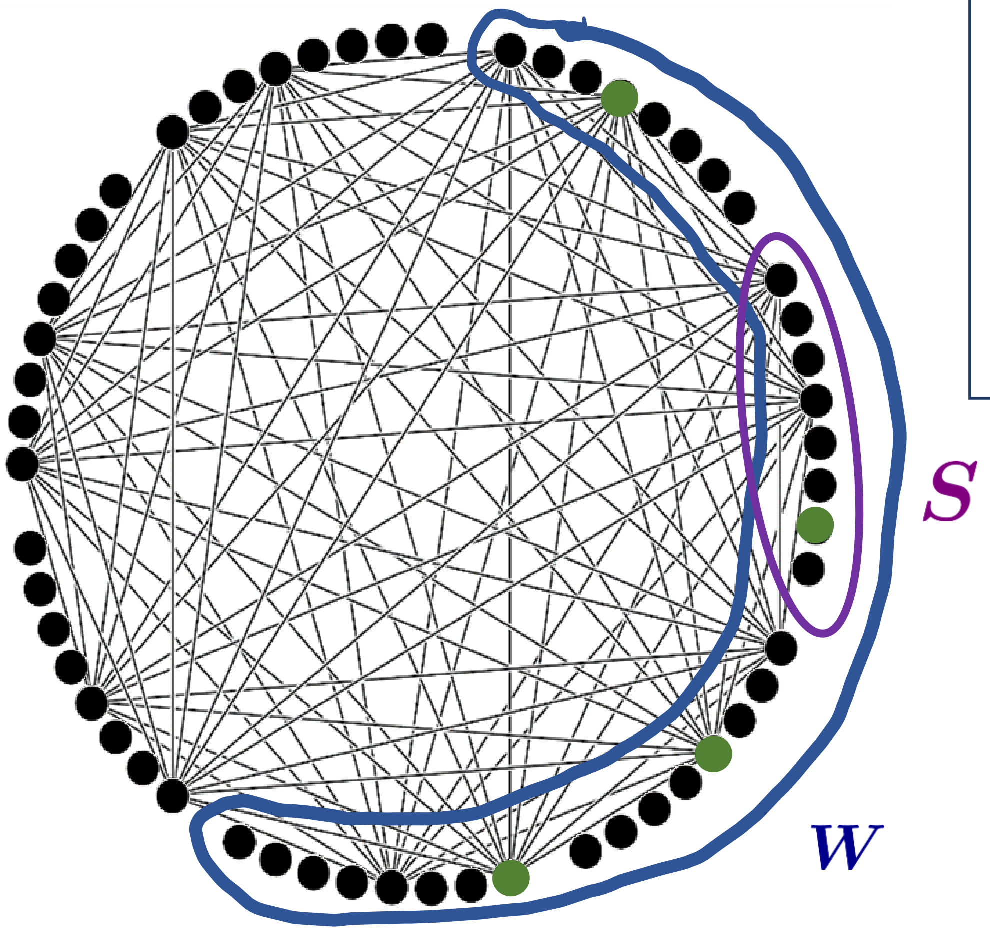
$\forall W \subseteq V$ that can extend any r-clique in many ways:

$\exists$ *small* set S s.t. any ℓ-clique $\ell \leq 10r$

that cannot be extended in W in many ways

must intersect S in many vertices

(k-1)-partite graph does not satisfy clique-denseness



$r = k/100$

$\forall W \subseteq V$ that can extend any r-clique in many ways:

$\exists$ *small* set S s.t. any ℓ-clique $\ell \leq 10r$

that cannot be extended in W in many ways

must intersect S in many vertices

Summary

- Average-case proof complexity of three NP-hard problems
  - Primarily interested in size of proofs

Summary

- Average-case proof complexity of three NP-hard problems
  - Primarily interested in size of proofs
- Imply lower bounds for algorithms

Summary

- Average-case proof complexity of three NP-hard problems
  - Primarily interested in size of proofs
- Imply lower bounds for algorithms
- Candidate hard instances for strong proof systems

Summary

- Average-case proof complexity of three NP-hard problems
  - Primarily interested in size of proofs
- Imply lower bounds for algorithms
- Candidate hard instances for strong proof systems
- Lower bounds: identify structure in randomness

Summary

- Average-case proof complexity of three NP-hard problems
  - Primarily interested in size of proofs
- Imply lower bounds for algorithms
- Candidate hard instances for strong proof systems
- Lower bounds: identify structure in randomness
- Many open problems

Average-case hardness results

	k-clique	k-coloring	3-SAT	3-XOR
Tree-like Resolution	HARD (size $n^{\Omega(k)}$) [Beyersdorff, Galesi, Lauria '11]	HARD [Beame, Culberson, Mitchell, Moore '05]	HARD [Chvátal, Szemerédi '88] Improved [Ben-Sasson, Galesi '01] (size $\exp(n/\Delta^{1+\epsilon})$) $\Delta = m/n$	
Resolution	OPEN Some partial results*		HARD [Chvátal, Szemerédi '88] $\exp(n/\Delta^{2+\epsilon})$ Improved [Beame, Karp, Pitassi, Saks '98], [Ben-Sasson '01]	
Polynomial Calculus	OPEN	OPEN	$\mathbb{F} \neq 2$: HARD [Ben-Sasson, Impagliazzo '99]	
			$\mathbb{F} = 2$: HARD [Alekhnovich, Razborov '01]	EASY
Sum of Squares	OPEN Some partial results** $\mathcal{G}(n, 1/2)$: degree $= \Theta(\log n)$	OPEN [Kothari, Manohar '21] $\mathcal{G}(n, 1/2)$: $d \geq \Omega(\log n)$	HARD [Grigoriev '01, Schoenebeck '08]	
Cutting Planes	OPEN	OPEN	OPEN $\Theta(\log n)$-SAT [Fleming, Pankratov, Pitassi, Robere '17] [Hrubeš, Pudlák '17]	Quasi-poly EASY [Fleming, Göös, Impagliazzo, Pitassi, Robere, Tan, Wigderson '21] [Dadush, Tiwari '20]

* [Beame, Impagliazzo, Sabharwal '01], [Pang '21], [Atserias, Bonacina, **dR**, Lauria, Nordström, Razborov '18], [Lauria, Pudlák, Rödl, Thapen '13]
** [Meka, Potechin and Wigderson '15], ..., [Barak, Hopkins, Kelner, Kothari, Moitra, Potechin '16]

Average-case hardness results

Thank you!

	k-clique	k-coloring	3-SAT	3-XOR
Tree-like Resolution	HARD (size $n^{\Omega(k)}$) [Beyersdorff, Galesi, Lauria '11]	HARD [Beame, Culberson, Mitchell, Moore '05]	HARD [Chvátal, Szemerédi '88] Improved [Ben-Sasson, Galesi '01] (size $\exp(n/\Delta^{1+\epsilon})$) $\Delta = m/n$	
Resolution	OPEN Some partial results*		HARD [Chvátal, Szemerédi '88] $\exp(n/\Delta^{2+\epsilon})$ Improved [Beame, Karp, Pitassi, Saks '98], [Ben-Sasson '01]	
Polynomial Calculus	OPEN	OPEN	$\mathbb{F} \neq 2$: HARD [Ben-Sasson, Impagliazzo '99] $\mathbb{F} = 2$: HARD [Alekhnovich, Razborov '01]	$\mathbb{F} = 2$: EASY
Sum of Squares	OPEN Some partial results** $\mathcal{G}(n, 1/2)$: degree $= \Theta(\log n)$	OPEN [Kothari, Manohar '21] $\mathcal{G}(n, 1/2)$: $d \geq \Omega(\log n)$	HARD [Grigoriev '01, Schoenebeck '08]	
Cutting Planes	OPEN	OPEN	OPEN $\Theta(\log n)$-SAT [Fleming, Pankratov, Pitassi, Robere '17] [Hrubeš, Pudlák '17]	Quasi-poly EASY [Fleming, Göös, Impagliazzo, Pitassi, Robere, Tan, Wigderson '21] [Dadush, Tiwari '20]

* [Beame, Impagliazzo, Sabharwal '01], [Pang '21], [Atserias, Bonacina, **dR**, Lauria, Nordström, Razborov '18], [Lauria, Pudlák, Rödl, Thapen '13]
** [Meka, Potechin and Wigderson '15], ..., [Barak, Hopkins, Kelner, Kothari, Moitra, Potechin '16]