

Notice that in a partition, once we've fixed 3 vertices of C , the rest of C plays more or less the same role as one of the A_i blocks, since it also has size $k - 3$ and also needs to be taken either *in toto* or not at all.

Recall that we fix a rectangle $\mathcal{R} = \mathcal{U} \times \mathcal{M}$ and any matching $H \subseteq E(C \cup D)$, and define

$$p_{\mathcal{M},T}(H) = \Pr_{M \in \mathcal{M}_{\text{all}}(T)} (M \in \mathcal{M} \mid H \subseteq M)$$

and

$$p_{\mathcal{U},T}(H) = \Pr_{U \in \mathcal{U}_{\text{all}}(T)} (U \in \mathcal{U} \mid U \cap C = V(H) \cap C)$$

Last time, we saw the following lemma:

Lemma 1

$$\begin{aligned} \mu_k(\mathcal{R}) &= \mathbb{E}_T \mathbb{E}_{\substack{F \subseteq C \times D \\ \text{perf. matching}}} [p_{\mathcal{M},T}(F)p_{\mathcal{U},T}(F)] \\ \mu_3(\mathcal{R}) &= \mathbb{E}_T \mathbb{E}_{\substack{F \subseteq C \times D \\ \text{perf. matching}}} \mathbb{E}_{H \in \binom{F}{3}} [p_{\mathcal{M},T}(H)p_{\mathcal{U},T}(H)] \end{aligned}$$

Recall that our goal is to prove the following ‘‘Main Lemma’’:

Lemma 2 *For any valid rectangle \mathcal{R} ,*

$$\mu_3(\mathcal{R}) \leq \frac{400}{k^2} \mu_k(\mathcal{R}) + 2^{-\delta(k) \cdot n}$$

Definition 3 *For a given partition T and a 3-edge matching $H \subseteq C \times D$, we distinguish three cases:*

(a) *We call (T, H) unbiased if*

$$0 < p_{\mathcal{U},T}(H) \in (1 \pm \varepsilon)p_{\mathcal{U},T}(F)$$

where F is a perfect matching on $C \times D$, and additionally

$$0 < p_{\mathcal{M},T}(H) \in (1 \pm \varepsilon)p_{\mathcal{M},T}(H')$$

for every matching $H \subseteq H' \subseteq E(C \cup D)$.

(b) *We call (T, H) unlikely if*

$$p_{\mathcal{M},T} \cdot p_{\mathcal{U},T} \leq 2^{-\delta m}$$

where m is the number of A_i (or B_i) blocks, so that $m = \Theta(n)$.

(c) *Otherwise, we call (T, H) bad. We say that it's good if either (a) or (b) happens.*

We can now split up our formula for $\mu_3(\mathcal{R})$ into these three cases:

$$\mu_3(\mathcal{R}) = \mathbb{E}_{T,F,H} \mathbb{E}[(\text{unbiased}(T, H) + \text{unlikely}(T, H) + \text{bad}(T, H))p_{\mathcal{M},T}(H)p_{\mathcal{U},T}(H)]$$

where unbiased, unlikely, and bad denote the indicator random variables of these events. We will show that both of the good cases give us what we want, and that the bad case contributes very little.

Lemma 4 *Fix a partition T and a perfect matching F on $C \times D$. Then*

$$\Pr_{H \in \binom{F}{3}} [\text{unbiased}(T, H)] \leq \frac{100}{k^2}$$

Proof: We claim that if $(T, H), (T, H^*)$ are both unbiased, then $|H \cap H^*| \geq 2$. Indeed, if $|H \cap H^*| \leq 1$, let $u, v \in (V(H^*) \setminus V(H)) \cap C$ be distinct vertices. Consider the matching $H' = H \cup \{(u, v)\}$. By the definition of the unbiased case, we know that $p_{\mathcal{M},T}(H') > 0$. Thus, there exists a perfect matching $M \in \mathcal{M}_{\text{all}}(T) \cap \mathcal{M}$ that extends H' . But since H^* is unbiased, we know that $p_{\mathcal{U},T}(H^*) > 0$, so there exists a cut $U \in \mathcal{U}(T)$ with $U \cap C = V(H^*) \cap C$. But since $u, v \in V(H^*) \subseteq U$, we get that $|\delta(U) \cap M| = 1$. So we have an entry in \mathcal{R} with zero slack, which means that \mathcal{R} is not valid.

With this fact, we see that there can be at most $3k$ (in fact far fewer, namely $k - 2$). Therefore,

$$\Pr[\text{unbiased}(T, H)] \leq \frac{3k}{\binom{k}{3}} \leq \frac{100}{k^2}$$

□

This gives us the first term in our bound of $\mu_3(\mathcal{R})$, namely $400/k^2 \mu_k(\mathcal{R})$. Similarly, the unbiased ones contribute our second term, $2^{-\delta m}$. So all we need is to bound the bad cases.

Lemma 5

$$\mathbb{E}_{T,F,H} [\text{bad}(T, H)p_{\mathcal{M},T}(H)p_{\mathcal{U},T}(H)] \leq \varepsilon \mu_3(\mathcal{R})$$

This implies that

$$\mu_3(\mathcal{R}) \leq \frac{200}{k^2} \mu_k(\mathcal{R}) + 2^{-\delta m} + \varepsilon \mu_3(\mathcal{R})$$

Then rearranging and dividing by $1 - \varepsilon$ gives us what we wanted. So it boils down to proving this bound on the bad set, which we will not prove fully. Instead, we will prove the following:

Lemma 6 (Simpler lemma) *Call (T, H) \mathcal{U} -good if $p_{\mathcal{U},T} \leq 2^{-\delta m}$ or $p_{\mathcal{U},T}(H) \in (1 \pm \varepsilon)p_{\mathcal{U},T}(C)$. Then $\forall \varepsilon \exists \delta$ such that*

$$\Pr_{(T,H)} [(T, H) \text{ is } \mathcal{U}\text{-good}] \geq 1 - \varepsilon$$

To actually prove the full claim, we have to do this, and the analogous thing for matching, and then do a bunch more conditioning to actually get everything to work out. Morally, this suffices, since it tells us that in most cases we will either be unbiased or unlikely, at least as far as the cut-sets are concerned.

To prove this claim, we will use some pseudorandom properties of large subsets of $2^{[m]}$.

Lemma 7 Let $\mathcal{Y} \subseteq 2^{[m]}$ with $|\mathcal{Y}| \geq 2^{(1-\Theta(\varepsilon^3))m}$. Then at least $(1-\varepsilon)m$ elements in $[m]$ are ε -unbiased, meaning that

$$\Pr_{\substack{S \in \mathcal{Y} \\ \text{uniform}}} [i \in S] \in \left[\frac{1-\varepsilon}{2}, \frac{1+\varepsilon}{2} \right]$$

Proof: We will use that the entropy function is submodular. In fact, we will just need the fact that

$$H(X_1, \dots, X_m) \leq \sum_{i=1}^m H(X_i)$$

We will let X_i be the indicator of the event $i \in S$ where $S \in \mathcal{Y}$ is uniformly random. Then

$$H(X_1, \dots, X_m) = H(S) = \log |\mathcal{Y}|$$

On the other hand

$$H(X_i) = -p_i \log p_i - (1-p_i) \log(1-p_i)$$

where $p_i = \Pr(X_i = 1) = \Pr(i \in S)$. So if i is ε -biased, then

$$H(X_i) \leq 1 - \Theta(\varepsilon^2)$$

by Taylor approximation to H and the fact that it attains a maximum at $1/2$. So if we had more than εm ε -biased variables, then

$$\log |\mathcal{Y}| = H(S) \leq \sum_{i=1}^m H(X_i) < m - \varepsilon m \cdot \Theta(\varepsilon^2) = (1 - \Theta(\varepsilon^3))m$$

which contradicts the assumption on the size of \mathcal{Y} . □

Equivalently to the above fact (via e.g. Bayes' rule),

$$\Pr_{\substack{R \subseteq [m] \\ \text{uniform}}} [R \in \mathcal{Y} \mid i \in R] \in (1 \pm \varepsilon) \Pr(R \in \mathcal{Y})$$

Now we choose (T, H) as follows. First, we pick a random partition \tilde{T} uniformly. Then, we pick a matching $H \subseteq C \times D$ with $|H| = 3$ uniformly at random. Finally, we swap $C \setminus V(H)$ with a random block A_i , or keep it as is, each option with probability $1/(m+1)$. This produces (T, H) uniformly randomly. For $I \subseteq [m+1]$, let

$$f(I) = \bigcup_{i \in I} \widetilde{A}_i \cup (C \cap V(H))$$

where $\widetilde{A}_i = A_i$ for $1 \leq i \leq m$ and $\widetilde{A}_{m+1} = C \setminus V(H)$. Recall that a valid cut-set U takes roughly half the blocks \widetilde{A}_i , so really we only want to consider I of roughly half-size. So set

$$\mathcal{Y} = \{I \subseteq [m+1] : f(I) \in \mathcal{U}\}$$

Then the intuition is that either \mathcal{Y} is small, in which case very few extensions of H are good, so $p_{\mathcal{U}, T} \leq 2^{-\delta m}$, or else \mathcal{Y} is large, in which case we can use the pseudorandom structure and get the unbiased conclusion. More formally, the cases are as follows.

Case 1: If $|\mathcal{Y}| \leq 2^{(1-\delta)m}$ where $\delta = \Theta(\varepsilon^3)$, then

$$p_{\mathcal{U},T}(H) \leq \frac{2^{(1-\delta)m}}{\binom{m}{\lceil m/2 \rceil}} \leq \sqrt{m}2^{-\delta m}$$

and we can modify δ a little bit to get the bound we want.

Case 2: If $|\mathcal{Y}| > 2^{(1-\delta)m}$, then by the pseudorandom property $\Pr_{R \subseteq [m]}[R \in \mathcal{Y} \mid i \in R] \in (1 \pm \varepsilon) \Pr(R \in Y)$, we get

$$p_{\mathcal{U},T}(H) \in (1 \pm \varepsilon)p_{\mathcal{U},T}(C)$$

But this will only hold for $\geq (1 - \varepsilon)m$ choices of $C \setminus V(H) = \widetilde{A}_i$. This is precisely the $1 - \varepsilon$ that we had in the claim of the simpler lemma, namely the probability that we're in neither the unlikely nor the unbiased case.