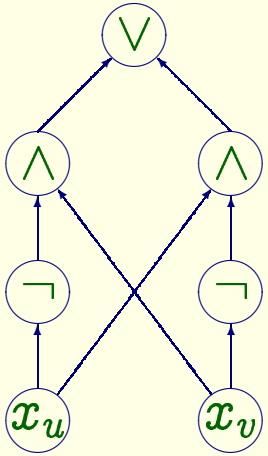
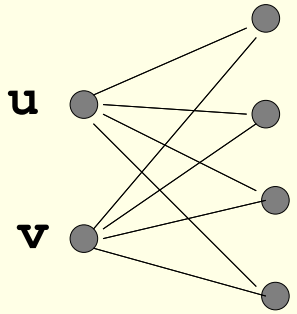
Graphs, Circuits and Communication
A possible attack on old problems?

S. Jukna



Circuit computes $x_u \oplus x_v$ $\Rightarrow$ "represents" the graph



The Problem

$C(f_m)$ = min # of AND, OR, NOT gates to compute f_m

Ultimate Goal: Prove $C(f_m) \geqslant n^{\alpha}$ with $\alpha \to \infty$ for $f_m \in$ NP

$\Rightarrow$ P $\neq$ NP and many other good things (security of RSA, etc)

$\Rightarrow$ 1.000.000,- USD among others ...

"Moderate" Goal: Prove $C(f_m) \geqslant \alpha m$ for $\alpha \to \infty$

"Minor" Goal: Prove $C(f_m) \geqslant \alpha m$ for Log-depth circuits

Parallel time $O(\log m)$ $\Rightarrow$ super-linear # of processors

$>$ 50 years intensive research $\Rightarrow$ even $C(f_m) \geqslant 5 \cdot m$ not known !

Razborov/Rudich (1994): "Natural" proofs will not work!

"Natural" = proof works for "almost all" functions $\Rightarrow$ largeness condition

Idea: Use graphs to avoid this obstacle !

The Plan

- Attack strategy: use graphs to define "complex" functions
- Attack on boolean formulas $\Rightarrow$ the "edge/non-edge" game
- Attack on Log-depth circuits $\Rightarrow$ Σ_3 circuits!
- The case of $\Sigma_3^{\oplus}$ circuits cracked!
- How to crack "pure" Σ_3? $\Rightarrow$ Graph covering problems
- Disproof of the Single Level Conjecture
- Open problems

The idea

Goal: Define an explicit "complicated" boolean function f

Idea: "complicated" $\Rightarrow$ complicated graph structure

Graph $G = (U, W, E)$ with $U = W = \{0,1\}^m$ $\Rightarrow$ gives boolean function

$$f_{2m}(uv) = 1 \iff uv \in E$$

$\Rightarrow$ characteristic function of G

Want: Graph G complicated $\Rightarrow$ $C(f_{2m}) \geqslant$ large

Example: "Complicated" = needs many cliques to cover all edges

$\Rightarrow$ nondeterministic communication complexity of f_{2m}

Graph complexity

Graph $G = (U, W, E)$ complicated := needs large circuits to represent it

Circuits: Inputs = stars Operations $\cup, \cap$

Variables X = vertices (not edges!) = $\{x_u : u \in V\}$, $V = U \cup W$

Circuit $F : 2^V \to \{0,1\}$ $\Rightarrow$ accepts/rejects subsets $S \subseteq V$

Circuit $F(X)$ represents $G = (U, W, E)$ if it accepts all edges and rejects all non-edges:

$$F(0, \ldots, 0, \overset{u}{1}, 0, \ldots, 0, \overset{v}{1}, 0, \ldots, 0) = 1 \quad \Longleftrightarrow \quad uv \in E$$

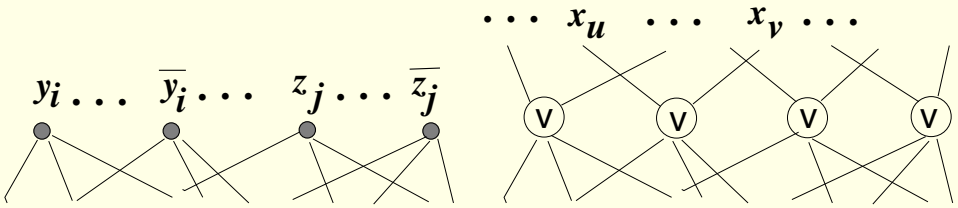
$\Rightarrow$ on inputs $a \in \{0,1\}^V$ with $|a| \neq 2$ can do what it wants !

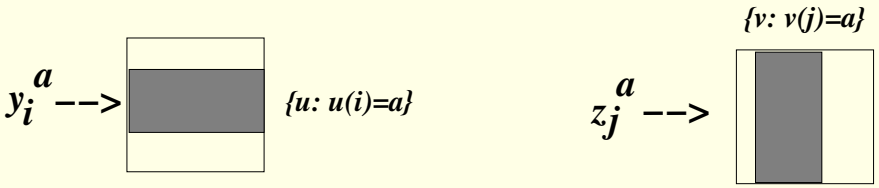
$\Rightarrow$ needs not to compute $f_G = \vee_{uv \in E} x_u x_v$!

Graph complexity = min. size of a circuit representing a graph

From Graphs to Boolean Functions

F computes $f_{2m}(uv) = 1$ iff $uv \in G$ $\Rightarrow$ replace y_i^a by $\bigvee_{u \in U : u(i)=a} x_u$





$\Rightarrow$ no negated inputs in F^+ $\Rightarrow$ monotone circuit ! ! !

Magnification Lemma: *If F computes f then F^+ represents G*

OR or Parity gates on the top $\Rightarrow$ $size(F^+) = size(F)$!

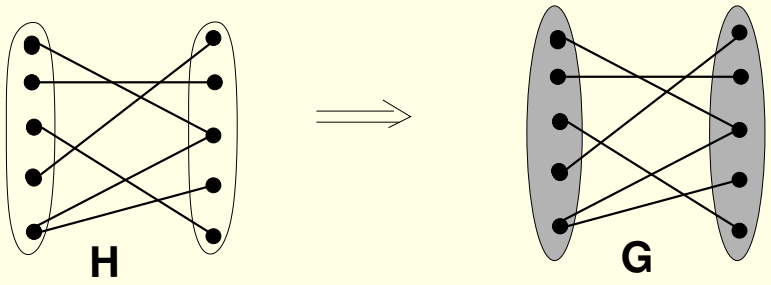
G has $n = 2^m$ vertices $\Rightarrow$
$size(F^+) \geqslant n^\epsilon$ $\Rightarrow$ $size(F) \geqslant 2^{\epsilon m}$ $\Rightarrow$ exponential in m ! ! !

Graphs and Quadratic Functions

$G = (V, E) \Rightarrow f_G = \vee_{uv \in E} x_u x_v \Rightarrow f_G(S) = 0$ iff S indep. set

F computes $f_G \Rightarrow F$ represents G. But $\not\Leftarrow$ in general !

$G = (V, E)$ saturated $\Rightarrow$



Lemma: F monotone circuit, G saturated graph $\Rightarrow$

F represents G iff F computes f_G

Proof: F monotone and represents G

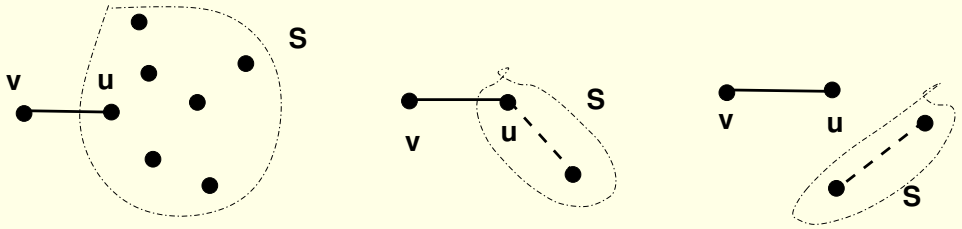
$f_G(S) = 1 \Rightarrow S$ contains an edge $uv \Rightarrow F(\{uv\}) = 1 \Rightarrow F(S) = 1$

$f_G(S) = 0 \Rightarrow S$ independent set $\Rightarrow S =$ vertex or non-edge $\Rightarrow F(S) = 0$ $\square$

Need: Lower bounds for saturated graphs

Attack on formulas $\Rightarrow$ let's play a game!

- Alice gets an edge $e \in E$
- Bob gets an independent set $S \subseteq V$
- Determine a vertex $v \in e \setminus S$ (Bob must also know v!)



$c(G)$ = comm. complexity (deterministic, two-way)

Bob must know the answer $\Rightarrow$ $c(G) \geqslant \log_2 n$

Alice can send here edge $\Rightarrow$ $c(G) \leqslant \log |E| \leqslant 2 \log n$

Why interesting?

Formula = circuit with all gates of fanout 1

$L(f_m)$ = minimal formula over $\{\wedge, \vee, \neg\}$

Best known: $L(f_m) \geqslant m^3$ [Khrapchenko/Andreev/Hastad]

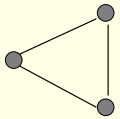
$c_2(G)$ = when Bob gets a non-edge $\Rightarrow$ $|S| = 2$

Lemma:
f_m with $m = \log n$ char. funct. of G $\Rightarrow$ $L(f_m) \geqslant 2^{c_2(G)-m}$

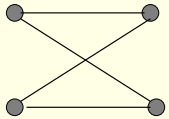
Proof: Magnification lemma + Karchmer–Wigderson

$c_2(G) \geqslant \log n + k \cdot \log\log n$ $\Rightarrow$ $L(f_m) \geqslant (\log n)^k = m^k$!

Games with large independent sets



triangle



C_4 oder $K_{2,2}$

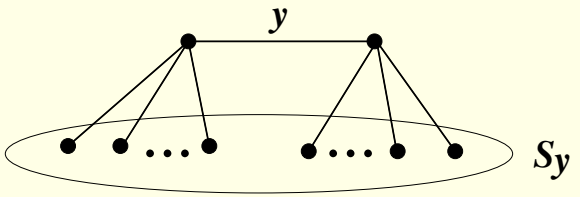
Theorem:
No triangles and no 4-cycles in $G = (V, E)$ $\Rightarrow$ $c(G) \geqslant \log |E| - 1$

Quadratic function of $G = (V, E)$ $\Rightarrow$ $f_G = \bigvee_{uv \in E} x_u x_v$ $\Rightarrow$ $L_+(f_G) \leqslant |E|$

Corollary:
No triangles and no 4-cycles in $G = (V, E)$ $\Rightarrow$ $L_+(f_G) \geqslant |E|/2$

Plane graph $G = (V, E)$ $\Rightarrow$ $|E| = \Theta(n^{3/2})$ $\Rightarrow$ $L_+(f_G) = \Omega(n^{3/2})$

Proof



$y \in E \;\Rightarrow\; S_y =$ set of proper neighbors of y

C_4-free $\;\Rightarrow\;$ sets $S_y \subseteq V$ are independent sets

Protocol $P(x, S_y)$ outputs some vertex $v \in x \setminus S_y$

New protocol:

$$P'(x, S_y) = \begin{cases} 1 & \text{if } v \in y \\ 0 & \text{if } v \notin y \end{cases}$$

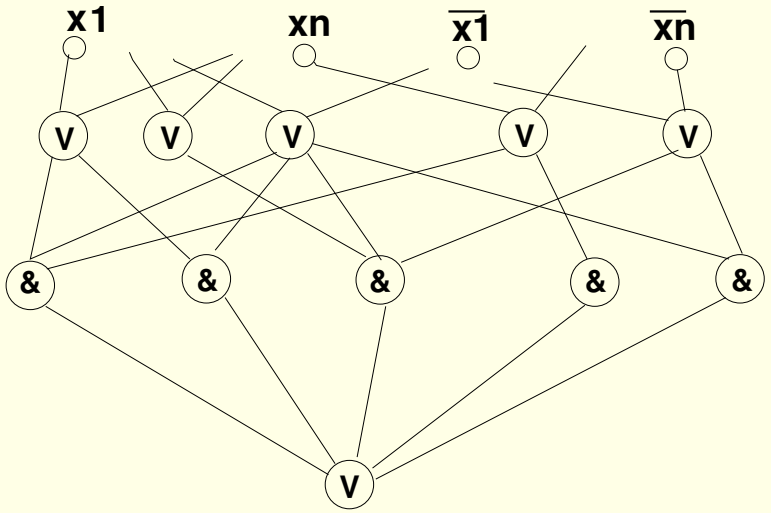
Observation: $P'(x, S_y) = 1 \iff x \cap y \neq \emptyset$

Why? $x = uv$ and $v \notin y \;\Rightarrow\; u \notin y$ Why? Because otherwise $v \in S_y$

Comm. matrix of $P' =$ intersection matrix $\;\Rightarrow\;$ has full rank $= |E|$

$\Rightarrow$ comm. complexity of P' is $\geqslant \log |E| \;\Rightarrow$ comm. compl. P is $\geqslant \log |E| - 1$ $\quad \square$

Σ_3 circuits



= just an OR of CNFs ... so simple! ... why interesting?

Σ_3 and log-depth circuits

Theorem [Valiant 1977]:
Lower bound $2^{\alpha m/\log\log m}$ for Σ_3 circuits
$\Rightarrow$ super-linear lower bound αm for log-depth circuits.

A lot of progress, but ... known only $\Sigma_3(f_m) \geqslant 2^{\Omega(\sqrt{m})}$ $\Rightarrow$ too weak!

Need $n \times n$ graphs G with $\Sigma_3(G) \geqslant n^{\epsilon}$, $\epsilon = \omega(\frac{\alpha}{\log\log\log n})$ $(n = 2^m)$

Conjecture [Pudlák, Rödl, Savický 1988]:
If G is C_4-free then $\Sigma_3(G) = \Omega(|E|/n)$

Explicit graphs with $|E| = \Omega(n^{3/2})$ exist (projective planes)

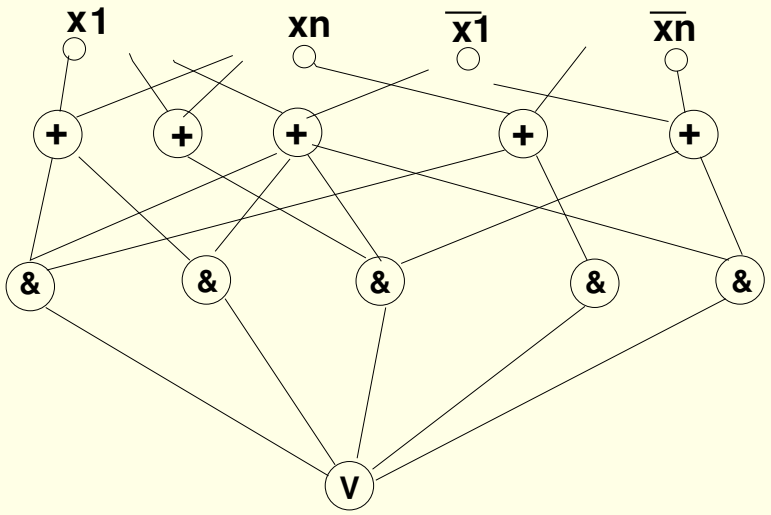
$\Rightarrow$ would yield $\Sigma_3(G) \geqslant n^{\epsilon}$ with $\epsilon = 1/2$

$\Rightarrow$ lower bound $\Omega(m \log\log m)$ for log-depth circuits!

But .... the conjecture remains open!

Conjecture is true for $\Sigma_3^{\oplus}$ circuits

$\Sigma_3^{\oplus}$ circuit = Σ_3 circuit with $\oplus$-gates on the bottom



Computes a union of affine spaces

Conjecture is true for $\Sigma_3^{\oplus}$ circuits (cntd.)

$\omega(G) =$ max k s.t. G contains complete $k \times k$ subgraph

Theorem: For every $n \times n$ graph $G \subseteq U \times V$

$$\Sigma_3^{\oplus}(G) \geqslant \frac{|G|}{2n \cdot \omega(G)}$$

Plane graph $G \quad \Rightarrow \quad |G| = \Theta(n^{3/2})$ and $\omega(G) = 1$

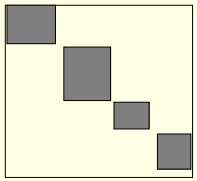
$\Rightarrow \quad \Sigma_3^{\oplus}(\mathsf{f}_m) \geqslant 2^{m/2}$ for plane function f_m

Hadamard graph $G \quad \Rightarrow \quad |G| = \Theta(n^2)$ and $\omega(G) = O(\sqrt{n})$

$\Rightarrow \quad \Sigma_3^{\oplus}(IP_m) \geqslant 2^{m/2}$ for inner product function

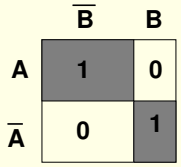
$$IP_m(x_1, \ldots, x_m, y_1, \ldots, y_m) = \sum_{i=1}^{m} x_i y_i \,(\mathrm{mod}\, 2)$$

Proof



Fat matching = collection of vertex-disjoint cliques

$\mathrm{fat}(G) := \min\{t \mid G \text{ is a union of } t \text{ fat matchings}\}$



$h(X) = \bigoplus_{v \in A \cup B} x_v$ represents a fat matching

Observation: Intersection of fat matchings is a fat matching !

$\Rightarrow \; g(X) = \bigwedge_{i=1}^{t} \bigoplus_{v \in S_i} x_v$ represents a fat matching $\Rightarrow \; \Sigma_3^{\oplus}(G) \geqslant \mathrm{fat}(G)$

$H = \cup_{i=1}^{t} A_i \times B_i \subseteq G$ and $I = \{i : |A_i| \leqslant k\} \; \Rightarrow$

$|H| = \Sigma_{i=1}^{t} |A_i| \cdot |B_i| \leqslant \Sigma_{i \in I} |B_i| \cdot k + \Sigma_{i \notin I} |A_i| \cdot k \leqslant 2nk$

$\Rightarrow \; \mathrm{fat}(G) \geqslant |G|/(2nk)$ □

$\Sigma_3^{\oplus}$ with arbitrary threshold gates on the top

Threshold covering number $\mathrm{thr}_{\mathcal{H}}(G) = \min t$ s.t. $\exists k \geqslant 0$ and $\exists H_1, \ldots, H_t \in \mathcal{H}$ s.t.

$$uv \in G \iff uv \text{ belongs to } \geqslant k \text{ of } H_i\text{'s}$$

Discriminator Lemma [Hajnal/Maass/Pudlák/Szegedy/Turán 1993]:
If

$$\left| \frac{|G \cap H|}{|G|} - \frac{|\overline{G} \cap H|}{|\overline{G}|} \right| \leqslant \frac{1}{M} \quad \text{for every } H \in \mathcal{H}$$

then

$$\mathrm{thr}_{\mathcal{H}}(G) \geqslant M$$

Theorem:
Any $\Sigma_3^{\oplus}$ circuit which has an arbitrary threshold gate on the top and represents an $n \times n$ Hadamard graph must have top fanin $\Omega(\sqrt{n})$.

Proof

$H \in \mathcal{H} = \{\text{all fat matchings}\} \;\Rightarrow\; H = \cup_{i=1}^{t} A_i \times B_i \;\Rightarrow$

$$\begin{aligned}
\Big||G \cap H| - |\overline{G} \cap H|\Big| &= \sum_{i=1}^{t} \Big||G \cap (A_i \times B_i)| - |\overline{G} \cap (A_i \times B_i)|\Big| \\
&\leqslant \sum_{i=1}^{t} \sqrt{a_i b_i n} \quad \text{(Lindsey's lemma)} \\
&\leqslant \sqrt{n} \sum_{i=1}^{t} \frac{a_i + b_i}{2} \quad \text{(arithm./geom. means)} \\
&\leqslant n^{3/2}.
\end{aligned}$$

G is Hadamard graph $\;\Rightarrow\;$ both $|G|$ and $|\overline{G}|$ are $\Theta(n^2)$

$$\Rightarrow \left|\frac{|G \cap H|}{|G|} - \frac{|\overline{G} \cap H|}{|\overline{G}|}\right| = O\left(\frac{1}{\sqrt{n}}\right) \;\Rightarrow\; \mathrm{thr}_{\mathcal{H}}(G) = \Omega(\sqrt{n})$$

How to crack Σ_3?

Σ_3 = OR of CNFs $g(X) = \left(\vee_{u \in S_1} x_u\right) \wedge \cdots \wedge \left(\vee_{u \in S_t} x_u\right)$

Clique Covering $\mathrm{cc}(G) = \min t$: $E = A_1 \times B_1 \cup \cdots \cup A_t \times B_t$

Intersection Number $\mathrm{int}(G) = \min t$: $\exists V \ni u \mapsto A_u \subseteq \{1, \ldots, t\}$ s.t.

$$uv \in E \iff A_u \cap A_v = \emptyset$$

Lemma: $\mathrm{cnf}(G) \overset{(1)}{=} \mathrm{cc}(\overline{G}) \overset{(2)}{=} \mathrm{int}(G)$

Proof: (1) $\overline{S_1}, \ldots, \overline{S_t}$ indep. sets in $G \Rightarrow$ cliques in $\overline{G}$

(2): $I_1, \ldots, I_t$ indep. sets in $G \Rightarrow$ take $A_u = \{i : u \in I_i\}$

$uv \in E \iff \neg\exists j$: $\{u, v\} \subseteq I_j \iff A_u \cap A_v = \emptyset$ $\square$

Upper bounds

Lemma: $\Sigma_3(G) \leqslant \min\{\mathrm{cnf}(G), \mathrm{cnf}(\overline{G})\}$

Proof: $t = \mathrm{cnf}(\overline{G}) = \mathrm{cc}\,(G)$

$\Rightarrow$ $E = \cup_{i=1}^{t} A_i \times B_i$ $\Rightarrow$ OR of t CNFs $\left(\vee_{u \in A_i} x_u\right) \wedge \left(\vee_{v \in B_i} x_v\right)$

$\Rightarrow$ $\Sigma_3(G) \leqslant t$. □

[Alon 1986]: $d =$ max degree of G $\Rightarrow$ $\mathrm{cnf}(G) = \mathrm{cc}\,(\overline{G}) = O(d^2 \log n)$

Proof: Probabilistic argument □

Matching M_n $\Rightarrow$ $\mathrm{cnf}(M_n) = O(\log n)$ $\Rightarrow$ $\Sigma_3(\overline{M}_n) = O(\log n)$

Lower bounds?

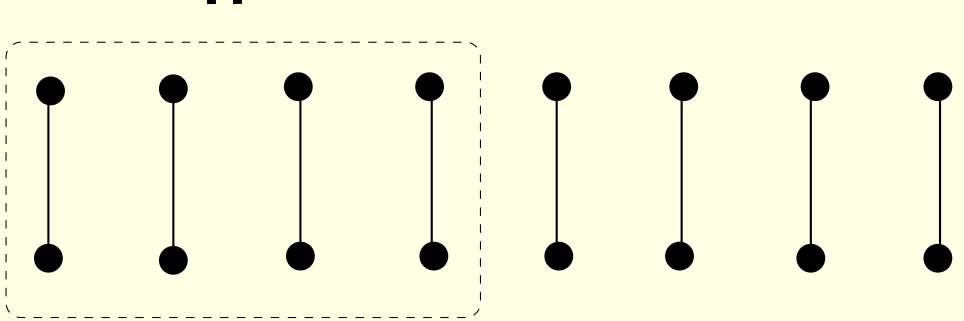
Best known: $\Sigma_3(H_n) \geqslant (\log n)^{3/2}$ for Hadamard graph [Lokam 2003]

$\Rightarrow$ just a bit more than trivial !

Matching $M_n \quad \Rightarrow \quad \Sigma_3(M_n) = \Omega(\log n)$

Proof: $t = \Sigma_3(M_n) \quad \Rightarrow$ $\exists$ matching with $|H| \geqslant n/t$ edges and

$\mathrm{cnf}(H) = \mathrm{int}(H) \leqslant t$



$\Rightarrow \quad \exists u \mapsto A_u \subseteq \{1, \ldots, t\}$ s.t. $A_{u_i} \cap A_{v_j} = \emptyset \iff i = j$

$\Rightarrow$ sets $A_{u_1}, \ldots, A_{u_k}$ distinct $\quad \Rightarrow \quad 2^t \geqslant k \geqslant n/t \quad \Rightarrow \quad t = \Omega(\log n)$ $\square$

For almost all $n \times n$ graphs $G \quad \Rightarrow \quad \Sigma_3(G) = \Omega(\sqrt{n})$

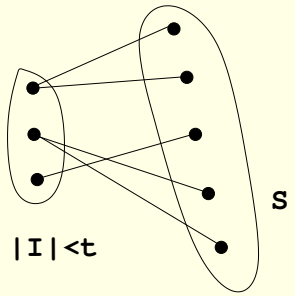
Why difficult? An easy case

$G = (V, E)$ $\Rightarrow$ quadratic function $f_G = \vee_{uv \in E} x_u x_v$

Theorem: G star-free $\Rightarrow$ $\mathrm{cnf}(G) \geqslant |E|/d^2$
$\Rightarrow$ $\Sigma_3(G) \geqslant \sqrt{|E|}/d$

Proof: CNF F of length $t = \mathrm{cnf}(G)$ computes f_G
$\Rightarrow$ take $F' = F \setminus \{C\}$ where $C = \vee_{u \in S} x_u$ $\Rightarrow$ $F = F' \wedge \left(\vee_{u \in S} x_u\right)$
F accepts all edges $e \in E$ $\Rightarrow$ $\forall e \in E$: $e \cap S \neq \emptyset$ $\Rightarrow$ $|S| \geqslant |E|/d$

F' must make an error $\Rightarrow$ $F'(I) = 1$ for indep. set I of G, $|I| \leqslant t$



C must correct the error $\Rightarrow$ `For all u in S exists v in I with uv in E`

$\Rightarrow$ $\exists v \in I$: $\deg(v) \geqslant |S|/|I| \geqslant |E|/td$ $\Rightarrow$ $t \geqslant |E|/d^2$ $\square$

Expander graphs?

Plane graph $G = (U, W, E)$ $\Rightarrow$ $\forall X \subseteq U$: $|N(X)| \geqslant n - \frac{n^{3/2}}{|X|}$ $\Rightarrow$ good expander!

Adversary must cover all non-edges by few indep. sets

$S = X \times Y$ indep. set $\Rightarrow$ $Y \cap N(X) = \emptyset$ $\Rightarrow$ $|S| \leqslant |X|(n - |N(X)|) \leqslant n^{3/2}$

$\Rightarrow$ need $\frac{n^2 - |E|}{n^{3/2}} \geqslant \sqrt{n}$ indep. sets!

But ... adversary is allowed to remove a $1/t$ fraction of edges

$\Rightarrow$ expanding property may be destroyed ! Why?

Remove $\frac{n}{C} \times \frac{n}{C}$ clique $S = X \times Y$

$\Rightarrow$ removed only constant fraction $|S \cap E| \leqslant (n/C)^{3/2}$ edges (due to C_4-freeness)

But ... very large indep. set $|S \cap \overline{E}| \geqslant \left(\frac{n}{C}\right)^2 = \Omega(n^2)$!

The Single-Level Conjecture

Single level circuit $\Rightarrow$ only one level of AND gates

Single level conjecture:
$L_1(f_G)/L(f_G) \leqslant const$ for every G

[Krichevski 1964] $\Rightarrow$ holds for $G = K_n$

[Mirwald–Schnorr 1987] $\Rightarrow$ holds in basis $\{\oplus, \wedge, 0, 1\}$ for $f_G^{\oplus} = \oplus_{uv \in E} x_u x_v$

[Bublitz, Lenz–Wegener 1991] $\Rightarrow$ examples with $L_1(f_G) = L(f_G) + 1$

[Lenz–Wegener 1991] $\Rightarrow$ What about $\{\vee, \wedge, 0, 1\}$?

[S.J.] $\Rightarrow$ For formulas conjecture is even not near to the truth !

Theorem
$\exists$ n-vertex graphs such that $L_1(f_G)/L(f_G) = \Omega(n^{\epsilon})$, $\epsilon > 0$

Disproof of the Conjecture

$\mathrm{cc_b}(G) = \min \Sigma_{i=1}^{t}(|A_i| + |B_i|)$ over all covers $E = \cup_{i=1}^{t} A_i \times B_i$

single level $\Rightarrow$ Σ_3 with ANDs of fanin 2 $\Rightarrow$ $L_1(f_G) = \mathrm{cc_b}(G)$

Theorem: $\mathrm{cc_b}(G) \geqslant \alpha \cdot |E|$ where $\alpha = \min\left\{\frac{a+b}{ab} : \ G \text{ contains } K_{a,b}\right\}$

$H \subseteq U \times W$ Kneser graph with $U = W = 2^{[r]}$ and $uv \in H$ iff $u \cap v = \emptyset$

$\Rightarrow$ no $K_{a,b}$ with $\frac{a+b}{ab} < \alpha = 2^{-r/2}$

$\Rightarrow$ $|H| = \Sigma_{u \in U}\, d(u) = \Sigma_{i=0}^{r} \binom{r}{i} 2^{r-i} = 3^r$

Take G = saturated version of H

$\Rightarrow$ $L_1(G) = \mathrm{cc_b}(G) \geqslant \mathrm{cc_b}(H) \ \geqslant \alpha \cdot |E| \geqslant 3^r/2^{r/2} \geqslant n^{1+c}$, $c > 0$

H can be represented by CNF F of length $|F| = \mathrm{int}(H) = r = \log n$

$\Rightarrow$ $F' = F \vee T_2^U \vee T_2^W$ computes f_G $\Rightarrow$ $L(f_G) = O(n \log n)$

$\Rightarrow$ $L_1(f_G)/L(f_G) = \Omega(n^{1+c}/n \log n) = \Omega(n^{\epsilon})$ $\square$

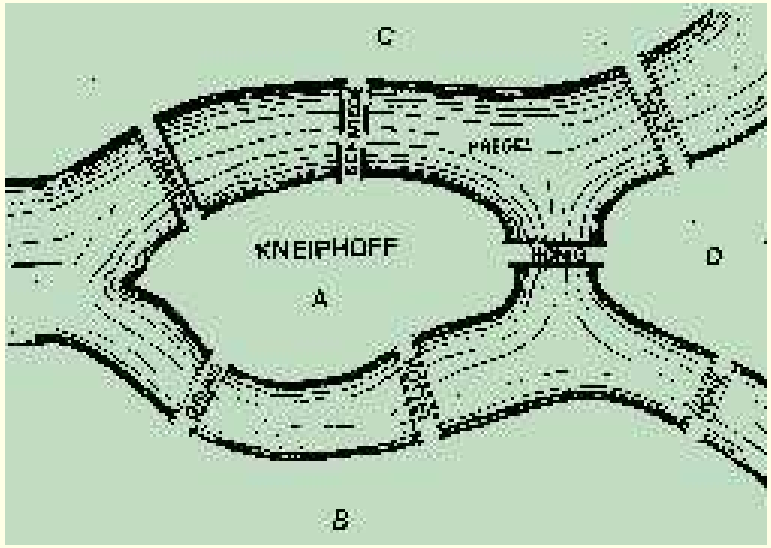
Open problems

1. Disprove the SL Conjecture for circuits $\Rightarrow$ not too hard $\Rightarrow$ diploma thesis!
2. Find graphs G with $\Sigma_3(G) \gg \Sigma_3(\overline{G})$ $\Rightarrow$ harder $\Rightarrow$ PhD thesis
   $\Rightarrow$ separation of comm. compl. classes $\Sigma_2^{cc} \neq \Pi_2^{cc}$
   $\Rightarrow$ a 20 years old problem!
4. Prove $\Sigma_3(G) \geqslant n^{\alpha}$ $\Rightarrow$ hard $\Rightarrow$ breakthrough !
   - $\alpha = \omega(1/\sqrt{\log n})$ $\Rightarrow$ highest LB for Σ_3 circuits $2^{\omega(\sqrt{m})}$
   - $\alpha = \omega(1/\log\log\log n)$ $\Rightarrow$ Super-linear LB for Log-depth circuits
     $\Rightarrow$ 30+ years old problem!
5. Prove $c_2(G) \geqslant \log n + k \cdot \log\log n$ for edge/non-edge game $\Rightarrow$ hard !
   $\Rightarrow$ $L_{\{\wedge,\vee,\neg\}}(f_m) \geqslant m^k$ $\Rightarrow$ beat the best LB $L_{\{\wedge,\vee,\neg\}}(f_m) \geqslant m^3$
6. Improve Razborov $\Rightarrow$ Monotone LBs also when maxterms are long!
   - Razborov needs: both DNF and CNF "dispersed"
   - Find arguments based only on properties of DNFs
   - Show that DNF $f_G = \vee_{uv \in E} x_u x_v$ cannot be compressed s

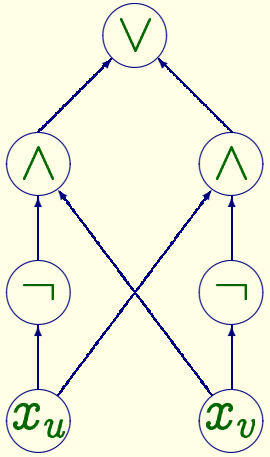
Conclusion

- Need to consider only monotone circuits $\Rightarrow$ a hope !
- Exist graphs with very special properties $\Rightarrow$ unnatural proofs $\Rightarrow$ a hope !
- Something can be already done:
  - $\Rightarrow$ high LBs for $\Sigma_3^{\oplus}$ $\Rightarrow$ $\Sigma_3^{\oplus}(IP_m) \geqslant 2^{m/2}$
  - $\Rightarrow$ even with threshold gates on the top
  - $\Rightarrow$ first high LBs for quadratic functions
  - $\Rightarrow$ disproof of the Single Level Conjecture for formulas
- Do this for Σ_3 $\Rightarrow$ super-linear LB for log-depth circuits !

A bridge between Computational Complexity and Graph Theory



$\Rightarrow$



Problems for circuits $\Rightarrow$ purely graph-theoretic problems

Graph theory is 250+ years old but very rich lady $\Rightarrow$ a light at the end of tunnel?