Clique is hard on average for regular resolution

Ilario Bonacina, UPC Barcelona Tech
July 20, 2018

RaTLoCC, Bertinoro
How hard is to certify that a graph is Ramsey?

Ilario Bonacina, UPC Barcelona Tech
July 20, 2018
RaTLoCC, Bertinoro
Talk based on a joint work with:

A. Atserias

S. de Rezende

M. Lauria

J. Nordström

A. Razborov
How hard is to certify that a graph is Ramsey?

A graph $G$ with $n$ vertices we say that is $k$-Ramsey if it has no set of $k$ vertices forming a clique or an independent set. If $k = \lceil 2 \log_2 n \rceil$ we just say that $G$ is Ramsey.
How hard is to certify that a graph is Ramsey?

A graph $G$ with $n$ vertices we say that is $k$-Ramsey if it has no set of $k$ vertices forming a clique or an independent set. If $k = \lceil 2 \log_2 n \rceil$ we just say that $G$ is Ramsey.

Erdős-Rényi random graphs
A graph $G = (V, E) \sim G(n, p)$ is such that $|V| = n$ and each edge $\{u, v\} \in E$ independently with prob. $p \in [0, 1]$
How hard is to certify that a graph is Ramsey?

A graph $G$ with $n$ vertices we say that is $k$-Ramsey if it has no set of $k$ vertices forming a clique or an independent set. If $k = \lceil 2 \log_2 n \rceil$ we just say that $G$ is Ramsey.

**Erdős-Rényi random graphs**

A graph $G = (V, E) \sim \mathcal{G}(n, p)$ is such that $|V| = n$ and each edge ${u, v} \in E$ independently with prob. $p \in [0, 1]$

- if $p \ll n^{-2/(k-1)}$ then a.a.s. $G \sim \mathcal{G}(n, p)$ has no $k$-cliques
- A.a.s. $G \sim \mathcal{G}(n, \frac{1}{2})$ is Ramsey
How hard is to certify that a graph is Ramsey?

Construct a propositional formula $\Psi_{G,k}$ unsatisfiable if and only if “$G$ is $k$-Ramsey”
How hard is to certify that a graph is Ramsey?

Construct a propositional formula $\Psi_{G,k}$ unsatisfiable if and only if “$G$ is $k$-Ramsey”

$x_{v,j} \equiv \text{“$v$ is the } j\text{-th vertex of a } k\text{-clique in } G$

or the $j$-th vertex of a $k$-independent set”).
How hard is to certify that a graph is Ramsey?

Construct a propositional formula $\Psi_{G,k}$ unsatisfiable if and only if “$G$ is $k$-Ramsey”

$x_{v,j} \equiv \text{“} v \text{ is the } j\text{-th vertex of a } k\text{-clique in } G$

or the $j\text{-th vertex of a } k\text{-independent set”}.$

$$\bigvee_{v \in V} x_{v,i} \quad \text{for } i \in [k]$$

and

$$y \lor \neg x_{u,i} \lor \neg x_{v,j} \quad \text{for } i \neq j \in [k], u \neq v \in V, (u,v) \notin E$$

and

$$\neg y \lor \neg x_{u,i} \lor \neg x_{v,j} \quad \text{for } i \neq j \in [k], u \neq v \in V, (u,v) \in E$$
How hard is **to certify** that a graph is Ramsey?

**Resolution**

\[ y \lor \neg z \]

\[ x \]

\[ y \lor \neg c \]

\[ x \lor c \]

\[ \neg x \lor z \]

\[ \neg y \]

\[ \text{clause}_1 \lor \text{var} \]

\[ \text{clause}_2 \lor \neg \text{var} \]

\[ \text{clause}_1 \lor \text{clause}_2 \]

Tree-like = the proof DAG is a tree

Regular = no variable resolved twice in any source-to-sink path

Size = # of nodes in the proof DAG
How hard is to certify that a graph is Ramsey?

Resolution

\[ y \lor \neg z \]

\[ \neg x \]

\[ y \lor \neg c \]

\[ x \lor y \]

\[ x \lor c \]

\[ \neg x \lor z \]

\[ \neg y \]

\[ \text{clause}_1 \lor \text{var} \]

\[ \text{clause}_1 \lor \text{clause}_2 \]

\[ \text{clause}_2 \lor \neg \text{var} \]

Tree-like = the proof DAG is a tree

Regular = no variable resolved twice in any source-to-sink path

Size = # of nodes in the proof DAG
How hard is **to certify** that a graph is Ramsey?

Resolution

\[ y \lor \neg z \]

\[ \neg x \]

\[ y \lor \neg c \]

\[ x \lor y \]

\[ x \lor c \]

\[ y \lor z \]

\[ \neg x \lor z \]

\[ \neg y \]

\[ \text{clause}_1 \lor \text{var} \]

\[ \text{clause}_2 \lor \neg \text{var} \]

\[ \text{clause}_1 \lor \text{clause}_2 \]
How hard is to certify that a graph is Ramsey?

Resolution

\[ \neg y \lor \neg z \]
\[ \neg x \]
\[ \text{clause}_1 \lor \text{var} \]
\[ \text{clause}_2 \lor \neg \text{var} \]

\[ y \lor \neg c \]
\[ x \lor y \]
\[ x \lor c \]
\[ x \lor y \]
\[ y \lor z \]
\[ \neg x \lor z \]
\[ \neg y \]

Tree-like = the proof DAG is a tree
Regular = no variable resolved twice in any source-to-sink path
Size = # of nodes in the proof DAG

4
How hard is to certify that a graph is Ramsey?

Resolution

\[
\begin{align*}
\neg y \lor \neg z \\
\neg x \\
x \lor \neg z \\
\neg x \lor z \\
y \lor \neg c \\
x \lor c
\end{align*}
\]

\[
\begin{align*}
y \lor \neg c \\
x \lor y \\
x \lor y
\end{align*}
\]

\[
\begin{align*}
x \lor \neg z \\
y \lor z \\
z
\end{align*}
\]

\[
\begin{align*}
\neg x \lor z \\
\neg y
\end{align*}
\]

Tree-like = the proof DAG is a tree
Regular = no variable resolved twice in any source-to-sink path
Size = # of nodes in the proof DAG

\[
\begin{align*}
\text{clause}_1 \lor \text{var} \\
\text{clause}_2 \lor \neg \text{var} \\
\text{clause}_1 \lor \text{clause}_2
\end{align*}
\]
How hard is to certify that a graph is Ramsey?

Resolution

\[
\begin{align*}
&\neg x \lor \neg z \\
&x \lor y \\
&\neg y \lor \neg z \\
&\neg x \lor \neg c \\
&x \lor c \\
&y \lor z \\
&\neg x \lor \neg z \\
&\neg y \\
\end{align*}
\]
How hard is to certify that a graph is Ramsey?

Resolution

\[ y \lor \neg z \]
\[ x \lor \neg z \]
\[ x \lor \neg c \]
\[ y \lor \neg c \]
\[ x \lor c \]
\[ \neg y \lor z \]
\[ \neg x \lor z \]
\[ \neg \neg y \]

\[ \text{Tree-like} = \text{the proof DAG is a tree} \]
\[ \text{Regular} = \text{no variable resolved twice in any source-to-sink path} \]
\[ \text{Size} = \# \text{of nodes in the proof DAG} \]

\[ \text{clause}_1 \lor \text{var} \]
\[ \text{clause}_2 \lor \neg \text{var} \]
\[ \text{clause}_1 \lor \text{clause}_2 \]
How hard is to certify that a graph is Ramsey?

Resolution

Tree-like = the proof DAG is a tree
Regular = no variable resolved twice in any source-to-sink path
Size = # of nodes in the proof DAG
Regular?
Regular? No.
Regular? No. And none of the shortest proofs is regular \([HY87]\).

\[\text{[HY87]} \quad \text{Huang and Yu, 1987. A DNF without regular shortest consensus path.}\]
What is Resolution good for?

- algorithms routinely used to solve NP-complete problems (hardware verification, ...) are *somewhat* formalizable in resolution
- the state-of-the-art algorithms to solve k-clique (Bron-Kerbosch, Östergård, Russian dolls algorithms, ...) are formalizable in *regular* resolution
What is Resolution good for?

- algorithms routinely used to solve NP-complete problems (hardware verification, ...) are *somewhat* formalizable in resolution
- the state-of-the-art algorithms to solve k-clique (Bron-Kerbosch, Östergård, Russian dolls algorithms, ...) are formalizable in *regular* resolution

**[HKM16]** All possible 2-colorings of \{1, \ldots, 7825\} have a monochromatic Pythagorean triple.
What is Resolution good for?

- algorithms routinely used to solve NP-complete problems (hardware verification, ...) are somewhat formalizable in resolution
- the state-of-the-art algorithms to solve k-clique (Bron-Kerbosh, Östergård, Russian dolls algorithms, ...) are formalizable in regular resolution

[HKM16] All possible 2-colorings of \{1, \ldots, 7825\} have a monochromatic Pythagorean triple.

This slide is too small to contain the 200Terabyte resolution proof...

Resolution size

Let $\phi$ be an conjunction of clauses in $N$ variables with $|\phi| = N^{O(1)}$

$S(\phi) = \text{minimum size of a resolution refutation of } \phi$

$S_{tree}(\phi) = \text{minimum size of a tree-like resolution refutation of } \phi$

$S_{reg}(\phi) = \text{minimum size of a regular resolution refutation of } \phi$
Resolution size

Let $\phi$ be an conjunction of clauses in $N$ variables with $|\phi| = N^{O(1)}$

- $S(\phi) = \text{minimum size of a resolution refutation of } \phi$
- $S_{\text{tree}}(\phi) = \text{minimum size of a tree-like resolution refutation of } \phi$
- $S_{\text{reg}}(\phi) = \text{minimum size of a regular resolution refutation of } \phi$

- for every $\phi$, $S(\phi) \leq S_{\text{reg}}(\phi) \leq S_{\text{tree}}(\phi)$
  (and there are examples of exponential separations)
- for every $\phi$, $S_{\text{tree}}(\phi) = 2^{\mathcal{O}(N)}$
How **hard** is to certify that a graph is Ramsey?

Theorem? (folklore)

\[ \Psi_{G,k}, \text{ whenever unsatisfiable, has } S_{\text{tree}}(\Psi_{G,k}) = n^{O(k)} \]
How hard is to certify that a graph is Ramsey?

Theorem? (folklore)

Ψ_{G,k}, whenever unsatisfiable, has S_{tree}(Ψ_{G,k}) = n^{O(k)}

Theorem [LPRT17]

If G is a Ramsey graph in n vertices and k = \lceil 2 \log n \rceil then
S_{tree}(Ψ_{G,k}) = n^{\Omega(\log n)}.

[LPRT17] Lauria, Pudlák, Rödl, and Thapen, 2017. The complexity of proving that a graph is Ramsey.
How hard is to certify that a graph is Ramsey?

Theorem? (folklore)
\[ \psi_{G,k}, \text{ whenever unsatisfiable, has } S_{\text{tree}}(\psi_{G,k}) = n^{O(k)} \]

Theorem [LPRT17]
If \( G \) is a Ramsey graph in \( n \) vertices and \( k = \lceil 2 \log n \rceil \) then
\[ S_{\text{tree}}(\psi_{G,k}) = n^{\Omega(\log n)}. \]

Theorem
If \( G \sim \mathcal{G}(n, \frac{1}{2}) \) (hence in particular a.a.s. \( G \) is Ramsey) and \( k = \lceil 2 \log n \rceil \) then \( S_{\text{reg}}(\psi_{G,k}) \text{ a.a.s. } n^{\Omega(\log n)}. \)

[LPRT17] Lauria, Pudlák, Rödl, and Thapen, 2017. The complexity of proving that a graph is Ramsey.
How **hard** is to certify that a graph is Ramsey?

Theorem? (folklore)
\[ \Psi_{G,k}, \text{ whenever unsatisfiable, has } S_{tree}(\Psi_{G,k}) = n^{O(k)} \]

**Theorem [LPRT17]**
If \( G \) is a Ramsey graph in \( n \) vertices and \( k = \lceil 2 \log n \rceil \) then
\[ S_{tree}(\Psi_{G,k}) = n^{\Omega(\log n)}. \]

**Theorem**
If \( G \sim G(n, \frac{1}{2}) \) (hence in particular a.a.s. \( G \) is Ramsey) and \( k = \lceil 2 \log n \rceil \) then \( S_{reg}(\Psi_{G,k}) \text{ a.a.s. } n^{\Omega(\log n)}. \)

**Open Problem**
Let \( G \) be a Ramsey graph in \( n \) vertices and let \( k = \lceil 2 \log n \rceil \). Is it true that \( S(\Psi_{G,k}) = n^{\Omega(\log n)} \)?

---

How hard is to certify that a graph does not contain a $k$-clique?

Construct a propositional formula $\Phi_{G,k}$ unsatisfiable if and only if “$G$ does not contain a $k$-clique”

We already have it: $\Phi_{G,k} = \Psi_{G,k} \upharpoonright y = 0$
How hard is to certify that a graph does not contain a $k$-clique?

Construct a propositional formula $\Phi_{G,k}$ unsatisfiable if and only if “$G$ does not contain a $k$-clique”

We already have it: $\Phi_{G,k} = \Psi_{G,k} \upharpoonright y=0$

$x_{v,j} \equiv \text{“} v \text{ is the } j\text{-th vertex of a } k\text{-clique in } G \text{”}. $
How hard is to certify that a graph does not contain a $k$-clique?

Construct a propositional formula $\Phi_{G,k}$ unsatisfiable if and only if “$G$ does not contain a $k$-clique”

We already have it: $\Phi_{G,k} = \Psi_{G,k} \upharpoonright y = 0$

$x_{v,j} \equiv \text{“} v \text{ is the } j\text{-th vertex of a } k\text{-clique in } G\text{”}$.

$$\bigvee_{v \in V} x_{v,i} \quad \text{for } i \in [k]$$

and

$$\neg x_{u,i} \lor \neg x_{v,j} \quad \text{for } i \neq j \in [k], u \neq v \in V, (u, v) \notin E$$
How hard is to certify that a **graph does not contain a k-clique**?

Construct a propositional formula $\Phi_{G,k}$ unsatisfiable if and only if

“$G$ does not contain a $k$-clique”

We already have it: $\Phi_{G,k} = \Psi_{G,k} \upharpoonright y=0$

$$x_{v,j} \equiv \text{“} v \text{ is the } j\text{-th vertex of a } k\text{-clique in } G \text{”}.$$ 

$$\bigvee_{v \in V} x_{v,i} \quad \text{for } i \in [k]$$

and

$$\neg x_{u,i} \lor \neg x_{v,j} \quad \text{for } i \neq j \in [k], u \neq v \in V, (u, v) \notin E$$

lower bounds on $S(\Phi_{G,k})$ imply lower bounds on $S(\Psi_{G,k})$
Overview of the literature: Upper Bounds

[\sim BGL13] if $G$ is $(k - 1)$-colorable then
\[ S_{\text{reg}}(\Phi_{G,k}) \leq 2^k k^2 n^2 \]

[folklore] $\Phi_{G,k}$, whenever unsatisfiable, has
\[ S_{\text{tree}}(\Phi_{G,k}) = n^{O(k)} \]

If $G$ is the complete $(k-1)$-partite graph, then $S_{tree}(\Phi_{G,k}) = n^{\Omega(k)}$.

The same holds for $G \sim \mathcal{G}(n,p)$ with suitable edge density $p$.

for $n^{5/6} \ll k < \frac{n}{3}$ and $G \sim \mathcal{G}(n,p)$ (with suitable edge density $p$), then $S(\Phi_{G,k}) \overset{\text{a.a.s.}}{=} 2^{n^{\Omega(1)}}$

if we encode $k$-clique using some other propositional encodings (e.g. in binary) we get $n^{\Omega(k)}$ size lower bounds for resolution

---


Main Result (simplified versions)

Main Theorem (version 1)
Let $G \sim G(n, p)$ be an Erdős-Rényi random graph with, for simplicity, $p = n^{-4/(k-1)}$ and let $k \leq n^{1/2-\epsilon}$ for some arbitrary small $\epsilon$. Then, $S_{reg}(\Phi_{G,k}) \overset{\text{a.a.s.}}{=} n^{\Omega(k)}$. 

Main Theorem (version 2)
Let $G \sim G(n, 1/2)$, then $S_{reg}(\Phi_{G,k}) \overset{\text{a.a.s.}}{=} n^{\Theta(\log n)}$ for $k = O(\log n)$ and $S_{reg}(\Phi_{G,k}) \overset{\text{a.a.s.}}{=} n!$ for $k = o(\log^2 n)$. 

Main Result (simplified versions)

Main Theorem (version 1)
Let $G \sim \mathcal{G}(n, p)$ be an Erdős-Rényi random graph with, for simplicity, $p = n^{-4/(k-1)}$ and let $k \leq n^{1/2-\epsilon}$ for some arbitrary small $\epsilon$. Then, $S_{reg}(\Phi_{G,k}) \overset{a.a.s.}{=} n^{\Omega(k)}$.

the actual lower bound decreases smoothly w.r.t. $p$
Main Result (simplified versions)

Main Theorem (version 1)
Let $G \sim \mathcal{G}(n, p)$ be an Erdős-Rényi random graph with, for simplicity, $p = n^{-4/(k-1)}$ and let $k \leq n^{1/2-\epsilon}$ for some arbitrary small $\epsilon$. Then, $S_{\text{reg}}(\Phi_{G,k}) \text{ a.a.s. } n^{\Omega(k)}$.

the actual lower bound decreases smoothly w.r.t. $p$

Main Theorem (version 2)
Let $G \sim \mathcal{G}(n, \frac{1}{2})$, then

$$S_{\text{reg}}(\Phi_{G,k}) \text{ a.a.s. } n^{\Omega(\log n)} \text{ for } k = \mathcal{O}(\log n)$$

and

$$S_{\text{reg}}(\Phi_{G,k}) \text{ a.a.s. } n^{\omega(1)} \text{ for } k = o(\log^2 n).$$
Focus on proving the following.

**Theorem**

Let $k = \lceil 2 \log n \rceil$ and $G \sim G(n, \frac{1}{2})$, then $S_{reg}(\Phi_{G,k}) \overset{\text{a.a.s.}}{=} n^{\Omega(\log n)}$
\( \hat{N}_W(R) \) is the set of common neighbors of \( R \) in \( W \)
\( \hat{N}_W(R) \) is the set of common neighbors of \( R \) in \( W \).

\( W \) is \((r,q)\)-dense if for every subset \( R \subseteq V \) of size \( \leq r \), it holds

\[ |\hat{N}_W(R)| \geq q \]
$\hat{N}_W(R)$ is the set of common neighbors of $R$ in $W$

$W$ is $(r, q)$-dense if for every subset $R \subseteq V$ of size $\leq r$, it holds $|\hat{N}_W(R)| \geq q$

**Theorem 1**

Let $k = \lceil 2 \log n \rceil$. A.a.s. $G = (V, E) \sim \mathcal{G}(n, \frac{1}{2})$ satisfies the following:

(✳) $V$ is $(\frac{k}{50}, \Theta(n^{0.9}))$-dense; and

(✳✳) For every $(\frac{k}{10000}, \Theta(n^{0.9}))$-dense $W \subseteq V$ there exists $S \subseteq V$, $|S| \leq \sqrt{n}$ s.t. for every $R \subseteq V$, with $|R| \leq \frac{k}{50}$ and $|\hat{N}_W(R)| < \tilde{\Theta}(n^{0.6})$ it holds that $|R \cap S| \geq \frac{k}{10000}$. 
\( \hat{N}_W(R) \) is the set of common neighbors of \( R \) in \( W \).

\( W \) is \((r, q)\)-dense if for every subset \( R \subseteq V \) of size \( \leq r \), it holds \(|\hat{N}_W(R)| \geq q\).

**Theorem 1**

Let \( k = \lceil 2\log n \rceil \). A.a.s. \( G = (V, E) \sim \mathcal{G}(n, \frac{1}{2}) \) satisfies the following:

\( (*) \) \( V \) is \((\frac{k}{50}, \Theta(n^{0.9}))\)-dense; and

\( (**) \) For every \((\frac{k}{10000}, \Theta(n^{0.9}))\)-dense \( W \subseteq V \) there exists \( S \subseteq V \), \(|S| \leq \sqrt{n} \) s.t. for every \( R \subseteq V \), with \(|R| \leq \frac{k}{50} \) and \(|\hat{N}_W(R)| < \tilde{\Theta}(n^{0.6}) \) it holds that \(|R \cap S| \geq \frac{k}{10000} \).

**Theorem 2**

Let \( k = \lceil 2\log n \rceil \). For every \( G \) satisfying properties \((*)\) and \((**)\),

\[ S_{reg}(\Phi_G, k) = n^{\Omega(\log n)} \]
Regular resolution $\equiv$ Read-Once Branching Programs
Regular resolution $\equiv$ Read-Once Branching Programs

\[
\neg x \lor y \lor \neg z
\]

\[
x \lor \neg y
\]

\[
x \lor c
\]

\[
y \lor \neg c
\]

\[
x \lor y
\]

\[
\neg x \lor z
\]

\[
y \lor \neg x
\]
**Haken bottleneck counting idea**

*Lemma 1*

Every random path \( \sim \mathcal{D} \) in the ROBP passes through a bottleneck node.

*Lemma 2*

Given any bottleneck node \( b \) in the ROBP, \( \Pr \sim \mathcal{D}[b^2] \leq n \cdot (k) \). Then, it is trivial to conclude:

\[
\frac{1}{\Pr \sim \mathcal{D}[b^2 | \text{ROBP bottleneck and } b^2] \cdot \max_{b \text{bottleneck in the ROBP}} \Pr \sim \mathcal{D}[b^2]} \leq n \cdot (k) \]

\[16\]
“Lemma 1”
Every random path $\gamma \sim D$ in the ROBP passes through a bottleneck node.
“Lemma 1”
Every random path $\gamma \sim D$ in the ROBP passes through a bottleneck node.

“Lemma 2”
Given any bottleneck node $b$ in the ROBP,

$$\Pr_{\gamma \sim D} [b \in \gamma] \leq n^{-\Theta(k)}.$$
Haken bottleneck counting idea

“Lemma 1”
Every random path $\gamma \sim D$ in the ROBP passes through a bottleneck node.

“Lemma 2”
Given any bottleneck node $b$ in the ROBP,

\[
\Pr_{\gamma \sim D} [b \in \gamma] \leq n^{-\Theta(k)}.
\]

Then, it is trivial to conclude:

\[
1 = \Pr_{\gamma \sim D} \left[ \exists b \in \text{ROBP} \text{ bottleneck and } b \in \gamma \right]
\leq |\text{ROBP}| \cdot \max_{b \text{ bottleneck in the ROBP}} \Pr_{\gamma \sim D} [b \in \gamma]
\leq |\text{ROBP}| \cdot n^{-\Theta(k)}
\]
The real bottleneck counting
\( \beta(c) = \max \) (partial) assignment contained in all paths from the source to \( c \)
\( \beta(c) = \max \text{ (partial) assignment contained in all paths from the source to } c \)

\( j \in [k] \) is forgotten at \( c \) if no sink reachable from \( c \) has label

\( \bigvee_{v \in V} x_{v,j} \)
\[ \beta(c) = \max \text{ (partial) assignment contained in all paths from the source to } c \]

\( j \in [k] \) is forgotten at \( c \) if no sink reachable from \( c \) has label \( \bigvee_{v \in V} x_{v,j} \)

**The random path** \( \gamma \)

- if \( j \) forgotten at \( c \) or \( \beta(c) \cup \{x_{v,j} = 1\} \) falsifies a short clause of \( \Phi_{G,k} \)
  then continue with \( x_{v,j} = 0 \)
- otherwise toss a coin and with prob. \( \Theta(n^{-0.6}) \)
  continue with \( x_{v,j} = 1 \)
\[ V_j^0(a) = \{ v \in V : \beta(a)(x_{v,j}) = 0 \} \]
$$V_j^0(a) = \{ v \in V : \beta(a)(x_{v,j}) = 0 \}$$

**Lemma 1**
For every random path $\gamma$, there exists two nodes $a, b$ in the ROBP s.t.

1. touches $a$, sets 6 $\ll 200$ variables to 1 and then touches $b$;
2. there exists a $j^\ast \geq \ll 20000$ not-forgotten at $b$ and such that $V_j^0(b)$ is $\ll 20000, j^\ast$-dense.

**Lemma 2**
For every pair of nodes $(a, b)$ in the ROBP satisfying point (2) of Lemma 1, $\Pr$
\[ V_j^0(a) = \{ v \in V : \beta(a)(x_{v,j}) = 0 \} \]

**Lemma 1**

For every random path \( \gamma \), there exists two nodes \( a, b \) in the ROBP s.t.

1. \( \gamma \) touches \( a \), sets \( \leq \left\lceil \frac{k}{200} \right\rceil \) variables to 1 and then touches \( b \);
\( V_j^0(a) = \{ v \in V : \beta(a)(x_{v,j}) = 0 \} \)

**Lemma 1**

For every random path \( \gamma \), there exists two nodes \( a, b \) in the ROBP s.t.

1. \( \gamma \) touches \( a \), sets \( \leq \left\lfloor \frac{k}{200} \right\rfloor \) variables to 1 and then touches \( b \);
2. there exists a \( j^* \in [k] \) not-forgotten at \( b \) and such that \( V_{j^*}^0(b) \setminus V_{j^*}^0(a) \) is \( (\frac{k}{10000}, \Theta(n^{0.9})) \)-dense.

Go to Conclusions
\[ V_j^0(a) = \{ v \in V : \beta(a)(x_{v,j}) = 0 \} \]

**Lemma 1**

For every random path \( \gamma \), there exists two nodes \( a, b \) in the ROBP s.t.

1. \( \gamma \) touches \( a \), sets \( \leq \left[ \frac{k}{200} \right] \) variables to 1 and then touches \( b \);
2. there exists a \( j^* \in [k] \) not-forgotten at \( b \) and such that \( V_j^0(b) \setminus V_j^0(a) \) is \( \left( \frac{k}{10000}, \Theta(n^{0.9}) \right) \)-dense.

**Lemma 2**

For every pair of nodes \((a, b)\) in the ROBP satisfying point (2) of Lemma 1,

\[
\Pr[\gamma \text{ touches } a, \text{ sets } \leq \left[ \frac{k}{200} \right] \text{ vars to 1 and then touches } b] \leq n^{-\Theta(k)}
\]
Proof sketch of Lemma 2

Let $E=\\gamma$ touches $a$, sets $\leq \lceil k/200 \rceil$ vars to 1 and then touches $b$” and let $W = V_j^0(b) \setminus V_j^0(a)$
Proof sketch of Lemma 2

Let $E$ = “$\gamma$ touches $a$, sets $\leq \lceil k/200 \rceil$ vars to 1 and then touches $b$” and let $W = V_{j^*}^0(b) \setminus V_{j^*}^0(a)$

**Case 1:** $V^1(a) = \{ v \in V : \exists i \in [k] \beta(a)(x_{v,i}) = 1 \}$ has large size ($\geq k/20000$). Then $\Pr[E] \leq n^{-\Theta(k)}$ because of the prob. of 1s in the random path $\gamma$ and a Markov chain argument.
Proof sketch of Lemma 2

Let \( E = \\text{“} \gamma \text{ touches } a, \text{ sets } \leq \lfloor k/200 \rfloor \text{ vars to 1 and then touches } b \text{”} \)
and let \( W = V_j^0(b) \setminus V_j^0(a) \)

**Case 1:** \( V_1^1(a) = \{v \in V : \exists i \in [k] \beta(a)(x_{v,i}) = 1\} \) has large size (\( \geq k/20000 \)). Then \( \Pr[E] \leq n^{-\Theta(k)} \) because of the prob. of 1s in the random path \( \gamma \) and a Markov chain argument.

**Case 2.1:** \( V_1^1(a) \) is not large but many (\( \geq \tilde{\Theta}(n^{0.6}) \)) vertices in \( W \) are set to 0 by coin tosses.

So \( \Pr[E \land W \text{ has many coin tosses}] \leq n^{-\Theta(k)} \) again by a Markov chain argument as in **Case 1**.
Proof sketch of Lemma 2

Let $E=$ “$\gamma$ touches $a$, sets $\leq \lceil k/200 \rceil$ vars to 1 and then touches $b$” and let $W = V^0_{j^*}(b) \setminus V^0_{j^*}(a)$

**Case 1:** $V^1(a) = \{ v \in V : \exists i \in [k] \beta(a)(x_{v,i}) = 1 \}$ has large size ($\geq k/20000$). Then $\Pr[E] \leq n^{-\Theta(k)}$ because of the prob. of 1s in the random path $\gamma$ and a Markov chain argument.

**Case 2.1:** $V^1(a)$ is not large but many ($\geq \tilde{\Theta}(n^{0.6})$) vertices in $W$ are set to 0 by coin tosses. So $\Pr[E \land W$ has many coin tosses] $\leq n^{-\Theta(k)}$ again by a Markov chain argument as in **Case 1**.

**Case 2.2:** $V^1(a)$ is not large and not many vertices in $W$ are set to 0 by coin tosses. Then many of the 1s set by the random path $\gamma$ between $a$ and $b$ must belong to a set of size at most $\sqrt{n}$, by the new combinatorial property (**)).

So $\Pr[E \land W$ has not many coin tosses] $\leq n^{-\Theta(k)}$. 

19
Open Problem: How hard is to prove that a graph is Ramsey?

Let $G$ be a Ramsey graph in $n$ vertices and let $k = \lceil 2 \log n \rceil$. Is it true that $S(\Psi_{G,k}) = n^{\Omega(\log n)}$?

([LPRT17] proved this but for a binary encoding of “$G$ is Ramsey”)

Thanks!

bonacina@cs.upc.edu