RESOLUTION AND TSEITIN FORMULAS

LAST LECTURE

Exponential lower bounds on resolution refutations of pigeonhole principle (PHP) formulas

$$\boxed{L_{\mathcal{R}}\left(PHP_n^{n+1} \vdash \bot\right) = \exp(\Omega(n))}$$

PHP_n^{n+1} has size $N = \Theta(n^3)$, so the lower bounds are $\exp(\Omega(\sqrt[3]{N}))$ in terms of formula size.

Proof idea (high-level)

Use Prosecutor–Defendant game.

Good strategies for Defendant $\Rightarrow$ resolution lower bounds

Defendant picks random matching of $n/4$ pigeons to $n/4$ pigeonholes. Exponentially many different choices. Before Prosecutor wins, has to write down noticeable fraction of (information about) this random matching

$\Rightarrow$ Prosecutor needs exponentially many records

$\Rightarrow$ exponential resolution length/size lower bound

TODAY

Study formulas encoding (contradictions of) the **HANDSHAKING LEMMA**: "Sum of vertex degrees in undirected graph is an even number"

Encoded in somewhat sneaky way as **TSEITIN FORMULAS**. We will prove truly exponential lower bounds $\exp(\Omega(N))$ in terms of formula size N for Tseitin formulas.

Prosecutor – Defendant Game [Pudlák '00]

Unsatisfiable CNF formula F

Prosecutor maintains record R = partial truth value assignment to $Vars(F)$

Every record R has an instruction of type (a) or (b)

(a) ask about x

Defendant answers $b \in \{0,1\}$; prosecutor adds $x = b$ to R

(b) Forget values, i.e., shrink assignment to $R' \subsetneq R$

A <u>winning position</u> R for prosecutor is an assignment falsifying some $C \in F$

A <u>(complete) strategy</u> $\mathcal{S}$ for prosecutor is a (finite) set of record such that

- there is a record for every possible defendant response
- regardless of how defendant plays, prosecutor always wins in the end.

The size of a strategy $\mathcal{S}$ is the number of records in it.

LEMMA [Pudlák '00]

F has a resolution refutation in length L iff prosecutor has a strategy of size $\Theta(L)$

<u>Proof</u> ($\Rightarrow$) We did this last lecture

($\Leftarrow$) Exercise (not super hard, but not trivial either)

Use notation x^b for literals, $b \in \{0,1\}$

$\boxed{x^1 = x}$ $\boxed{x^0 = \overline{x}}$ (literal x^b satisfied by $x=b$)

TSEITIN FORMULA

$G = (V, E)$ undirected graph over $|V| = n$ vertices

$\chi : V \to \{0,1\}$ CHARGE FUNCTION

Identify every edge $e \in E$ with variable x_e (or sometimes overload notation and write e for variable also)

$PARITY_{v,\chi}$ = set of clauses encoding that # true edges incident to v is equal to $\chi(v)$ (mod 2) – i.e., the parity of the sum is $\chi(v)$

$$PARITY_{v,\chi} = \text{"} \sum_{e \ni v} x_e \equiv \chi(v) \pmod 2 \text{"}$$

Formally, this is set of clauses

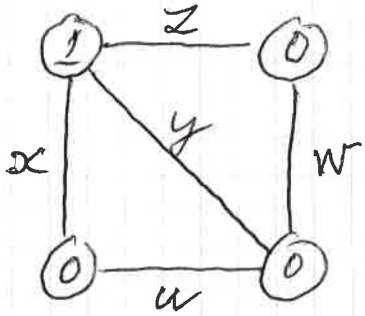
$$\boxed{PARITY_{v,\chi} = \left\{ \bigvee_{e \ni v} x_e^{1-b_e} \;\middle|\; \sum_{e \ni v} b_e \not\equiv \chi(v) \pmod 2 \right\}}$$

For every assignment to $\{x_e \mid e \ni v\}$ that gets parity wrong, i.e. $\sum b_e \not\equiv \chi(v) \pmod 2$, at least one literal $x_e^{1-b_e}$ has to be satisfied, showing that we do not have this assignment

The TSEITIN FORMULA over G with respect to χ is

$$\boxed{Ts(G, \chi) \doteq \bigwedge_{v \in V} PARITY_{v,\chi}}$$

EXAMPLE



$$
\begin{aligned}
&(x \lor y \lor z) && \land (u \lor w \lor \overline{y}) \\
\land\ &(x \lor \overline{y} \lor \overline{z}) && \land (u \lor \overline{w} \lor y) \\
\land\ &(\overline{x} \lor y \lor \overline{z}) && \land (\overline{u} \lor w \lor y) \\
\land\ &(\overline{x} \lor \overline{y} \lor z) && \land (\overline{u} \lor \overline{w} \lor \overline{y}) \\
\land\ &(x \lor \overline{u}) && \land (w \lor \overline{z}) \\
\land\ &(\overline{x} \lor u) && \land (\overline{w} \lor z)
\end{aligned}
$$

If G has bounded degree d (i.e., all vertices have $\leq d$ incident edges), then $Ts(G, \chi)$ is

- d-CNF formula
- $\leq d|V|/2$ variables
- $\leq 2^{d-1}|V|$ clauses

When talking about Tseitin formulas over G, assume bounded degree $\boxed{d = O(1)}$ unless stated otherwise

$\chi : V \to \{0,1\}$ is an ODD-CHARGE function if $\sum_{v \in V} \chi(v) \equiv 1 \pmod 2$

PROPOSITION 1

If G is connected, then $Ts(G, \chi)$ is unsatisfiable iff χ is an odd-charge function

Proof sketch If χ odd-charge, then sum-up parity constraints over all $v \in V$ should be odd number, but is equal to $\sum_{e \in E} 2 \cdot x_e$ since every edge counted exactly twice

If χ even-charge, start with assignment setting all edges false. Pick any two violated vertices. Flip all edges on path between them. Only changes parity of these two vertices. #violations decrease by 2. Repeat ☐

Even # of violations

In what follows, always assume G connected unless stated otherwise.

Today we want to prove that if G is a well-connected graph, then resolution refutations of $TS(G, \chi)$ require exponential length (truly exponential $\exp(\Omega(N))$ in size N of formula)

Two comments:

- For even-charge χ this is obviously true (satisfiable formulas are very hard to refute)
- For odd-charge χ we only really care about the charge. All odd-charge functions on V are equivalent from the point of view of resolution. (Won't need this; it is just a side note.)

Well-connected graphs are known as expanders. Several different ways of measuring expansion

(1) Vertex expansion
Every small-to-medium vertex set $U \subseteq V(G)$ has many neighbours in $N(U) \setminus U$.

(2) Edge expansion
Every small-to-medium-large $U \subseteq V$ has many outgoing edges to $V \setminus U$.

(3) Algebraic expansion
The gap between the two largest eigenvalues in the (normalized) adjacency matrix is large.

These notions are all tightly connected (but proving this is out of scope for this course). We will use edge expansion

For $G = (V, E)$ and $U \subseteq V$, let ∂U denote the set of outgoing edges from U, i.e.,

$$\partial U = \{ (u, v) \in E \mid u \in U, v \in V \setminus U \}.$$

DEFINITION 2 (Edge expansion)

An undirected graph $G = (V, E)$ G has bounded degree d and is a (d, δ)-edge expander if for every vertex set $U \subseteq V$ of size $|U| \leq |V|/2$ it holds that $|\partial U| \geq \delta |U|$

That is, a constant fraction of edges incident to U are exiting U.

Now we can state the goal of today's lecture

THEOREM 3 [Urquhart '87]

Fix $d \in \mathbb{N}^+$, $\delta > 0$, and suppose that $\{G_n\}_{n=3}^{\infty}$ is a family of n-vertex, (d, δ)-edge expanders. Then for any family of odd-charge functions $\chi_n : V(G_n) \to \{0, 1\}$ it holds that the CNF formula family $\{Ts(G_n, \chi_n)\}_{n=3}^{\infty}$ requires resolution refutations of length $\exp(\Omega(n))$.

Remarks: (i) Formula size is $\Theta(n)$, so truly exponential

(ii) Concrete constants will depend on d, δ.

Will not do Urquhart's proof, but our own Prosecutor-Defendant-style proof (cooked up together with Massimo Lauria & Per Austrin, but any errors are the responsibility of the lecturer, of course).

We will prove a slightly weaker result requiring expansion $\delta > 1$. This condition can be removed with some work, but we keep it for simplicity

But first: Is this a non-vacuous theorem? I.e., are there such edge expanders? Yes, picking a random d-regular graph will do.

THEOREM 4 [Bollobás '88]
For any fixed $d \in \mathbb{N}^+$, $d \geq 3$, there exists a universal constant δ such that asymptotically almost surely a random d-regular graph is a (d, δ)-expander.

A family of events $\{E_n\}_{n=1}^{\infty}$ happens ASYMPTOTICALLY ALMOST SURELY (a.a.s.) if $\lim_{n \to \infty} \Pr[E_n] = 1$

Sometimes also referred to as "with high probability" (whp), but we will try to stick to the more precise terminology a.a.s. in this course

In fact, it is possible to give explicit constructions of families of expander graphs, but the analysis of such constructions is typically highly nontrivial.

MORE ABOUT EDGE EXPANSION

VII 1/2

The maximal edge expansion of a graph is known as the ISOPERIMETRIC NUMBER $h(G)$ (or the CHEEGER CONSTANT). i.e.,

$$h(G) = \min_{U \subseteq V, |U| \leq |V|/2} \frac{|\partial U|}{|U|}$$

For a random graph and any ~~partition~~ subset $U \subseteq V(G)$ of size $|U| \leq |V(G)|/2$, we would expect roughly half of the edges in U to go to $V \setminus U$. So a random d-regular graph might have ~~boundary~~ edge expansion something like $d/2$ if we are lucky. This is indeed the case for d large enough.

THEOREM 4' [Bollobás '88]
For every $\varepsilon > 0$ there is a $d \in \mathbb{N}^+$ such that a random d-regular n-vertex graph has edge expansion at least $\frac{d}{2} - \varepsilon$ asymptotically almost surely as $n \to \infty$.

Bollobás actually calculates a more precise expression from which it follows that random 6-regular graphs have edge expansion at least > 1 a.a.s and random 4-regular graphs have expansion 0.4 a.a.s.

In the proof, it will be convenient to use another (fourth) notion of expansion

DEFINITION 5 (CONNECTIVITY EXPANSION)
An undirected graph $G=(V,E)$ is a $\boxed{(d,c)\text{-CONNECTIVITY EXPANDER}}$ if G has bounded degree d and for every edge set $E' \subseteq E$, $|E'| \leq c \cdot |V|$ it holds that the graph $G' = (V, E \setminus E')$ has a connected component of size strictly larger than $|V|/2$

PROPOSITION 6
Every (d,δ)-edge expander is a (d,c)-connectivity expander for $c = \delta/4$

Proof. Suppose not. Fix counter-example E' of size $|E'| = cn$. Then we can find $V^* \subseteq V$, $n/4 < |V^*| \leq n/2$ (for $|V| = n$) such that there is no edge from V^* to $V \setminus V^*$ in G' (Just look at connected components and take unions until we reach a set V^* of this size.)
But since G is an edge expander, there are $\geq \delta |V^*| > \frac{\delta n}{4} = c \cdot n = |E'|$ edges from V^* to $V \setminus V^*$, which are more edges than E' removed. Contradiction $\boxtimes$

To save typing in what follows, say that any edge set $E' \subseteq E$ of size $|E'| \leq c \cdot n$ is of $\boxed{\text{MODERATE SIZE}}$ and that the (unique) connected component of size $> n/2 = |V|/2$ is the $\boxed{\text{LARGE COMPONENT}}$.

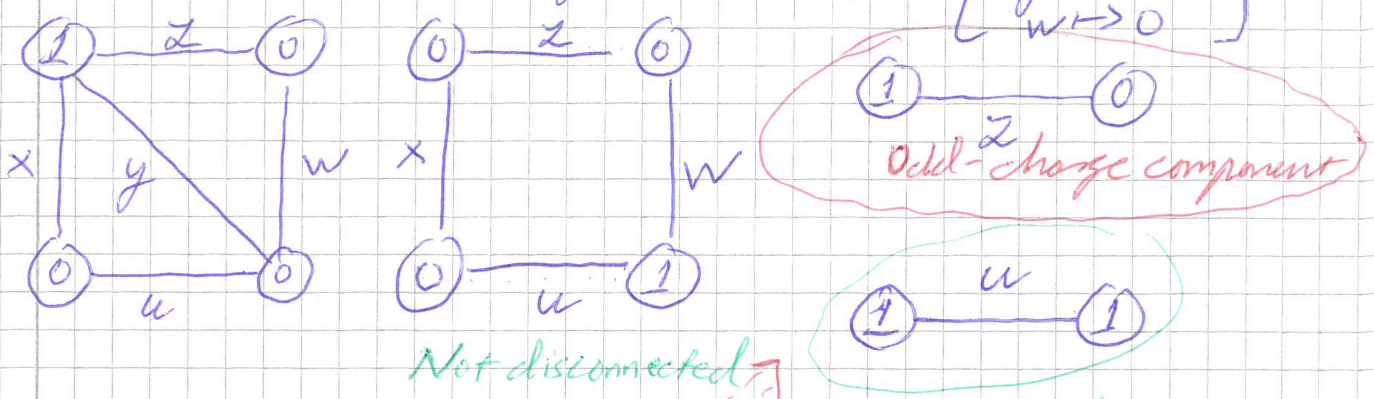
For any moderate-size E', an assignment $\rho : E' \to \{0,1\}$ (identifying edges and variables) is $\boxed{\text{CHARGE-PRESERVING}}$ if in $G' = (V, E \setminus E')$ the large component has odd charge and any small component has even charge

(where the charge function χ is updated for G' by plugging in values from ρ)

Observe: If E' is a set of edges that disconnects G, then where the odd charge ends up depends only on the parity of the cut $E_{cut} \subseteq E'$

EXAMPLE

$\rho_2 = \{y \mapsto 1\}$ $\qquad$ $\rho_2 = \left\{ \begin{array}{l} x \mapsto 1, \\ y \mapsto 1, \\ w \mapsto 0 \end{array} \right\}$ $\qquad$ IX



OBSERVATION Where the odd charge ends up only depends on the parity of the cut.

DEFENDANT STRATEGY (a bit high level)

$G = (V, E)$ is a (d, δ)-edge expander and (d, c)-connectivity expander for $c = \delta/4$.

Pick uniformly at random a set $E' \subseteq E$ of size $cn/2$.
Sample uniformly at random a charge-preserving assignment to E'.

NB! Different edges are not set independently.
E.g. above, if $x, u \in E'$, then once we have randomly set x we must set $u = x$.
But many edges in E' will be set independently at random (need to argue this carefully).

Call this assignment ρ_{init}.
Defendant will maintain ρ such that

a) $\rho \supseteq \rho_{init}$

b) ρ consistent with $\mathcal{R}$ (Prosecutor record)

in fact, will have $\boxed{\rho = \rho_{init} \cup \mathcal{R}}$

When Prosecutor asks about value of x_e (= edge e) X

(i) If $e \in Dom(\rho)$, answer according to $\rho(x_e)$
[In this case $e \in Dom(\rho_{init})$, otherwise Prosecutor knows the answer and wouldn't ask]

(ii) If $e \notin Dom(\rho)$, answer with value $b \in \{0,1\}$ so that the odd charge in $G' = (V, E \setminus (Dom(\rho) \cup \{e\}))$ stays in the large component.

If this is not possible (since the large component disappeared) give up, or answer arbitrarily or whatever.

When Prosecutor forgets to get $R' \subsetneq R$

Simply update ρ to $\rho = \rho_{init} \cup R'$

Now we want to implement the same lower bound strategy as for PHP

1. Before Prosecutor wins, has to have "informative record" containing lots of edges
2. Such records contain lots of information about Defendant's initial random choice
3. Hence, any given informative record is exponentially unlikely to be consistent with a particular random choice
4. So strategy contains exponentially many records.

OBSERVATION 7
Before Prosecutor wins there must be a record with more than $cn/2$ edges

Proof A winning position for Prosecutor is a falsified vertex constraint = an odd-charge disconnected component of size 1.

Let E' initial random choice by Defendant. Let E'' be edges in Prosecutor record. As long as $|E' \cup E''| \le cn$ Defendant is making sure the odd charge is in the large component. Hence, before Prosecutor wins we must have $|E' \cup E''| = |E''| + |E' \setminus E''| \ge cn$ or $|E''| \ge cn - |E' \setminus E''| \ge cn/2$. ∎

Call a record with $\ge cn/2$ edges informative
We want to prove that a fixed informative record R has exponentially small probability of being consistent with Defendant's initial random choice E'

For any edge $e \in R$ $\Pr_{g_{init}}[e \in E'] = \dfrac{cn/2}{\#\text{edges } |E(G)|} \ge$

$$\ge \frac{cn/2}{dn/2} = \frac{c}{d}$$

By linearity of expectation

$$\mathbb{E}_{g_{init}}\left[|\mathrm{Dom}(R) \cap E'|\right] \ge \frac{cn}{2} \cdot \frac{c}{d} = \frac{c^2}{2d} n$$

By concentration of measure (same calculations (kind of) as for PHP) it holds that

$$\Pr_{\rho_{init}}\left[\, |Dom(R) \cap E'| \le \frac{c^2}{4d} n \,\right] \le 2^{-\varepsilon' n} \qquad (\dagger)$$

for some $\varepsilon' > 0$ (for n large enough).

Fix $E_1 = Dom(R) \cap E'$ and assume for now that

$$|E_1| \ge \frac{c^2}{4d} n$$

By construction also have

$$|E_1| \le |E'| = cn/2$$

RECALL PROPOSITION 6

LEMMA 8 (Key technical lemma)

Suppose that G is a (d,δ)-edge expander for $\delta = 1$ and let $E_1 \subseteq E(G)$ be any moderate-size edge set (i.e., $|E_1| \le cn$ for $c = \delta/4 = 1/4$).

Then there is a subset $E_2 \subseteq E_1$ of size $\ge \gamma |E_1|$ for some $\gamma > 0$ such that if ρ is a uniformly randomly sampled charge-preserving assignment to E_1, it holds that all edges in E_2 are assigned uniformly and independently at random.

Taking Lemma 8 on faith for now, we can prove Theorem 3 (for $\delta \geq 1$).

Proof (of Thm 3)

Let S be a complete strategy for prosecutor for $Ts(G_n, \mathcal{V}_n)$ for n large enough.

Any game goes through informative record R with probability 1. This R is consistent with ρ_{init}.

If we can prove for any fixed informative R that

$$\Pr[R \,\&\, \rho_{init} \text{ consistent}] \leq 2^{-\varepsilon n}$$

it follows that size of $S \geq$ # informative records in $S \geq 2^{\varepsilon n}$

$$\Pr[R \,\&\, \rho_{init} \text{ consistent}] \leq \qquad (\dagger)$$

$$\Pr\left[|Dom(R) \cap E'| \leq \frac{c^2}{4d} n\right] + \qquad (*)$$

$$\Pr\left[R \,\&\, \rho_{init} \text{ consistent} \;\middle|\; |Dom(R) \cap E'| \geq \frac{c^2}{4d} n\right] \qquad (**)$$

By $(\dagger)$ above we have $(*) \leq 2^{-\varepsilon' n}$

For $(**)$ we have that $\gamma |E_1^*| \geq \gamma \frac{c^2}{4d} n$ edges are set uniformly and independently at random. Agreement with R with probability $\leq \left(\frac{1}{2}\right)^{\frac{\gamma c^2}{4d} n} = 2^{-\varepsilon'' n}$

Combining this we get

$$\Pr[\, R \,\&\, g_{init} \text{ consistent}] \leq 2^{-\varepsilon' n} + 2^{-\varepsilon'' n}$$
$$\leq 2^{-\varepsilon n}$$

for some $\varepsilon > 0$ and Theorem 3 follows $\boxed{}$

Lemma 8 follows from the following two lemmas

Lemma 9

For $G = (V, E)$ a (d, δ)-expander with $\delta > 0$, suppose that $E_1 \subseteq E(G)$ is a moderate-size set and that $E_2 \subseteq E_1$ does not disconnect G (i.e., $G' = (V, E \setminus E_2)$ is a connected graph). Then uniformly random sampling of a charge-preserving assignment to E_1 gives a uniformly random sample of $\{0,1\}^{|E_2|}$ for the edges in E_2.

(We will use the short-hand $\{0,1\}^{E_2}$ for the set of all possible assignments to edges in E_2.)

Lemma 10

Let $G = (V, E)$ be an (d, δ)-expander with $\delta \geq 1$ and let $E_1 \subseteq E$ be any moderate-size set. Then there is a subset of edges $E_2 \subseteq E_1$ of size $|E_2| = \Omega(|E_1|)$ such that E_2 does not disconnect G.

i.e., $\exists$ global constant $\gamma > 0$ such that $|E_2| \geq \gamma |E_1|$

Proof of Lemma 8

Consider the set E_1 in Lemma 8, which is of moderate size.
Lemma 10 guarantees existence of E_2 of size $|E_2| \geq \gamma |E_1|$ for some $\gamma > 0$ such that $G' = (V, E \setminus E_2)$ connected.
Now Lemma 9 says when we randomly sample a charge-preserving assignment to E_1 we get uniform and independent random bits in E_2 $\boxtimes$

Lemma 9 is mostly some linear algebra juggling and will probably appear on pset 1. Let us do Lemma 10 first.

Proof of Lemma 10

Let $E_1 \subseteq E(G)$ be any moderate-size set i.e. $|E_1| \leq cn$ for $c = \delta/4$.
Look at all ~~the~~ small ~~connected~~ components in $G' = (V, E \setminus E_1)$. Let sum of their sizes $= s$
By assumption $s < n/2$.

Case 1 $s \leq |E_1|/2d$
(Sum of sizes of small components is rather small.)
Then the total #edges not completely inside the large component is at most

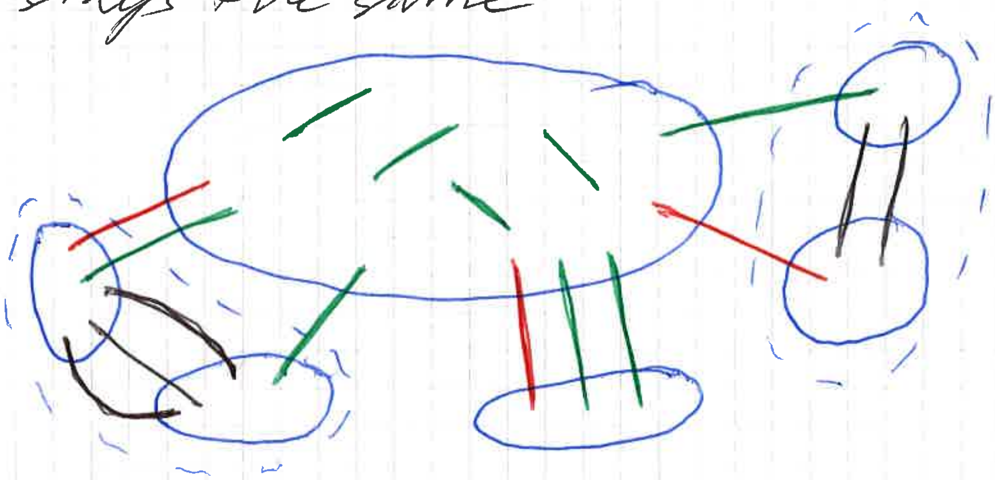
$$sd \leq |E_1|/2$$

so if we pick E_2 to be the edges in E_1 between vertices in the large component we have $|E_2| \geq |E_1|/2$.

<u>Case 2</u> $s > |\mathcal{E}_1| / 2d$

The sum of the sizes of the small components is larger, but not too large (by connectivity expansion)

Remove from $\mathcal{E}_1$ any edges not incident to the large component (BLACK EDGES below) to get $\mathcal{E}^* \subseteq \mathcal{E}$. This might merge some small components (dashed lines), but they are still smallish and their total sizes stays the same



All edges in $\mathcal{E}^*$ are incident to large component
By edge expansion (from the small components) we have

$$|\mathcal{E}^*| \geq \delta s > \delta |\mathcal{E}_1| / 2d$$

Fix one edge per merged small component going to the large component (RED EDGES) Let

$$\mathcal{E}^{**} = \mathcal{E}^* \setminus \{\text{red edges}\}$$

Since $\delta > 1$, every small component has ≥ 2 edges to the large component

Hence, red edges are at most half of $\mathcal{E}^*$. The remaining edges in $\mathcal{E}^{**}$ are the GREEN EDGES.

By construction

$$G^{**} = (V, \mathcal{E}^{**})$$

is connected and

$$|\mathcal{E}^{**}| \geq |\mathcal{E}^*|/2 > \frac{\delta |E_1|}{4d} = \Omega(|E_1|)$$

The lemma follows. $\blacksquare$

It remains to prove Lemma 9 that if $\mathcal{E}'$ does not disconnect G, then sampling randomly a charge-preserving assignment to a superset $\mathcal{E}'' \supseteq \mathcal{E}'$ yields uniformly random bits on $\mathcal{E}'$!

We will mostly leave this as a linear algebra exercise (that might appear on the first problem set), but some useful background facts are provided in the following notes.

LECTURE 3: LINEAR ALGEBRA APPENDIX

Suppose that V is a vector space over some field $\mathbb{F}$. From before, you might be used to $\mathbb{F}$ being the real numbers $\mathbb{R}$ or the complex number $\mathbb{C}$, but here we will have $\mathbb{F} = GF(2)$, i.e., the field with two elements $0, 1$ such that

$$0+0=0 \qquad 0+1=1 \qquad 1+1=0$$
$$0\cdot f = 0 \qquad 1\cdot f = f \quad \text{for } f\in\{0,1\}$$

Most of the basic facts about vector spaces still hold in this setting.

An AFFINE SUBSPACE A of V is a set $A = \{ a + u \mid u \in U \}$ for some fixed $a \in V$ and some fixed (linear) subspace $U \subseteq V$.

That is, an affine space is a linear space shifted by a constant vector.
If $u, v \in A$, then $u - v \in A$, but in general $u + v$ might not be in A
If always $u, v \in A$ implies $u + v \in A$ then A is linear; i.e. we can choose the offset $a = 0$.

For simplicity, in what follows let us focus on affine subspaces of $\{0,1\}^m = GF(2)^m$

FACT A

Any affine subspace $A \subseteq \{0,1\}^m$ of dimension $n \leq m$ is generated by $Mx + a$, for some fixed (but not unique) $m \times n$ matrix M of (full) rank n and some fixed size-m column vector a, when we let x range over (all column vectors in) $\{0,1\}^n$. Uniform random sampling from A can be performed by choosing a uniformly random $x \in \{0,1\}^n$.

We will just accept this as true — you can take it as a definition if you like.

PROPOSITION B

Let $A \subseteq \{0,1\}^m$ be an affine subspace. Suppose for a subset of coordinates $S \subseteq [m]$ that all bitstrings in $\{0,1\}^S$ are supported by A (i.e., there are vectors $u \in A$ that agree with any bitpattern in $\{0,1\}^S$). Then a uniformly random sample from A yields uniformly random and independent bits when restricted to $\{0,1\}^S$

Proof Exercise, but let us hint at two possible solutions.

Approach 1 Argue that (by the way we have defined things) in any affine subspace $A \subseteq \{0,1\}^m$ any bit that is not fixed to 0 or 1 is 0 in exactly half of the vectors and 1 in exactly half of the vectors (why?)

Repeat this on bit after bit in S, using that fixing a bit yields another affine subspace $A' \subseteq A$ (why?)

Argue that every fixed bit pattern in $\{0,1\}^S$ must appear in a fraction $1/2^{|S|}$ of the vectors in A.

Approach 2

The rows $\{R_i \mid i \in S\}$ in M must be linearly independent. (Why?)
This means that the submatrix consisting of these rows has rank $|S|$, and hence there exists a set of T columns, $|T| = |S|$, such that the submatrix on rows S and columns T is invertible. Argue that for any choice of the values of x outside of the coordinates in T the values of $Mx + a$ over $\{0,1\}^T$, restricted to the coordinates in S, is one-to-one and hence uniform over random x.

Phrased differently, Proposition B says that if an affine subspace is supported on the uniform distribution of some set of coordinates S, then sampling from A uniformly at random and restricting to the coordinates in S yields the uniform distribution over $\{0,1\}^S$.

OBSERVATION C

Let $G = (V, E)$ be any connected graph, let $\chi : V \to \{0,1\}$ be any odd-charge function, and let $E' \subseteq E$ be a set of edges such that $G' = (V, E \setminus E')$ has a (unique) connected component of size $> |V|/2$. Then the set of charge-preserving assignments $A \subseteq \{0,1\}^{E'}$ form an affine subspace

To prove this we need another observation

OBSERVATION D

Let $G = (V, E)$ be a connected graph with an charge function χ and let $E' \subseteq E$ be a minimal set of edges that disconnects G into two connected subgraphs G_1 and G_2. Then total charges of the subgraphs G_1 and G_2 resulting from any assignment $\rho : E' \to \{0,1\}$ depends only on χ and on the parity of

$$\sum_{e \in E'} \rho(e)$$

Proof Exercise.

Using Observation D it is straightforward, if a bit tedious, to prove Observation C

Proof of Observation C (sketch)

Consider $G' = (V, E \setminus E')$. Let G_0 be the unique large component and G_i, $i = 1, \ldots, s$, the small (connected) components.
Let $E_{i,j}$ be the edges between G_i and G_j

Look first at $G^0 = (V, E \setminus \bigcup_{j=1}^{s} E_{0,j})$
This yields one affine constraint per ~~connected~~ small component in G^0 requiring that the odd charge is pushed into the large component.
~~Given any~~ Now consider $E_{i,j} \neq \emptyset$ ~~for~~ in order for $1 \leq i < j \leq s$.
If adding $E_{i,j}$ to the set of previously considered edges adds ~~a~~ new connected components, i.e., splits ~~one~~ small component into two smaller ones, the constraint that both new small components should get even charge is an affine constraint on the edges $E_{i,j}$ and the previously assigned ~~edges~~. All of these constraints can be collected in matrix form $M^* y = b$, and the set of solutions can be written as $Mx + b$ for some x of suitable dimension. $\boxtimes$

PROPOSITION $\mathcal{E}$

Suppose $G = (V, E)$ is a connected graph with an odd-charge function χ and let $\mathcal{E}_1 \subseteq E$ be such that $G_1 = (V, \mathcal{E} \setminus \mathcal{E}_1)$ has a unique connected component of size $> |V|/2$. Then for any fixed $\mathcal{E}_2 \subseteq \mathcal{E}_1$ the set of charge-preserving assignments to $\mathcal{E}_1$ has full support on $\{0,1\}^{\mathcal{E}_2}$ if and only if $G_2 = (V, \mathcal{E} \setminus \mathcal{E}_2)$ is connected.

<u>Proof</u> ($\Leftarrow$) If G_2 is connected, then clearly we can assign all edges $e \in E_2$ arbitrarily, since there is only one component and its charge stays the same. Assignments to edges in $\mathcal{E}_1 \setminus E_2$ will take care of charges when the graph is disconnected. Hence we have full support on $\{0,1\}^{E_2}$

($\Rightarrow$) Pick a minimal set $\mathcal{E}^* \subseteq \mathcal{E}_2$ that disconnects G. The parity of this set of edges must be such that the odd charge stays in the large components. Hence we can only have half of the assignments to $\{0,1\}^{E^*}$ and, in particular, do not have full support on $\{0,1\}^{E_2}$.

Now it is straightforward to prove Lemma 9 in the notes for Lecture 3, stated here again for reference.

LEMMA 9

For $G = (V, E)$ a (d, δ)-expander with $\delta > 0$, suppose that $\mathcal{E}_1 \subseteq E$ is a moderate-size set and that $\mathcal{E}_2 \subseteq \mathcal{E}_1$ does not disconnect G. Then uniformly random sampling of charge-preserving assignments to $\mathcal{E}_1$ yields uniformly random samples of $\{0, 1\}^{\mathcal{E}_2}$.

Proof An exercise in putting together the Facts, Propositions and Observations we have covered so far.