Independent Sets in Bounded-Degree, 3-colorable graphs

Input: - $G = (V, E)$
- $\deg(v) \leq d \forall v$
- Promise: $G$ is 3-colorable

Output: - $I \subseteq V$, $I$ is independent

Obj: - Maximize $|I|$

Randomized Rounding with a fix (even w/o $\oplus$):

- Select $I_1$ where $v \in V$ is sampled w.r.t. $p$
- Remove any pair $i, j$ from $I_1$ if $(i, j) \in E$
- Return remaining set $I = I_1 - I_2$

$\mathbb{E}[|I_1|] = np$

$\mathbb{E}[|I_2|] \leq \sum_{(i, j) \in E} \mathbb{P}[i \in I_1 \land j \in I_1]$

$\leq \frac{nd}{2} \cdot p^2$

$\therefore \mathbb{E}[|I|] \geq np - \frac{nd}{2} \cdot p^2$
\[ \mathbb{E} \left[ |I| \right] \geq n[p - \frac{nd}{2}p^2] \\
= n \left( p - \frac{dp^2}{2} \right) \]

\[ \therefore \text{If } p \text{ was chosen so that } p = \frac{dp^2}{2} \text{ i.e. } p = \frac{1}{d} \text{; then...} \]

\[ \mathbb{E} \left[ |I| \right] \geq \frac{n}{2d} \]

linear in d.

Today, we see how to get a sublinear dep.

on d when \( \ast \), i.e., G is 3-colorable.

Uses SDPs and a "simple" rounding trick.

**Two ideas:**

1. Use G is 3-colorable to obtain an embedding of G to \( S_n \), the n-dimn sphere, with edges' esp. being "far apart".

2. Use a similar idea as above (RR with a fix) to get a better IS. Except of "ind sampling", it'll be correlated
What can be said about 3-colorable Graphs?

→ Checking if \( G \) is 3-COL or not is NP-complete. No exact characterization is known.

→ We use a necessary condition.

Note: \( G \) is 3-COL

\[ \max IS \geq \frac{n}{3} \]

We will be nowhere close to finding this large a IS.

\[ \exists \text{ an embedding } \phi : V \rightarrow \mathbb{R}^2 \]

s.t. \[ \forall i , \quad \| \phi(i) \|_2 = 1 \quad (\text{unit circle}) \]

\[ \text{if } (i,j) \in E , \quad \angle \phi(i), \phi(j) = 120^\circ \]
If \( G \) is 3-col, the following has a feasible soln:
\[
\{ (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n : \\
\forall (i,j) \in E : \langle v_i, v_j \rangle = -\frac{1}{2} \\
\|v_i\|_2 = 1 \}
\]

SDP formulation: The following system has a feasible soln
\[
\{ X \in \mathbb{R}^{n \times n} : \\
X_{ii} = 1, \quad \forall i = 1 \ldots n, \\
X_{ij} = -\frac{1}{2}, \quad \forall (i,j) \in E \\
X \geq 0 \}
\]

\( G \) is 3-col \( \implies \) SDP-col is feasible.

We may assume we have unit vectors
\[
\{ v_1, \ldots, v_n \} \text{ st } \langle v_i, v_j \rangle = -\frac{1}{2} \text{ for } i \neq j.
\]
all edges \((i,j) \in E\)  

--- End of Part 1. --- 

**Randomized Rounding with a finer part deux**

- Two forces at loggerheads
  
  - Want to sample so that lots of vertices in \(I_1\).
  
  - But not so aggressively that many edges enter \(I_1\).

- Pick an edge \((i,j)\):

  \[
  \langle v_i, v_j \rangle = -\frac{1}{2}
  \]

  \[
  \| v_i + v_j \|^2 = \| v_i \|^2 + \| v_j \|^2 + 2 \langle v_i, v_j \rangle \\
  = 2 - 1 = 1
  \]
\[ \therefore \|w_i + w_j\| = 1 \quad \ldots \quad \text{while for non-edges it could be as large as 2.} \]

\[ \text{ALGO: (Karger-Motwani-Sudan aka KMS alg)} \]

- Sample a random unit gaussian \( g \) in \( \mathbb{R}^n \).

\[ I_1 := \{ i \mid \langle v_i, g \rangle \geq c \} \]

for some \( c \) to be chosen later.

- \( I = I_1 - I_2 \), where \( I_2 \) are the left of edges in \( I_1 \).

\[ \text{Facts about Gaussians} \]

\( 1 \)-dimn: \( X \sim N(0, \sigma) \)

\[ \Rightarrow P_X[ X = x ] = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/(2\sigma^2)} \]

\( n \)-dimn: \( g = (g_1, \ldots, g_n) \)
each $g_i \sim N(0, 1)$ independent

- Unit Gaussian: $g_i \sim \frac{p}{||g||_2}$

- If $g$ is a unit Gaussian in $\mathbb{R}^n$, and $v$ is any fixed vector in $\mathbb{R}^n$, then $\langle v, g \rangle$ is a random variable with $\langle v, g \rangle \sim N(0, ||v||)$

- Sum of independent Gaussian r.v.s is Gaussian
- Variances add up

- "Error function" / "Quantile function"

$$erf(t) := P_x \left[ X > t \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} dx$$

Bounds:

$\forall t > 0$:

$$\left( \frac{1}{t} \cdot \frac{1}{t^3} \right) e^{-t^2/2} \leq \sqrt{2\pi} \cdot erf(t) \leq \frac{1}{t} \cdot e^{-t^2/2}$$
Analysis of the KMS algorithm

\[
\mathbb{E}[|I_1|] = \sum_{i \in V} \mathbb{P}(\langle \nu_i, g \rangle \geq c) \\
= n \cdot \text{erf}(c)
\]

\[
\mathbb{E}[|I_2|] = \sum_{(i,j) \in E} \mathbb{P}(\langle \nu_i, g \rangle \geq c \land \langle \nu_j, g \rangle \geq c) \\
\leq \sum_{(i,j) \in E} \mathbb{P}(\langle \nu_i + \nu_j, g \rangle \geq 2c) \\
\leq \frac{nd}{2} \cdot \text{erf} \left( \frac{2c}{\|\nu_i + \nu_j\|} \right)
\]
\[ \mathbb{E}[|I|] \geq n \operatorname{erf}(c) - \frac{nd}{2} \operatorname{erf}(2c) \]

\[ \approx n \left[ \frac{1}{c} e^{-c^2/2} - \frac{d}{2} \frac{1}{2c} e^{-2c^2} \right] \]

if \[ e^{-c^2/2} \approx \frac{d}{2} e^{-2c^2} \]

i.e. \[ e^{\frac{3c^2}{2}} \approx \frac{d}{2} \]

\[ c = \sqrt{\frac{2}{3} \ln\left(\frac{d}{2}\right)} \]

\[ \approx \frac{n}{\sqrt{\frac{2}{3} \ln\left(\frac{d}{2}\right)}} \cdot \left[ \frac{1}{4} e^{-\frac{1}{3} \ln\left(\frac{d}{2}\right)} \right] \]

\[ \geq \frac{n}{A \cdot d^{1/3} \sqrt{\ln d}} \]

for some constant \( A \)
Theorem: If $G$ has max-degree $d$ and is $3$-colorable, then one can find an independent set of size $\geq \frac{n}{\log d}$.

Corollary: Once we find an independent set, we can "pluck" it out giving it one color and repeat. Thus we can color $3$-colorable graphs with $O(\sqrt[3]{d})$ colors.

Using "another trick", this gives a coloring of $3$-col. graphs using $\tilde{O}(n^{0.25})$ colors.... sounds ridiculous?

The best known algorithm till date is $\tilde{O}(n^{0.38..})$ colors. From 2012!