

PROOF COMPLEXITY AS A COMPUTATIONAL LENS

LECTURE 5

Recap

Clique-colouring formula hard for resolution

"There exists an n -vertex graph G such that

(a) G has m -clique

(b) G is $(m-1)$ -colourable"

Set $m = n^{\delta}$

Proof technique: Given resolution refutation,

use FEASIBLE(MONOTONE)INTERPOLATION to

extract monotone circuit C that given graph G determines whether (a) or (b) applies

Circuit size = $\Theta(\text{resolution proof length})$

But can prove there are no small

monotone circuits for this problem

Hence no short resolution refutations either

Question

What if we fix graph G ?

i.e., consider fixed assignment

$g : \{0,1\}^{\binom{n}{2}} \rightarrow \{0,1\}$ to edge indicator variables?

GRAPH COLOURING

Strong lower bounds for resolution and polynomial calculus - will see this later in the course

CLIQUE

Much less well understood

CLIQUE PROBLEM

Given undirected n -vertex graph G , does G contain a k -clique?

(i.e., k vertices $v_1, v_2, \dots, v_k \in V(G)$

such that $(v_i, v_j) \in E(G)$ for all $1 \leq i < j \leq k$)

Trivial algorithm in $\underline{\text{time}}\left(\frac{n}{k}\right)^{O(1)} n^{O(k)}$

Check all subsets of vertices of size k

Can do slightly better with algebraic techniques: $O(n^{\omega k/3})$ where

$\omega \approx 2.373$ is matrix multiplication exponent

But $\underline{n^{O(k)}}$ upper bound best known

Believed to be hard

- even to meaningfully approximate size of largest clique

- even on average for random graphs

$\underline{n^{\Omega(k)}}$ believed to be correct lower bound

True if SAT problem requires exponential time

(Exponential Time Hypothesis or ETH)

$n^{\Omega(k)} \Leftrightarrow \exists \delta > 0$ such that running time is $\underline{\Omega}(n^{\delta k})$ independent of k

But in practice there are very good algorithms for computing max cliques

What about unconditional results? | C1 II

In circuit complexity $n^{52(k)}$ lower bounds for

- bounded-depth circuits
- monotone circuits

Can we prove lower bounds for conflict-driven clause learning (CDCL) SAT solvers?

For state-of-the-art max clique algorithms?

What we will talk about today and next time

THEOREM (very informal)

For $k \ll \sqrt{n}$ it holds for suitably sampled random n -vertex graphs G that

- (a) G does not have k -clique a.a.s.
- (b) REGULAR resolution requires length n^{2k} to prove this a.a.s.

Recall that a sequence of events $\{E_n\}_{n=1}^{\infty}$ happens ASYMPTOTICALLY ALMOST SURELY if

$$\lim_{n \rightarrow \infty} \Pr[E_n] = 1.$$

COROLLARY

Essentially all Many state-of-the-art max clique algorithms require time n^{2k}

Corollary follows since reasoning on many max clique algorithms is captured by regular resolution

(This requires an argument that we will ignore. Also, there is a 2^k -loss in reasoning power, but this is dominated by the $n^{-\epsilon(k)}$ lower bound)



But we still have no good bounds for general resolution and CDCL



Need to make precise what we mean by

- (i) "random" graph
- (ii) "regular" resolution
- (iii) "prove" — which formula is refuted?



(i) Erdős-Rémyi random graphs $G(n,p)$

n -vertex undirected graph; think of $V = [n]$
For every potential edge (i,j) , $i < j$, flip independent coin and include edge with probability p

For $X \subseteq V$, $|X| = k$,

$$\Pr[X \text{ forms clique}] = p^{\binom{k}{2}}$$

Let $N_k = \# k\text{-cliques in random graph}$ C1 IV

$$E[N_k] = \binom{n}{k} p^{\binom{k}{2}} \quad (*)$$

by linearity of expectation

For $p = n^{-\frac{2}{k-1}}$ get $E[N_k] \approx 1$

Markov's inequality says that

$$\Pr[N_k \geq 1] \leq E[N_k]$$

Fix constant $\eta > 1$ and set

$$p = n^{-\frac{2\eta}{k-1}} \quad (**)$$

Then

$$\Pr[N_k \geq 1] \leq E[N_k] \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

$G \sim G(n, p)$ very unlikely to contain k -clique
(but will contain many cliques of size $\Omega(k)$)

This is the random graph distribution that we will study.

(ii) Regular resolution

General resolution: Proof DAG without restrictions

Tree-like resolution: Proof DAG is binary tree

Equivalent to decision tree for falsified clause search problem:

- Internal nodes query variables
- Edges labelled by values to queried variables
- Each leaf labelled by clause C falsified by truth value assignment defined by path to leaf

Regular resolution: DAG-like proof

But on each path every variable resolved over at most once

(once variable eliminated, not introduced again)

Regular resolution is exponentially stronger than tree-like resolution

(E.g., for XORified pebbling formulas, which we will see later in the course)

Regular seems "morally close" to general resolution
(Why re-introduce variables that have been eliminated?!)

Has been conjectured to be as strong as general resolution for some classes of formulas

But regular resolution is exponentially weaker than general resolution (but separating formulas are hard to find and are quite contrived)

Regular resolution = Read-once branching program (ROBP)

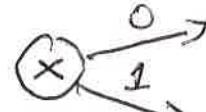
Read-once branching program (ROBP) for CNF formula F

Directed acyclic graph (DAG)

One source node (no incoming edges)
= start node

Every sink node (no outgoing edges) labelled by clause $C \in F$

Size = # nodes

Every non-sink node  C1 VI

- labelled by variable x
- 2 outgoing edges labelled 0 and 1

"Node queries x and takes edge with label agreeing with value of x "

Any assignment α to $\text{Vars}(F)$ defines source-to-sink path

Branching program solves (falsified clause search problem for) F if $\forall \alpha$ reach sink labelled by C such that $\alpha(C) = 0$.

Read-once Along any path, query any variable at most once

General resolution proof \Rightarrow general branching program

- make 1 start node
- flip direction of all edges (\times)
- let resolvent node query variable resolved over
- label edge to premise by value $\{0, 1\}$
falsifying literal over x in premise clause

So lower bounds for branching program size
 \Rightarrow lower bounds for resolution length

Unfortunately, general branching programs
much more powerful than resolution ...

Regular resolution proof \Rightarrow Read-once branching program

C1 VII

By definition of regularity, every variable queried at most once

In fact, regular resolution \Leftrightarrow ROBP,
but we only need the easy direction
(The harder direction tells us that this is a promising approach, though)

(iii) Clique formula

Basically clique colouring formula but with restriction to edge indicator variables
But with some constraints added to avoid formula being hard for "stupid reasons"

Variables $x_{v,i} = \text{"v is } i\text{th member of clique"}$
 $v \in V = \{1, 2, \dots, n\} = [n]$
 $i \in \{1, 2, \dots, k\} = [k]$

CLIQUE AXIOMS

$$\bigvee_{v \in V} x_{v,i}$$

$$i \in [k]$$

EDGE AXIOMS

$$\overline{x}_{u,i} \vee \overline{x}_{v,j}$$

$$i \neq j \in [k]$$

$$(u, v) \notin E$$

$$i \in [k]$$

$$u \neq v \in V$$

FUNCTIONALITY AXIOMS

$$\overline{x}_{u,i} \vee \overline{x}_{v,i}$$

$$i \in [k]$$

ORDERING AXIOMS

$$\overline{x}_{u,i} \vee \overline{x}_{v,j}$$

$$u < v \in V$$

$$(u, v) \in E$$

$$i > j \in [k]$$

"if u & v are both in clique and $u < v$,
then clique index of u < clique index of v "

THEOREM [Aceras, Bonacch, de Rezende, Lauria, Nordström, Razborov]

For any constant $\epsilon > 0$, $\eta > 1$, if
 $G \sim G(n, n^{-2\eta/(k-1)})$ for $k \leq n^{1+\epsilon}$
then asymptotically almost surely
 G does not contain a k -clique but
any regular resolution refutation of
the k -clique formula for G has length $n^{\Omega(k)}$.

Ordering axioms are important — but
we will cheat a bit and ignore them
for simplicity of exposition

All key ideas still there, and even
without ordering axioms the lower bound
is highly nontrivial and interesting

Will do slightly simpler proof from conference
paper [ABdRLNR '18]. Full details in
journal version [ABdRLNR '21].

PREVIOUSLY KNOWN

C18

TREE-LIKE RESOLUTION

Upper bound $n^{O(k)}$

Proof sketch: Build decision tree that tests all $\binom{n}{k}$ clique candidates (and for each candidate finds missing edge)

Lower bound $n^{\Omega(k)}$ for Erdős-Rényi random graphs [Beyersdorff, Galesi, Lauria '13]

Prosecutor-Defendant game w/o forgetting (= decision tree)

Defendant commits to k' -clique for $k' = \Omega(k)$

When asked about $x_{v,i}$, $v \in V$, say no/0 if $v \notin K$. Say yes/1 if $v \in K$ and not already answered 1 for $x_{v,j}$ & $j \neq i$

When Prosecutor wins, has identified clique $n^{\Omega(k)}$ different cliques \Rightarrow lower bound

Same lower-bound argument works for balanced, complete $(k-1)$ -partite graphs

GENERAL RESOLUTION

[Beame, Impagliazzo, Sabharwal '07]

Lower bound $\exp(\Omega(n^\varepsilon))$, $\varepsilon > 0$,
for $n^{5/6} \leq k \leq n/3$ for

very dense Erdős-Rényi graphs

ε depends on density. For $k = n/3$, get $\exp(-\Omega(n))$
But not right dependence on k

REGULAR RESOLUTION

C X

Complete $(k-1)$ -partite graphs

(and only $(k-1)$ -colourable graphs)

are ETSY - regular resolution repetition

$$\text{in length} \leq [2^k k^2 n^2] [BGL13]$$

$$V = V_1 \cup V_2 \cup \dots \cup V_{k-1}$$

$$\mathcal{E} = \bigcup_{1 \leq i < j \leq k-1} V_i \times V_j$$

	V_1	V_2	\dots	V_{k-1}
1				
2				
\vdots				
k				

Prosecutor strategy:

Prosecutor's strategy:

Ask $x_{v,k}$ for all $v \in V$

All answers = 0 \Rightarrow win (clique axiom)

Suppose $x_{v,k} = 1$ for $v \in V_i$

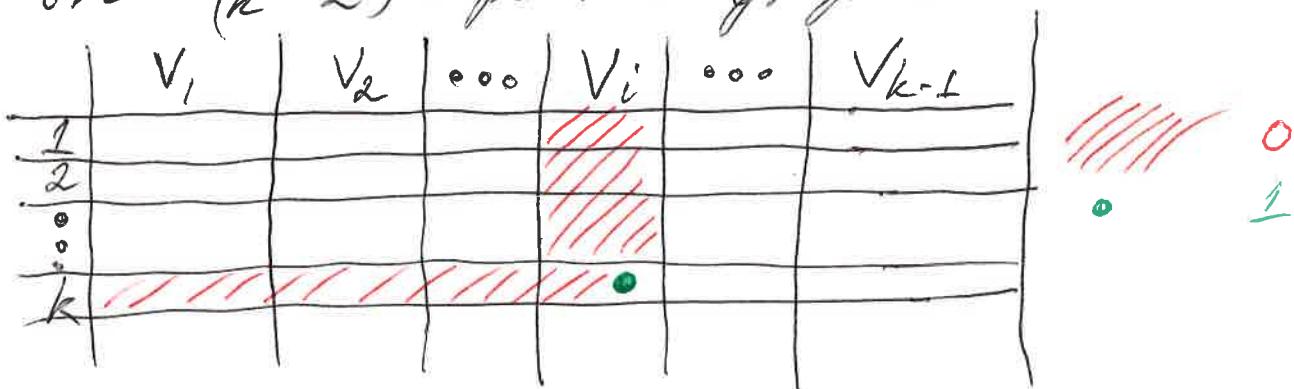
Ask sc_v for all $v \in V_i$, $\ell \leq k-1$

All answers 0, since otherwise violate edge axiom

Forget all variables x, y, z

Now we have $(k-1)$ -ligne problem

on $(k-2)$ -partite graph



or the round:

C1 XI

looking for $(k+1-r)$ -clique in
 $(k-r)$ partite graph

For purposes of illustration, fix $k=8, r=4$

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	
1								0
2								1
3								
4								
5								
6								
7								
8								forgotten

$(r-1)$ blocks v_i already ruled out

$\binom{k}{r-1}$ possible starting records

- ① Ask $x_{v_i, k+1-r}$ for all (remaining) $v_i \in V$
 $\leq n$ queries
until find a 1 in some v_i .

- ② Ask for all $u \in V$ and all $l \leq k+r$
about $x_{u,l}$
 $\leq k+n$ queries

- ③ Forget everything about variables $x_{v_i, k+1-r}$

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	
1								
2								
3								
4								
5								
6								
7								
8								

Total # records something like

$$\leq \sum_{r=0}^k \binom{k}{r} k \cdot n^2 = \underline{2^k \cdot k \cdot n^2} \quad (\text{missing factor } k?)$$

Note that we only use clique and edge axioms — no functionality or ordering axioms

Balanced, complete $(k-1)$ -partite graphs

(i.e., all V_i satisfy $|V_i| = \lceil \frac{|V|}{k-1} \rceil$ or $|V_i| = \lfloor \frac{|V|}{k-1} \rfloor$)

good testbed for lower bound ideas.

If a ^{promising} lower bound idea identifies a property that holds for $(k-2)$ -partite graphs, then it is no longer so promising...

Proof idea used for PHP and Tcritic

BOTTLENECK COUNTING

- (1) Define random walk from \mathbb{I} in proof DAG
 - (2) Show that all random walks R_x have to pass through special "bottleneck node" w (informative record)
 - (3) Prove that for any fixed bottleneck node w

$$\Pr [R_x \text{ passes through } w] \leq \text{small}$$
 - (4) But

$$\begin{aligned} 1 &= \Pr [R_x \text{ passes through some } w] \\ &\leq \sum_{\substack{\text{bottle-} \\ \text{node}}} \Pr [R_x \text{ passes through fixed } w] \\ &\leq (\# \text{ bottlenecks}) \cdot \text{small} \end{aligned}$$
- so $(\# \text{ bottlenecks}) = \text{large}$

Twists:

- = Typically, small = $\exp(-n^5)$
but we can't get below than $n^{-\Omega(\ell)}$
- We will need pair of bottleneck nodes

Phrase proof in terms of ROBPs

ROBP path \Leftrightarrow partial truth value assignment
(because of read-once property)

Branching program: NODES a, b, c

Graph $G = (V, E)$ VERTICES u, v, w

$\beta(a)$ = maximal partial assignment contained
in every path from source to a
= all variable queries & answers obtained
along every path to a
"ASSIGNMENTS REMEMBERED" at a

Sourcepath = path from source to some node
Node a queries x : $x(a) = x$
If on path α x is queried and edge
 $(\alpha, b) \rightarrow$ taken, say " α sets x to b "

If \exists sourcepath α to a such that
- α sets x
- $\beta(a)$ doesn't contain x

then we say " x is FORGOTTEN AT a"

OBSERVATION 1

If x forgotten at a and \exists path a \rightsquigarrow b,
then x forgotten at b

Proof use read-once property

$\exists \alpha$: source \rightsquigarrow a not setting x

$\exists \alpha'$: source \rightsquigarrow a setting x

If $\exists \gamma$: a \rightsquigarrow b queries x , then composed
path $\alpha' \circ \gamma$ violates read-once property \square

Focus now on clique formula

If $x_{v,i}$ forgotten at a for some $v \in V$,
say "**INDEX i FORGOTTEN AT a** "

If source path α ends at clique axiom
 $\vee_{v \in V} x_{v,i}$ (which α thus falsifies), say
"PATH α RULES OUT INDEX i "

OBSERVATION 2

If source path α rules out index i at sink b , then i is not forgotten at b .

Proof All ^{source} paths α ending at b must have queried all $x_{v,i}, v \in V$ and have received answers 0, so no $x_{v,i}$ can be forgotten \square

Fix from now on arbitrary ROPD Π solving k -clique formula for $G = (V, E)$ without k -clique (and omitting ordering axioms). Want to show Π has size $n^{-\Omega(k)}$ a.a.s. if $G \sim G(n, n^{-2/(k-1)})$

Define RANDOM DISTRIBUTION β OVER PATHS α by following process, starting with (current node a):= source

(a) If $\beta(a) \cup \{x(a) \mapsto 1\}$ falsifies edge or functionalizing axiom, take 0-edge from a
(b) If $x(a) = x_{v,i}$ and i forgotten at a ,
take 0-edge from a set $x_{v,i} = 0$

(c) Otherwise, flip $n^{-\delta}$ -biased coin, $\delta > 0$
 $x(a) = x_{v,i} \mapsto 1$ with probability $n^{-\delta}$
 $x_{v,i} \mapsto 0$ otherwise

(a) + (b) "FORCED CHOICES" always setting $x_{v,i} = 0$

(c) "FREE CHOICE / COIN FLIP" but still highly likely to set $x_{v,i} = 0$

OBSERVATION 3

- (i) Any $\alpha \sim D$ ends by ruling out some index i (never falsifies functionality or edge axiom)
- (ii) Any $\alpha \sim D$ sets at most k variables $x_{v,i}$ to 1.

Proof Straightforward exercise using definition of random path distribution D .

FIRST PROOF IDEA (WON'T WORK)

Given any $\alpha \sim D$, define **WAYPOINTS**
(for $t \in \mathbb{N}^+$ fixed)

$a_0 :=$ source

$a_{i+1} :=$ first time after a_i we have

$\lceil k/t \rceil$ additional assignments $x_{v,i} \mapsto 1$
on α or sink of path, whichever comes first

Suppose α rules out i

Let $V_i = \{v \mid x_{v,i} \mapsto 0 \text{ between } a_{i-1} \text{ and } a_i\}$

$V = V_1 \cup V_2 \cup \dots \cup V_t$ for $t \leq t$

Must exist j such that $|V_j| \geq n/t$
i.e., between a_{j-1} and a_j we have
 $\geq n/t$ assignments $x_{v,i} = 0$

ROBP Π makes lots of progress in ruling out vertices on subpath (α_{j-1}, α_j) CLB V

Call such a pair (α_{j-1}, α_j) USEFUL*
We just proved:

OBSERVATION 4

Any path $\alpha \in \mathcal{D}$ passes through a useful* pair (a, b) with probability 1.

Now we want to fix any pair (ab) of nodes in Π and argue that

$$\Pr_{\alpha \in \mathcal{D}} [(\alpha, b) \text{ useful* for } \alpha] = n^{-\Omega(k)}$$

If we can do this, then we're done, since

$$\begin{aligned} 1 &= \Pr_{\alpha \in \mathcal{D}} [\exists (\alpha, b) \in \Pi^2 \text{ useful* for } \alpha] \\ &\leq \sum_{(a, b) \in \Pi^2} \Pr_{\alpha} [\text{fixed } (a, b) \text{ useful* for } \alpha] \\ &\leq |\Pi|^2 \cdot n^{-\Omega(k)} \end{aligned}$$

little bit of notation
overload
 $\alpha \in \mathcal{D}$:
 α sampled according to \mathcal{D}
 α is path in support of \mathcal{D}

OBSERVATION 5 (CRUCIAL)

Suppose \exists path $\alpha \in \mathcal{D}$ such that

(i) α passes through a and b

(ii) α assigns $x_{v,i} = 0$ for $v \in V_j'$

(iii) α rules out index i

Then any path α' through a and b has to set $x_{v,i} = 0$ for $v \in V_j'$

Proof Exercise. (Use read-once property) CLB VI

HOPEFUL CLAIM 6 (BROKEN)

For any $(a, b) \in \Pi^2$ it holds that

$$\Pr_{\alpha \sim D} [(\alpha, b) \text{ useful* for } \alpha] \leq n^{-\Omega(k)}$$

Attempt at proof Case analysis's

- (1) If (a, b) not useful for any path
⇒ probability 0
- (2) Suppose $\beta(a)$ contains $\Omega(k)$ assignments $x_{v,j} \mapsto 1$. Note that 1-assignments are never forced, but are free (and unlikely) coin flips. Then

$$\Pr_{\alpha \sim D} [\alpha \text{ passes through } a] \leq n^{-\delta k}$$

(By read-once property, only have one chance for every $x_{v,j} \mapsto 1$ in $\beta(a)$.)

- (3) Relatively few 1s in $\beta(a)$: Then Π knows almost nothing about the candidate clique found up until this point
Yet between a and b makes $\geq n/\ell$ assignments $x_{v,i} \mapsto 0$. Some of these might be forced, to be sure. But a polynomial fraction should be free coin flips, say #free choices n^δ for $\delta \gg \delta$

$$\begin{aligned} \Pr_{\alpha} [\text{all these coin flips yield 0}] &= (1 - n^{-\delta})^{n^\delta} \\ &= \exp(-n^{\delta-\delta}) , \text{ so great!} \end{aligned}$$

Except case 3 is broken. We could have essentially all choices forced. (Sanity check: Did we do anything so far that would fail for $(k-1)$ -partite graphs?) Need a better notion of "useful" (without the asterisk *)

ChB VII

ACTUAL PROOF (EXCEPT FOR SOME CALCULATIONS)

$$\begin{aligned} N(v) &= \{u \mid (u,v) \in E\} \text{ neighbours of } v \\ \widehat{N}(R) &= \bigcap_{v \in R} N(v) \text{ common neighbours of } R \\ \widehat{N}_W(R) &= \widehat{N}(R) \cap W \text{ common neighbours of } R \text{ in } W \end{aligned}$$

W ⊆ V is q -NEIGHBOUR-DENSE FOR $R \subseteq V$
if $|\widehat{N}_W(R)| \geq q$

W ⊆ V is (r,q) -NEIGHBOUR-DENSE if it is
 q -neighbour-dense $\forall R \subseteq V \quad |R| \leq r$

Means, in particular, that any clique of size r can be enlarged with vertex from W in $\geq q$ different ways

Should be likely to hold for moderate r, q for random graphs. But as r, q get larger, property becomes more unlikely

If (r, g) -neighbour-dense fails, want to find "mostly neighbour-dense" set W that fail only for few sets R . CDB VIII

Namely, exist set S such that $R \cap S$ is large for sets R that make W fail

In what follows, think of parameters as

$$r, r' = R(k) \quad r' \geq r$$

$$g' = n^{\delta'} \quad 1 \gg \delta' \gg \gamma > 0$$

$$s = n^\delta$$

$W \subseteq V$ is (r', r, g', s) -MOSTLY NEIGHBOUR-DENSE
if $\exists S \subseteq V \mid |S| \leq s$ such that
for every $R \subseteq V$, $|R| \leq r'$ for which W is
not g' -neighbour-dense, it holds
that $|R \cap S| \geq r$

Graph $G = (V, E)$ is (k, t, s, ϵ) -CLIQUE-DENSE
for $k \in \mathbb{N}^+$, $t, s, \epsilon \in \mathbb{R}^+$, $1 \leq t \leq k$,

if $\exists r, g \in \mathbb{R}^+$, $r \geq 4k/t^2$, such that:

- (1) V is (tr, tg) -neighbour-dense.
- (2) Every (r, g) -neighbour-dense set $W \subseteq V$ is (tr, r, g', s) -mostly neighbour-dense
for $g' = 3\epsilon ks^{1+\epsilon} \log s$

THEOREM 7

For constants $\epsilon > 0$, $\eta > 1$ and n large enough, if $G \sim G(n, n - 2\eta/(k-1))$ then with probability $\geq 1 - \exp(-\sqrt{n})$ G is (k, t, s, ϵ) -clique-dense for $t = 64\eta/\epsilon$ and $s = (n/\eta)^{1/2}$.

Proof A few pages of careful calculations...

THEOREM 8

If $k \in \mathbb{N}^+$ and $t, s, \epsilon \in \mathbb{R}^+$ are such that G is (k, t, s, ϵ) -clique-dense, then regular resolution requires length at least

$$\frac{1}{\sqrt{2}} s^{\frac{\epsilon k t^2}{2}} \xrightarrow{\text{to refute the clique formula for } G} n^{2(k)}$$

(without ordering axioms)

The proof of Thm 8 is what we will focus on. Reuse our previous approach, but with more careful definition of "useful"

For each $a \in \Pi$, $i \in [k]$, define

$$V_i^0(a) = \{v \in V \mid \beta(a) sets x_{v,i} to 0\}$$

$$V_i^1(a) = \{v \in V \mid \beta(a) sets x_{v,i} to 1\}$$

$$\beta(a)^2 = V_i^0(a) V_i^1(a)$$

Keep distribution \mathcal{D}

Probability of coin flip 1 is
 $rs^{-(1+\epsilon)} / 2^{ek} \approx s^{-(1+\epsilon)}$

Observation 3 still true, since D is essentially the same

Say $(a, b) \in \Pi^2$ (TRUEY) USEFUL if there is index such that

- (a) $V_i^1(b) = \emptyset$
- (b) i is not forgotten at b
- (c) $V_i^0(b) \setminus V_i^0(a)$ is (r, q) -neighbour-dense

For each useful pair (a, b) , fix (arbitrarily) index $i = i(a, b)$ witnessing that (a, b) useful

Path α USEFULLY TRAVESES useful pair (a, b) if α goes through a & b in that order and sets $\leq \lceil k/t \rceil$ variables so 1 between a and b (with a included and b excluded)

Theorem 8 follows by proof we did before from following two lemmas

LEMMA 9 (\Leftrightarrow Observation 4)

Every path $\alpha \sim D$ usefully traverses a useful pair

LEMMA 10 (\Leftrightarrow Helpful Claim 6)

For every useful pair (a, b) $\Pr_{\alpha \sim D} [\alpha \text{ usefully traverses } (a, b)] \leq 2s^{-\epsilon r/2}$

Proof of Lemma 9

Consider any $\alpha \sim D$. Ends in i^* th clique axiom for some $i^* \in [k]$

Obs 3 $\Rightarrow \alpha$ sets $\leq k$ variables so 1

Split α into t pieces source = $a_0 \rightsquigarrow a_1 \rightsquigarrow \dots \rightsquigarrow a_t$ so that between a_j and $a_{j+1} \leq \lceil k/t \rceil$ variables set to 1.

Clearly $V_{i^*}^{\circ}(a_j) = \emptyset$ and i^* not forgotten at a_j for all j . CL8X1

Remains to prove for some j that

$$W_j = V_{i^*}^{\circ}(a_j) \setminus V_{i^*}^{\circ}(a_{j-1}) \quad (\dagger)$$

is (r, q) -neighbour-dense. If so, (a_{j-1}, a_j) is a useful pair that is usefully traversed.

Towards contradiction, suppose no W_j is (r, q) -neighbour-dense.

That is, $\forall j \in [t] \exists R_j \subseteq V, |R_j| \leq r$, such that $|\widehat{N}_{W_j}(R_j)| \leq q$ (‡)

$$\text{Let } R = \bigcup_{j \in [t]} R_j$$

Path α ends in i^{th} clique axiom

$$\Rightarrow V_{i^*}^{\circ}(a_t) = V$$

i^* is not forgotten along the path

$$\Rightarrow V_{i^*}^{\circ}(a_{j-1}) \subseteq V_{i^*}^{\circ}(a_j) \quad \forall j \in [t]$$

So sets W_1, W_2, \dots, W_t in (\dagger)

form partition of V , i.e.

$$V = W_1 \dot{\cup} W_2 \dot{\cup} \dots \dot{\cup} W_t$$

Hence

$$|\widehat{N}_V(R)| = \sum_{j \in [t]} |\widehat{N}_{W_j}(R)| \leq \sum_{j \in [t]} |\widehat{N}_{W_j}(R_j)| \leq tq$$

by (\ddagger) . But this contradicts that V is $(\epsilon r, tq)$ -neighbour-dense. Lemma 9 follows □