Expander Graphs and Their Applications (XV)

Yijia Chen
Shanghai Jiaotong University

Review of the Previous Lecture

Amplification Lemma

Lemma

Let $0 < \lambda < d$ and $|\Sigma|$ be constants. Then there exists a constant $\beta_2 = \beta_2(\lambda, d, |\Sigma|) > 0$ such that for every $t \in \mathbb{N}$ and for every d-regular constraint graph G with a selfloop on each vertex and $\lambda(G) \leq \lambda$,

$$\mathsf{unsat}(G^t) \geq \beta_2 \cdot \sqrt{t} \cdot \min\left(\mathsf{unsat}(G), \frac{1}{t}\right).$$

Amplification Lemma (stronger version)

Lemma

Let $0 < \lambda < d$ and $|\Sigma|$ be constants. Then there exists a constant $\beta_2 = \beta_2(\lambda, d, |\Sigma|) > 0$ such that for every $t \in \mathbb{N}$ and for every d-regular constraint graph G with a selfloop on each vertex and $\lambda(G) \leq \lambda$, the following hold.

For every $\vec{\sigma} : V \to \Sigma^{d^{\lceil t/2 \rceil}}$ let $\sigma : V \to \Sigma$ be defined according to "popular opinion" by setting, for each $v \in V$, $\sigma(v) := a$ such that

$$\Pr[\text{a random } \lceil t/2 \rceil\text{-step walk in } G \text{ from } v \text{ reaches a vertex } w \text{ with } \vec{\sigma}(w)_v = a],$$

where $\vec{\sigma}(w)_v \in \Sigma$ denotes the restriction of $\vec{\sigma}(w)$ to v, is maximized over all $a \in \Sigma$.

Then

$$\mathrm{unsat}_{\vec{\sigma}}(G^t) \geq \beta_2 \cdot \sqrt{t} \cdot \min\left(\mathrm{unsat}_{\sigma}(G), \frac{1}{t}\right).$$

Proof of the Amplification Lemma

Let $\vec{\sigma} : V \to \Sigma^{d^{\lceil t/2 \rceil}}$ be an assignment for G^t. Define the assignment $\sigma : V \to \Sigma$ as before, i.e., according to "popular opinion."

For every $v \in V$, let X_v be a random variable such that for every $a \in \Sigma$

$$\Pr[X_v = a] = \Pr\Big[\text{a random } \lceil t/2 \rceil\text{-step walk in } G \text{ from } v \text{ reaches a vertex } w \text{ with } \vec{\sigma}(w)_v = a\Big].$$

Then for every $a \in \Sigma$,

$$\Pr\big[X_v = \sigma(v)\big] \geq \Pr\big[X_v = a\big].$$

Hence,

$$\Pr\big[X_v = \sigma(v)\big] \geq \frac{1}{|\Sigma|}.$$

Proof of the Amplification Lemma (cont'd)

Let

$$F := \begin{cases} \{e \in E \mid \sigma \text{ violates } e\} & \text{if } \mathrm{unsat}_\sigma(G) < 1/t \\ \text{an arbitrary subset of the above } \{\ldots\} & \\ \qquad \text{with } |F| = \lfloor |E|/t \rfloor & \text{otherwise.} \end{cases}$$

Then

$$\Omega\left(\frac{|F|}{|E|}\right) = \min\left(\mathrm{unsat}_\sigma(G), \frac{1}{t}\right)$$

From now on, we fix $\vec{\sigma}$, σ, and F.

Proof of the Amplification Lemma (cont'd)

Let $\mathbf{E} := E(G^t) = E^t$. Recall there is a one-to-one correspondence between every edge $\mathbf{e} \in \mathbf{E}$ and every walk of length t in G.
With some abuse of notation we write $\mathbf{e} = (v_0, v_1, \ldots, v_t)$ where $(v_{i-1}, v_i) \in E$ for all $i \in [t]$.

Definition

A walk $\mathbf{e} = (v_0, v_1, \ldots, v_t)$ is hit by its i-th edge if

1. $(v_{i-1}, v_i) \in F$, and
2. Both $\vec{\sigma}(v_0)_{v_{i-1}} = \sigma(v_{i-1})$ and $\vec{\sigma}(v_t)_{v_i} = \sigma(v_i)$.

Remark.

- *$(v_{i-1}, v_i) \in F$ means that the edge (v_{i-1}, v_i) rejects σ.*
- *By 2, v_0 has the major opinion of v_{i-1} in G^t, (which implies that there is a $\lceil t/2 \rceil$-step walk from v_0 to v_{i-1}). Similarly, v_t has the majority opinion of v_i.*

Hence, (v_0, v_t) rejects $\vec{\sigma}$ by our definition of G^t.

Proof of the Amplification Lemma (cont'd)

Let

$$I := \left\{ \frac{t}{2} - \sqrt{\frac{t}{2}} < i \leq \frac{t}{2} + \sqrt{\frac{t}{2}} \right\} \subseteq \mathbb{N}$$

be the set of "middle" indices. For each walk $\mathbf{e}$, we define

$$N(\mathbf{e}) := \left| \left\{ i \in I \mid \mathbf{e} \text{ is hit by its } i\text{-th edge} \right\} \right|.$$

$N(\mathbf{e}) > 0$ implies that $\mathbf{e}$ rejects $\vec{\sigma}$.

Thus $\Pr_{\mathbf{e}}[N(\mathbf{e}) > 0] \leq \Pr_{\mathbf{e}}[\mathbf{e} \text{ rejects } \vec{\sigma}] = \mathrm{unsat}_{\vec{\sigma}}(G^t)$.

We will prove

$$\Omega(\sqrt{t}) \cdot \frac{|F|}{|E|} \leq \Pr_{\mathbf{e}}[N(\mathbf{e}) > 0].$$

Combined with $\Omega(|F|/|E|) = \min\left(\mathrm{unsat}_{\sigma}(G), 1/t\right)$,

$$\Omega(\sqrt{t}) \cdot \min\left(\mathrm{unsat}_{\sigma}(G), \frac{1}{t}\right) \leq \Omega(\sqrt{t}) \cdot \frac{|F|}{|E|} \leq \Pr_{\mathbf{e}}[N(\mathbf{e}) > 0] \leq \mathrm{unsat}_{\vec{\sigma}}(G^t).$$

Proof of the Amplification Lemma (cont'd)

To show:

$$\Omega(\sqrt{t}) \cdot \frac{|F|}{|E|} \leq \Pr_{\mathbf{e}}[N(\mathbf{e}) > 0].$$

we will prove two lemmas.

Lemma

$$\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot \frac{|F|}{|E|}.$$

Lemma

$$\mathbb{E}_{\mathbf{e}}\left[\left(N(\mathbf{e})\right)^2\right] \leq O(\sqrt{t}) \cdot \frac{|F|}{|E|}.$$

Proof of the Amplification Lemma (cont'd)

From probability theory:

Lemma

For every non-negative random variable X which is not identically zero,

$$\Pr[X > 0] \geq \frac{\mathbb{E}^2[X]}{\mathbb{E}[X^2]}.$$

Proof.

$$\mathbb{E}[X] = \mathbb{E}\left[X \cdot 1_{X>0}\right] \leq \sqrt{\mathbb{E}[X^2]} \cdot \sqrt{\mathbb{E}\left[(1_{X>0})^2\right]} = \sqrt{\mathbb{E}[X^2]} \cdot \sqrt{\Pr[X > 0]}.$$

□

Now by the previous lemmas

$$\Pr[N(\mathbf{e}) > 0] \geq \mathbb{E}^2[N(\mathbf{e})]/\mathbb{E}\left[\left(N(\mathbf{e})\right)^2\right] = \Omega(\sqrt{t}) \cdot \frac{|F|}{|E|}.$$

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$

Recall we define $I := \left\{ t/2 - \sqrt{t/2} < i \leq t/2 + \sqrt{t/2} \right\}$.

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$

Recall we define $I := \left\{ t/2 - \sqrt{t/2} < i \leq t/2 + \sqrt{t/2} \right\}$. For every $i \in I$ and every $\mathbf{e} \in \mathbf{E}$ we define an indicator variable

$$N_i(\mathbf{e}) := 1 \quad \Longleftrightarrow \quad \text{the walk } \mathbf{e} \text{ is hit by its } i\text{-th edge.}$$

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$

Recall we define $I := \left\{ t/2 - \sqrt{t/2} < i \leq t/2 + \sqrt{t/2} \right\}$. For every $i \in I$ and every $\mathbf{e} \in \mathbf{E}$ we define an indicator variable

$$N_i(\mathbf{e}) := 1 \quad \Longleftrightarrow \quad \text{the walk } \mathbf{e} \text{ is hit by its } i\text{-th edge.}$$

Then for every $\mathbf{e} \in \mathbf{E}$

$$N(\mathbf{e}) = \sum_{i \in I} N_i(\mathbf{e}).$$

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$

Recall we define $I := \left\{ t/2 - \sqrt{t/2} < i \leq t/2 + \sqrt{t/2} \right\}$. For every $i \in I$ and every $\mathbf{e} \in \mathbf{E}$ we define an indicator variable

$$N_i(\mathbf{e}) := 1 \quad \Longleftrightarrow \quad \text{the walk } \mathbf{e} \text{ is hit by its } i\text{-th edge.}$$

Then for every $\mathbf{e} \in \mathbf{E}$

$$N(\mathbf{e}) = \sum_{i \in I} N_i(\mathbf{e}).$$

And by **the linearity of expectation**

$$\mathbb{E}[N] = \sum_{i \in I} \mathbb{E}[N_i].$$

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

To estimate $\mathbb{E}[N_i]$, we choose $\mathbf{e} \in \mathbf{E}$ according to the following distribution:

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

To estimate $\mathbb{E}[N_i]$, we choose $\mathbf{e} \in \mathbf{E}$ according to the following distribution:

1. Choose $e = (u, v) \in E$ uniformly at random.

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

To estimate $\mathbb{E}[N_i]$, we choose $\mathbf{e} \in \mathbf{E}$ according to the following distribution:

1. Choose $e = (u, v) \in E$ uniformly at random.
2. Choose a random walk of length $i - 1$ from u, i.e., $u = v_{i-1}, v_{i-2}, \ldots, v_0$.

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

To estimate $\mathbb{E}[N_i]$, we choose $\mathbf{e} \in \mathbf{E}$ according to the following distribution:

1. Choose $e = (u, v) \in E$ uniformly at random.
2. Choose a random walk of length $i - 1$ from u, i.e., $u = v_{i-1}, v_{i-2}, \ldots, v_0$.
3. Choose a random walk of length $t - i$ from v, i.e., $v = v_i, v_{i+1}, \ldots, v_t$.

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

To estimate $\mathbb{E}[N_i]$, we choose $\mathbf{e} \in \mathbf{E}$ according to the following distribution:

1. Choose $e = (u, v) \in E$ uniformly at random.
2. Choose a random walk of length $i - 1$ from u, i.e., $u = v_{i-1}, v_{i-2}, \ldots, v_0$.
3. Choose a random walk of length $t - i$ from v, i.e., $v = v_i, v_{i+1}, \ldots, v_t$.
4. Output the walk $\mathbf{e} = (v_0, \ldots, v_t)$.

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

To estimate $\mathbb{E}[N_i]$, we choose $\mathbf{e} \in \mathbf{E}$ according to the following distribution:

1. Choose $e = (u, v) \in E$ uniformly at random.
2. Choose a random walk of length $i - 1$ from u, i.e., $u = v_{i-1}, v_{i-2}, \ldots, v_0$.
3. Choose a random walk of length $t - i$ from v, i.e., $v = v_i, v_{i+1}, \ldots, v_t$.
4. Output the walk $\mathbf{e} = (v_0, \ldots, v_t)$.

This yields the uniform distribution on $\mathbf{E}$.

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

For the step

1. Choose $e = (u, v) \in E$ uniformly at random.

$$\Pr[e \in F] = \frac{|F|}{|E|}.$$

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

For the step

1. Choose $e = (u, v) \in E$ uniformly at random.

$$\Pr[e \in F] = \frac{|F|}{|E|}.$$

For the steps:

2. Choose a random walk of length $i - 1$ from u, i.e., $u = v_{i-1}, v_{i-2}, \ldots, v_0$.
3. Choose a random walk of length $t - i$ from v, i.e., $v = v_i, v_{i+1}, \ldots, v_t$.

the choice of v_0 only depends on u, and the choice of v_t only depends on v.

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

For the step

1. Choose $e = (u, v) \in E$ uniformly at random.

$$\Pr[e \in F] = \frac{|F|}{|E|}.$$

For the steps:

2. Choose a random walk of length $i - 1$ from u, i.e., $u = v_{i-1}, v_{i-2}, \ldots, v_0$.
3. Choose a random walk of length $t - i$ from v, i.e., $v = v_i, v_{i+1}, \ldots, v_t$.

the choice of v_0 only depends on u, and the choice of v_t only depends on v. Therefore,

$$\Pr\left[N_i > 0\right] = \frac{|F|}{|E|} \cdot p_u \cdot p_v$$

where

$$p_u := \Pr_{v_0}\left[\vec{\sigma}(v_0)_u = \sigma(u)\right] \quad \text{and} \quad p_v := \Pr_{v_t}[\vec{\sigma}(v_t)_v = \sigma(v)].$$

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

We look at $p_u := \Pr_{v_0}\left[\vec{\sigma}(v_0)_u = \sigma(u)\right]$.

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

We look at $p_u := \Pr_{v_0}\left[\vec{\sigma}(v_0)_u = \sigma(u)\right]$. For every $\ell \in \mathbb{N}$, define a random variable $X_{u,\ell}$ such that for every $a \in \Sigma$

$$\Pr\left[X_{u,\ell} = a\right] = \Pr\Big[\text{a random } \ell\text{-step walk in } G \text{ from } u \text{ ends in a vertex } w \text{ for which } \vec{\sigma}(w)_u = a\Big].$$

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

We look at $p_u := \Pr_{v_0}\left[\vec{\sigma}(v_0)_u = \sigma(u)\right]$. For every $\ell \in \mathbb{N}$, define a random variable $X_{u,\ell}$ such that for every $a \in \Sigma$

$$\Pr\left[X_{u,\ell} = a\right] = \Pr\Big[\text{a random } \ell\text{-step walk in } G \text{ from } u \text{ ends in a vertex } w \text{ for which } \vec{\sigma}(w)_u = a\Big].$$

Therefore

$$p_u = \Pr\left[X_{u,i-1} = \sigma(u)\right]$$

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

We look at $p_u := \Pr_{v_0}\left[\vec{\sigma}(v_0)_u = \sigma(u)\right]$. For every $\ell \in \mathbb{N}$, define a random variable $X_{u,\ell}$ such that for every $a \in \Sigma$

$$\Pr\left[X_{u,\ell} = a\right] = \Pr\Big[\text{a random } \ell\text{-step walk in } G \text{ from } u \text{ ends in a vertex } w \text{ for which } \vec{\sigma}(w)_u = a\Big].$$

Therefore

$$p_u = \Pr\left[X_{u,i-1} = \sigma(u)\right]$$

and

$$\Pr\left[X_{u,t/2} = \sigma(u)\right] \geq \frac{1}{|\Sigma|}.$$

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

We look at $p_u := \Pr_{v_0}\left[\vec{\sigma}(v_0)_u = \sigma(u)\right]$. For every $\ell \in \mathbb{N}$, define a random variable $X_{u,\ell}$ such that for every $a \in \Sigma$

$$\Pr\left[X_{u,\ell} = a\right] = \Pr\Big[\text{a random } \ell\text{-step walk in } G \text{ from } u \text{ ends in a vertex } w \text{ for which } \vec{\sigma}(w)_u = a\Big].$$

Therefore

$$p_u = \Pr\left[X_{u,i-1} = \sigma(u)\right]$$

and

$$\Pr\left[X_{u,t/2} = \sigma(u)\right] \geq \frac{1}{|\Sigma|}.$$

So for $i - 1 = t/2$, we have $p_u \geq 1/|\Sigma|$.

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

We aim to show for all $\ell \in \mathbb{N}$

$$\left|\ell - \frac{t}{2}\right| \leq \sqrt{\frac{t}{2}} \quad \Longrightarrow \quad \Pr\left[X_{u,\ell} = \sigma(u)\right] \geq \frac{\tau}{2} \cdot \Pr\left[X_{u,t/2} = \sigma(u)\right]$$

for a constant $\tau > 0$ to be determined later.

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

We aim to show for all $\ell \in \mathbb{N}$

$$\left|\ell - \frac{t}{2}\right| \leq \sqrt{\frac{t}{2}} \quad \Longrightarrow \quad \Pr\left[X_{u,\ell} = \sigma(u)\right] \geq \frac{\tau}{2} \cdot \Pr\left[X_{u,t/2} = \sigma(u)\right]$$

for a constant $\tau > 0$ to be determined later.

Intuitively, the selfloops of G make the distribution of vertices reached by a random $t/2$-step walk from u roughly the same as distribution on vertices reached by an ℓ-step walk from u, for every $\ell \in I$.

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

Fix some $\ell \in I$.

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

Fix some $\ell \in I$. *Mark one selfloop on each vertex.*

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

Fix some $\ell \in I$. *Mark one selfloop on each vertex.* Observe that every length-ℓ walk from u in G can be equivalently described by

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

Fix some $\ell \in I$. *Mark one selfloop on each vertex.* Observe that every length-ℓ walk from u in G can be equivalently described by

(i) specifying in which which steps the *marked* edges were traversed; and

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

Fix some $\ell \in I$. *Mark one selfloop on each vertex.* Observe that every length-ℓ walk from u in G can be equivalently described by

(i) specifying in which which steps the *marked* edges were traversed; and

(ii) specifying the remaining steps conditioned on choosing only *non-marked* edges.

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

Fix some $\ell \in I$. *Mark one selfloop on each vertex.* Observe that every length-ℓ walk from u in G can be equivalently described by

(i) specifying in which which steps the *marked* edges were traversed; and

(ii) specifying the remaining steps conditioned on choosing only *non-marked* edges.

Let $X'_{u,k}$ be a random variable such that for every $a \in \Sigma$

$$\Pr\left[X'_{u,k} = a\right] = \Pr\left[\text{a } k\text{-step walk in } G \textit{ conditioned on walking only on non-marked edges} \text{ reaches a vertex } w \text{ with } \vec{\sigma}(w)_u = a\right].$$

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

For a binomial variable $B_{\ell,p}$ with

$$\Pr\left[B_{\ell,p} = k\right] = \binom{\ell}{k} \cdot p^k \cdot (1-p)^{\ell-k}$$

with $p := 1 - 1/d$.

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

For a binomial variable $B_{\ell,p}$ with

$$\Pr\left[B_{\ell,p} = k\right] = \binom{\ell}{k} \cdot p^k \cdot (1-p)^{\ell-k}$$

with $p := 1 - 1/d$.

$$\Pr\left[X_{u,\ell} = a\right] = \sum_{k=0}^{\ell} \Pr\left[B_{\ell,p} = k\right] \cdot \Pr\left[X'_{u,k} = a\right].$$

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$ (cont'd)

For a binomial variable $B_{\ell,p}$ with

$$\Pr\left[B_{\ell,p} = k\right] = \binom{\ell}{k} \cdot p^k \cdot (1-p)^{\ell-k}$$

with $p := 1 - 1/d$.

$$\Pr\left[X_{u,\ell} = a\right] = \sum_{k=0}^{\ell} \Pr\left[B_{\ell,p} = k\right] \cdot \Pr\left[X'_{u,k} = a\right].$$

The main point is that

if $\left|\ell_1 - \ell_2\right|$ is small, the the distributions $B_{\ell_1,p}$ and $B_{\ell_2,p}$ are similar.

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$

Lemma

For every $p \in (0,1)$ and $c > 0$ there exists some $\ell_0 > 0$ and $0 < \tau < 1$ such that for every $\ell_1, \ell_2 \in \mathbb{N}$ with $\ell_0 < \ell_1 - \sqrt{\ell_1} \leq \ell_2 < \ell_1 + \sqrt{\ell_1}$ we have

$$\text{for all } k \in \mathbb{N} \text{ if } \left|k - p \cdot \ell_1\right| \leq c \cdot \sqrt{\ell_1}, \text{ then } \tau \leq \frac{\Pr\left[B_{\ell_1,p} = k\right]}{\Pr\left[B_{\ell_2,p} = k\right]} \leq \frac{1}{\tau}$$

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$

Lemma

For every $p \in (0,1)$ and $c > 0$ there exists some $\ell_0 > 0$ and $0 < \tau < 1$ such that for every $\ell_1, \ell_2 \in \mathbb{N}$ with $\ell_0 < \ell_1 - \sqrt{\ell_1} \leq \ell_2 < \ell_1 + \sqrt{\ell_1}$ we have

$$\text{for all } k \in \mathbb{N} \text{ if } \left|k - p \cdot \ell_1\right| \leq c \cdot \sqrt{\ell_1}, \text{ then } \tau \leq \frac{\Pr\left[B_{\ell_1,p} = k\right]}{\Pr\left[B_{\ell_2,p} = k\right]} \leq \frac{1}{\tau}$$

We choose sufficiently large $c > 0$ such that for the set

$$K := \left\{k \in \mathbb{N} \middle| \left|k - p \cdot t/2\right| \leq c \cdot \sqrt{t/2}\right\} \qquad \text{where } p := 1 - \frac{1}{d}$$

we have

$$\Pr_{k \sim B_{t/2,p}}\left[k \notin K\right] = \sum_{k \notin K} \Pr\left[B_{t/2,p} = k\right] < \frac{1}{2 \cdot |\Sigma|}.$$

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$

Lemma

For every $p \in (0,1)$ and $c > 0$ there exists some $\ell_0 > 0$ and $0 < \tau < 1$ such that for every $\ell_1, \ell_2 \in \mathbb{N}$ with $\ell_0 < \ell_1 - \sqrt{\ell_1} \leq \ell_2 < \ell_1 + \sqrt{\ell_1}$ we have

$$\text{for all } k \in \mathbb{N} \text{ if } \left|k - p \cdot \ell_1\right| \leq c \cdot \sqrt{\ell_1}, \text{ then } \tau \leq \frac{\Pr\left[B_{\ell_1,p} = k\right]}{\Pr\left[B_{\ell_2,p} = k\right]} \leq \frac{1}{\tau}$$

We choose sufficiently large $c > 0$ such that for the set

$$K := \left\{k \in \mathbb{N} \middle| \left|k - p \cdot t/2\right| \leq c \cdot \sqrt{t/2}\right\} \qquad \text{where } p := 1 - \frac{1}{d}$$

we have

$$\Pr_{k \sim B_{t/2,p}}\left[k \notin K\right] = \sum_{k \notin K} \Pr\left[B_{t/2,p} = k\right] < \frac{1}{2 \cdot |\Sigma|}.$$

Apply the above lemma on $\ell_1 = t/2$ and $\ell_2 = \ell$: for some $\tau > 0$ and all $k \in K$

$$\Pr\left[B_{\ell,p} = k\right] \geq \tau \cdot \Pr\left[B_{t/2.p} = k\right].$$

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$

We now have for all $\ell \in I$

$$
\begin{aligned}
&\Pr\left[X_{u,\ell} = \sigma(u)\right] \\
&\geq \sum_{k \in K} \Pr\left[B_{\ell,p} = k\right] \cdot \Pr\left[X'_{u,k} = \sigma(u)\right] \\
&\geq \tau \cdot \sum_{k \in K} \Pr\left[B_{t/2,p} = k\right] \cdot \Pr\left[X'_{u,k} = \sigma(u)\right] \\
&\geq \tau \cdot \left(\sum_{k} \Pr\left[B_{t/2,p} = k\right] \cdot \Pr\left[X'_{u,k} = \sigma(u)\right] - \sum_{k \notin K} \Pr\left[B_{t/2,p} = k\right] \right) \\
&\geq \tau \cdot \left(\Pr\left[X_{u,t/2} = \sigma(u)\right] - \frac{1}{2 \cdot |\Sigma|} \right) \geq \frac{\tau}{2} \cdot \Pr\left[X_{u,t/2} = \sigma(u)\right]
\end{aligned}
$$

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$

We now have for all $\ell \in I$

$$\begin{aligned}
&\Pr\left[X_{u,\ell} = \sigma(u)\right] \\
&\geq \sum_{k \in K} \Pr\left[B_{\ell,p} = k\right] \cdot \Pr\left[X'_{u,k} = \sigma(u)\right] \\
&\geq \tau \cdot \sum_{k \in K} \Pr\left[B_{t/2,p} = k\right] \cdot \Pr\left[X'_{u,k} = \sigma(u)\right] \\
&\geq \tau \cdot \left(\sum_{k} \Pr\left[B_{t/2,p} = k\right] \cdot \Pr\left[X'_{u,k} = \sigma(u)\right] - \sum_{k \notin K} \Pr\left[B_{t/2,p} = k\right] \right) \\
&\geq \tau \cdot \left(\Pr\left[X_{u,t/2} = \sigma(u)\right] - \frac{1}{2 \cdot |\Sigma|} \right) \geq \frac{\tau}{2} \cdot \Pr\left[X_{u,t/2} = \sigma(u)\right]
\end{aligned}$$

As both $i - 1$ and $t - i$ are at most $\sqrt{t/2}$ away from $t/2$.

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$

We now have for all $\ell \in I$

$$\begin{aligned}
&\Pr\left[X_{u,\ell} = \sigma(u)\right] \\
&\geq \sum_{k \in K} \Pr\left[B_{\ell,p} = k\right] \cdot \Pr\left[X'_{u,k} = \sigma(u)\right] \\
&\geq \tau \cdot \sum_{k \in K} \Pr\left[B_{t/2,p} = k\right] \cdot \Pr\left[X'_{u,k} = \sigma(u)\right] \\
&\geq \tau \cdot \left(\sum_{k} \Pr\left[B_{t/2,p} = k\right] \cdot \Pr\left[X'_{u,k} = \sigma(u)\right] - \sum_{k \notin K} \Pr\left[B_{t/2,p} = k\right] \right) \\
&\geq \tau \cdot \left(\Pr\left[X_{u,t/2} = \sigma(u)\right] - \frac{1}{2 \cdot |\Sigma|} \right) \geq \frac{\tau}{2} \cdot \Pr\left[X_{u,t/2} = \sigma(u)\right]
\end{aligned}$$

As both $i-1$ and $t-i$ are at most $\sqrt{t/2}$ away from $t/2$.

$$p_u, p_v \geq \frac{\tau}{2 \cdot |\Sigma|}.$$

Proof of $\mathbb{E}_{\mathbf{e}}[N(\mathbf{e})] \geq \Omega(\sqrt{t}) \cdot |F|/|E|$

We now have for all $\ell \in I$

$$\begin{aligned}
&\Pr\left[X_{u,\ell} = \sigma(u)\right] \\
&\geq \sum_{k \in K} \Pr\left[B_{\ell,p} = k\right] \cdot \Pr\left[X'_{u,k} = \sigma(u)\right] \\
&\geq \tau \cdot \sum_{k \in K} \Pr\left[B_{t/2,p} = k\right] \cdot \Pr\left[X'_{u,k} = \sigma(u)\right] \\
&\geq \tau \cdot \left(\sum_{k} \Pr\left[B_{t/2,p} = k\right] \cdot \Pr\left[X'_{u,k} = \sigma(u)\right] - \sum_{k \notin K} \Pr\left[B_{t/2,p} = k\right] \right) \\
&\geq \tau \cdot \left(\Pr\left[X_{u,t/2} = \sigma(u)\right] - \frac{1}{2 \cdot |\Sigma|} \right) \geq \frac{\tau}{2} \cdot \Pr\left[X_{u,t/2} = \sigma(u)\right]
\end{aligned}$$

As both $i - 1$ and $t - i$ are at most $\sqrt{t/2}$ away from $t/2$.

$$p_u, p_v \geq \frac{\tau}{2 \cdot |\Sigma|}.$$

Recall $\Pr[N_i > 0] = |F|/|E| \cdot p_u \cdot p_v$, so $\mathbb{E}[N_i] \geq \Omega(1) \cdot |F|/|E|$. □

Lemma

For every $p \in (0,1)$ and $c > 0$ there exists some $0 < \tau < 1$ such that for every $n, m \in \mathbb{N}$ with $4 \leq n - \sqrt{n} \leq m < n + \sqrt{n}$ and all $k \in \mathbb{N}$ with $\left|k - p \cdot n\right| \leq c \cdot \sqrt{n}$ we have

$$\tau \cdot \Pr\left[B_{m,p} = k\right] \leq \Pr\left[B_{n,p} = k\right] \leq \frac{\Pr\left[B_{m,p} = k\right]}{\tau}$$

Lemma

For every $p \in (0,1)$ and $c > 0$ there exists some $0 < \tau < 1$ such that for every $n, m \in \mathbb{N}$ with $4 \leq n - \sqrt{n} \leq m < n + \sqrt{n}$ and all $k \in \mathbb{N}$ with $\left|k - p \cdot n\right| \leq c \cdot \sqrt{n}$ we have

$$\tau \cdot \Pr\left[B_{m,p} = k\right] \leq \Pr\left[B_{n,p} = k\right] \leq \frac{\Pr\left[B_{m,p} = k\right]}{\tau}$$

Proof.

Lemma

For every $p \in (0,1)$ and $c > 0$ there exists some $0 < \tau < 1$ such that for every $n, m \in \mathbb{N}$ with $4 \leq n - \sqrt{n} \leq m < n + \sqrt{n}$ and all $k \in \mathbb{N}$ with $\left|k - p \cdot n\right| \leq c \cdot \sqrt{n}$ we have

$$\tau \cdot \Pr\left[B_{m,p} = k\right] \leq \Pr\left[B_{n,p} = k\right] \leq \frac{\Pr\left[B_{m,p} = k\right]}{\tau}$$

Proof.

First assume $m \leq n$ and recall $n - \sqrt{n} \leq m$. Therefore

$$n = m + r \quad \text{for some } 0 \leq r \leq \sqrt{n}.$$

Lemma

For every $p \in (0,1)$ and $c > 0$ there exists some $0 < \tau < 1$ such that for every $n, m \in \mathbb{N}$ with $4 \leq n - \sqrt{n} \leq m < n + \sqrt{n}$ and all $k \in \mathbb{N}$ with $\left|k - p \cdot n\right| \leq c \cdot \sqrt{n}$ we have

$$\tau \cdot \Pr\left[B_{m,p} = k\right] \leq \Pr\left[B_{n,p} = k\right] \leq \frac{\Pr\left[B_{m,p} = k\right]}{\tau}$$

Proof.

First assume $m \leq n$ and recall $n - \sqrt{n} \leq m$. Therefore

$$n = m + r \quad \text{for some } 0 \leq r \leq \sqrt{n}.$$

Recall

$$\binom{m+1}{k} = \frac{m+1}{m+1-k} \cdot \binom{m}{k}.$$

Proof. (cont')

Proof. (cont')

$$
\begin{aligned}
&\Pr\left[B_{n,p} = k\right] \\
&= \binom{m+r}{k} \cdot p^k \cdot (1-p)^{m+r-k} \\
&= \frac{m+1}{m+1-k} \cdot \frac{m+2}{m+2-k} \cdots \frac{m+r}{m+r-k} \cdot \binom{m}{k} \cdot p^k \cdot (1-p)^{m-k} \cdot (1-p)^r \\
&= X \cdot p^k \cdot (1-p)^{m-k} \cdot \binom{m}{k} \\
&= X \cdot \Pr\left[B_{m,p} = k\right],
\end{aligned}
$$

where

$$
X = (1-p)^r \cdot \frac{m+1}{m+1-k} \cdot \frac{m+2}{m+2-k} \cdots \frac{m+r}{m+r-k}.
$$

Proof. (cont')

$$
\begin{aligned}
&\Pr\left[B_{n,p} = k\right] \\
&= \binom{m+r}{k} \cdot p^k \cdot (1-p)^{m+r-k} \\
&= \frac{m+1}{m+1-k} \cdot \frac{m+2}{m+2-k} \cdots \frac{m+r}{m+r-k} \cdot \binom{m}{k} \cdot p^k \cdot (1-p)^{m-k} \cdot (1-p)^r \\
&= X \cdot p^k \cdot (1-p)^{m-k} \cdot \binom{m}{k} \\
&= X \cdot \Pr\left[B_{m,p} = k\right],
\end{aligned}
$$

where

$$
X = (1-p)^r \cdot \frac{m+1}{m+1-k} \cdot \frac{m+2}{m+2-k} \cdots \frac{m+r}{m+r-k}.
$$

In the following, we bound X.

Proof. (cont'd)

Proof. (cont'd)

For every $a \leq r \leq \sqrt{n}$ we let

$$X_a = \frac{m+a}{m+a-k}$$

Proof. (cont'd)

For every $a \leq r \leq \sqrt{n}$ we let

$$X_a = \frac{m+a}{m+a-k} = \frac{1}{1-k/(m+a)}.$$

Proof. (cont'd)

For every $a \leq r \leq \sqrt{n}$ we let

$$X_a = \frac{m+a}{m+a-k} = \frac{1}{1-k/(m+a)}.$$

Observe

$$\frac{1}{1-k/m} \geq \frac{1}{1-k/(m+a)} \geq \frac{1}{1-k/n}.$$

Proof. (cont'd)

For every $a \leq r \leq \sqrt{n}$ we let

$$X_a = \frac{m+a}{m+a-k} = \frac{1}{1-k/(m+a)}.$$

Observe

$$\frac{1}{1-k/m} \geq \frac{1}{1-k/(m+a)} \geq \frac{1}{1-k/n}.$$

Since $k \geq p \cdot n - c \cdot \sqrt{n}$,

$$X_a \geq \frac{1}{1-k/n} \geq \frac{1}{1-p+c/\sqrt{n}} = \frac{1}{1-p} \cdot \frac{1}{1+\big(c/(1-p)\big) \cdot (1/\sqrt{n})}.$$

Proof. (cont'd)

For every $a \leq r \leq \sqrt{n}$ we let

$$X_a = \frac{m+a}{m+a-k} = \frac{1}{1-k/(m+a)}.$$

Observe

$$\frac{1}{1-k/m} \geq \frac{1}{1-k/(m+a)} \geq \frac{1}{1-k/n}.$$

Since $k \geq p \cdot n - c \cdot \sqrt{n}$,

$$X_a \geq \frac{1}{1-k/n} \geq \frac{1}{1-p+c/\sqrt{n}} = \frac{1}{1-p} \cdot \frac{1}{1+\big(c/(1-p)\big) \cdot (1/\sqrt{n})}.$$

Now let

$$\tau_1 := \left(1 + \frac{c}{1-p} \cdot \frac{1}{\sqrt{n}}\right)^{-r} \leq (1-p)^r \cdot \prod_{a=1}^{r} X_a = X.$$

Proof. (cont'd)

Proof. (cont'd)

Similarly, since $n - \sqrt{n} \leq m$, $p \cdot n + c \cdot \sqrt{n} \leq k$, $p \leq 1$, and $n \geq 4$

$$\frac{k}{m} \leq \frac{p \cdot n + c \cdot \sqrt{n}}{n - \sqrt{n}} \leq p + \frac{p \cdot \sqrt{n} + c \cdot \sqrt{n}}{n - \sqrt{n}} \leq p + \frac{2 \cdot c + 2}{\sqrt{n}}.$$

Proof. (cont'd)

Similarly, since $n - \sqrt{n} \leq m$, $p \cdot n + c \cdot \sqrt{n} \leq k$, $p \leq 1$, and $n \geq 4$

$$\frac{k}{m} \leq \frac{p \cdot n + c \cdot \sqrt{n}}{n - \sqrt{n}} \leq p + \frac{p \cdot \sqrt{n} + c \cdot \sqrt{n}}{n - \sqrt{n}} \leq p + \frac{2 \cdot c + 2}{\sqrt{n}}.$$

Hence,

$$1 - \frac{k}{m} \geq 1 - p - \frac{2 \cdot c + 2}{\sqrt{n}}.$$

Proof. (cont'd)

Similarly, since $n - \sqrt{n} \leq m$, $p \cdot n + c \cdot \sqrt{n} \leq k$, $p \leq 1$, and $n \geq 4$

$$\frac{k}{m} \leq \frac{p \cdot n + c \cdot \sqrt{n}}{n - \sqrt{n}} \leq p + \frac{p \cdot \sqrt{n} + c \cdot \sqrt{n}}{n - \sqrt{n}} \leq p + \frac{2 \cdot c + 2}{\sqrt{n}}.$$

Hence,

$$1 - \frac{k}{m} \geq 1 - p - \frac{2 \cdot c + 2}{\sqrt{n}}.$$

Then

$$X_a \leq \frac{1}{1 - k/m} \leq \frac{1}{1 - p - (2 \cdot c + 2)/\sqrt{n}}$$

Proof. (cont'd)

Similarly, since $n - \sqrt{n} \leq m$, $p \cdot n + c \cdot \sqrt{n} \leq k$, $p \leq 1$, and $n \geq 4$

$$\frac{k}{m} \leq \frac{p \cdot n + c \cdot \sqrt{n}}{n - \sqrt{n}} \leq p + \frac{p \cdot \sqrt{n} + c \cdot \sqrt{n}}{n - \sqrt{n}} \leq p + \frac{2 \cdot c + 2}{\sqrt{n}}.$$

Hence,

$$1 - \frac{k}{m} \geq 1 - p - \frac{2 \cdot c + 2}{\sqrt{n}}.$$

Then

$$X_a \leq \frac{1}{1 - k/m} \leq \frac{1}{1 - p - (2 \cdot c + 2)/\sqrt{n}}$$

Let

$$\tau_2 := \left(\frac{1}{1 - p - (2 \cdot c + 2)/\sqrt{n}} \right)^{-r} \geq (1 - p)^r \cdot \prod_{a=1}^{r} X_a = X.$$

Proof. (cont'd)

Proof. (cont'd)

Recall

$$\tau_1 := \left(1 + \frac{c}{1-p} \cdot \frac{1}{\sqrt{n}}\right)^{-r} \leq X$$

$$\tau_2 := \left(\frac{1}{1-p-(2\cdot c+2)/\sqrt{n}}\right)^{-r} \geq X.$$

Proof. (cont'd)

Recall

$$\tau_1 := \left(1 + \frac{c}{1-p} \cdot \frac{1}{\sqrt{n}}\right)^{-r} \leq X$$

$$\tau_2 := \left(\frac{1}{1-p-(2\cdot c+2)/\sqrt{n}}\right)^{-r} \geq X.$$

Since $r \leq \sqrt{n}$, both τ_1 and τ_2 can be bounded by constants which only depend on c and p.

Proof. (cont'd)

Recall

$$\tau_1 := \left(1 + \frac{c}{1-p} \cdot \frac{1}{\sqrt{n}}\right)^{-r} \leq X$$

$$\tau_2 := \left(\frac{1}{1-p-(2\cdot c+2)/\sqrt{n}}\right)^{-r} \geq X.$$

Since $r \leq \sqrt{n}$, both τ_1 and τ_2 can be bounded by constants which only depend on c and p.

We let

$$\tau := \min(\tau_1, 1/\tau_2).$$

Proof. (cont'd)

Proof. (cont'd)

Finally, if $n < m$, recall $n - \sqrt{n} \leq m < n + \sqrt{n}$, then

$$m - \sqrt{m} < n + \sqrt{n} - \sqrt{n} = n < m.$$

Proof. (cont'd)

Finally, if $n < m$, recall $n - \sqrt{n} \leq m < n + \sqrt{n}$, then

$$m - \sqrt{m} < n + \sqrt{n} - \sqrt{n} = n < m.$$

We can deduce the result using the above argument with the roles of m and n reversed, the same p, and

$$c' := c + 1.$$

□