Solution to First Proof Second Batch Problem C4

Anna Gilbert and Joon-Hyeok Yim, Yale

1 Setup and Statement

Fix $c > 1$. Let $G_n \sim G(n, c/n)$ and let $\mathcal{C}_{\max}$ denote its giant component. Let $\gamma = \gamma(c) \in (0, 1)$ be the unique solution of $\gamma = 1 - e^{-c\gamma}$, so that $|\mathcal{C}_{\max}| = \gamma n(1 + o_P(1))$.

Sampling convention. For each graph G_n, fix arbitrarily one geodesic $[u, v]$ for every ordered pair $(u, v) \in C^2_{\max}$. The argument below is pointwise in this choice, hence it also implies the stated result when geodesics are sampled uniformly at random.

Notation.

- $\Gamma_r(v) = \{u : d(u, v) = r\}$ and $N_r(v) = \{u : d(u, v) \leq r\}$.
- $N_r(S) = \bigcup_{v \in S} N_r(v)$ for a set $S \subseteq V(G_n)$.
- For $\delta \geq 0$, define the event

$$B_\delta(\Delta) := \big\{[x, y] \subseteq N_\delta([y, z] \cup [x, z])\big\} \cup \big\{[y, z] \subseteq N_\delta([x, y] \cup [x, z])\big\} \cup \big\{[x, z] \subseteq N_\delta([x, y] \cup [y, z])\big\}.$$

If $\neg B_\delta(\Delta)$ holds, then $\operatorname{slim}(\Delta) > \delta$. Note that we are using the fact that

$$\{\operatorname{slim}(\Delta) \leq \delta\} \subseteq B_\delta(\Delta).$$

Theorem 1. *For every $\epsilon > 0$ there exist constants $M = M(\epsilon, c) > 0$ and $N = N(\epsilon, c)$ such that, setting*

$$\delta_n := \left\lfloor \frac{1}{\log c}\big(\tfrac{1}{2}\log n - 9 \log\log n\big) \right\rfloor - M,$$

for all $n \geq N$,

$$\mathbb{P}_{G_n}\Big(\mathbb{P}\big(\operatorname{slim}(\Delta(x, y, z)) > \delta_n\big) > 1 - \epsilon\Big) > 1 - \epsilon.$$

2 Structural Inputs

We collect three facts, each holding with probability $1 - o(1)$ over G_n.

Lemma 2 (Giant component size)**.** There exist absolute constants $c_1, c_2 > 0$ such that

$$\mathbb{P}\Big(\big||\mathcal{C}_{\max}| - \gamma n\big| > 0.01\,\gamma n\Big) \leq c_1 e^{-c_2 n}.$$

In particular, $|\mathcal{C}_{\max}| \in [0.99\,\gamma n,\ 1.01\,\gamma n]$ with high probability.

Lemma 3 (Tightness of typical distance)**.** Conditional on $x, y \in \mathcal{C}_{\max}$ chosen uniformly, the family $\{d(x,y) - \lfloor \log_c n \rfloor\}_n$ is tight: for every $\eta > 0$ there is $K(\eta) > 0$ with

$$\mathbb{P}\Big(d(x,y) > \lfloor \log_c n \rfloor + K(\eta) \;\Big|\; x, y \in \mathcal{C}_{\max}\Big) < \eta^2$$

for all n sufficiently large.

Lemma 4 (Ball-size upper bound)**.** With probability $1 - o(1)$ over G_n, simultaneously for every $v \in V(G_n)$ and every $0 \le r \le n$,

$$|\Gamma_r(v)| \le (r+1)^2\, c^r\, \log n.$$

Lemmas 2–4 are standard; see [1] for Lemma 3 and [2] for Lemma 4.

3 Key Combinatorial Lemmas

Lemma 5 (Truncated traffic bound)**.** For any graph G, any vertex w, and any $L \ge 0$,

$$T_L(w) := \#\{(x,y) : d(x,y) \le L,\ w \in [x,y]\} \;\le\; \sum_{k+\ell \le L} |\Gamma_k(w)| \cdot |\Gamma_\ell(w)|.$$

Proof. If w lies on a geodesic from x to y with $d(x,y) \le L$, set $k = d(w,x)$ and $\ell = d(w,y)$; then $k + \ell = d(x,y) \le L$, $x \in \Gamma_k(w)$, $y \in \Gamma_\ell(w)$. □

Lemma 6 (Weighted bipartite color lemma)**.** Let $G = (U, V, E)$ be a bipartite (multi)graph with colored edges. Suppose each vertex $u \in U \cup V$ is incident to at most $t(u)$ distinct colors. Then some color class contains at least

$$\frac{|E|^2}{\big(\sum_{u \in U} t(u)\big)\big(\sum_{v \in V} t(v)\big)}$$

edges.

Proof sketch. Let $\mathcal{C}$ be the set of colors and let E_c denote the edges of color c. A double-counting argument (see [2]) yields the stated bound. □

4 Proof of the Main Theorem

Fix $\epsilon > 0$. Set

$$\eta := \epsilon/4, \qquad L_n := \lfloor \log_c n \rfloor + K(\eta),$$

where $K(\eta)$ comes from Lemma 3. Define $M = M(\epsilon)$ by

$$M := \max\left(\log_c\left(\frac{c^{K(\eta)/2}}{0.99\,\epsilon\,(\log c)^7\,\gamma}\right),\ 2\right),$$

and set $\delta_n := \left\lfloor \frac{1}{\log c}\big(\frac{1}{2}\log n - 9\log\log n\big)\right\rfloor - M$.

4.1 Step 1: The "good graph" event

Call G_n *good* if all of the following hold:

(G1) $0.99\,\gamma n \le |\mathcal{C}_{\max}| \le 1.01\,\gamma n$.

(G2) The quenched tail is small:

$$q_{L_n}(G_n) := \frac{\#\{(x,y) \in \mathcal{C}_{\max}^2 : d(x,y) > L_n\}}{|\mathcal{C}_{\max}|^2} \;<\; \eta.$$

(G3) For all v and all r, $|\Gamma_r(v)| \le (r+1)^2 c^r \log n$.

Claim. $\mathbb{P}(G_n \text{ is good}) > 1-\epsilon$ for n sufficiently large.

Proof. By Lemma 2, (G1) fails with exponentially small probability. By Lemma 4, (G3) fails with probability $o(1)$. For (G2), Lemma 3 gives

$$\mathbb{E}[q_{L_n}(G_n)] = \mathbb{P}(d(x,y) > L_n \mid x,y \in \mathcal{C}_{\max}) < \eta^2,$$

so by Markov's inequality $\mathbb{P}(q_{L_n}(G_n) \ge \eta) \le \eta = \epsilon/4$. A union bound yields $\mathbb{P}(\text{good}) > 1-\epsilon/2 > 1-\epsilon$. $\square$

It remains to show: *for any good G_n, $\mathbb{P}(B_{\delta_n}(\Delta)) < \epsilon$.*

4.2 Step 2: Reduction to short triangles

Call $\Delta(x,y,z)$ *L_n-long* if $\max(d(x,y), d(y,z), d(x,z)) > L_n$. By (G2) and a union bound over the three sides,

$$\mathbb{P}(\Delta \text{ is } L_n\text{-long}) \le 3\,q_{L_n}(G_n) < 3\eta = \tfrac{3\epsilon}{4}.$$

Thus

$$\mathbb{P}(B_{\delta_n}) \le \tfrac{3\epsilon}{4} + \mathbb{P}(B_{\delta_n} \text{ and } \Delta \text{ is not } L_n\text{-long})\,. \tag{1}$$

It suffices to show the second term is below $\epsilon/4$.

4.3 Step 3: Bounding the short-and-bad term via bipartite coloring

By symmetry over the three sides, we focus on

$$A := \big\{[x,y] \subseteq N_{\delta_n}([y,z] \cup [x,z])\big\} \cap \{\Delta \text{ not } L_n\text{-long}\}$$

and bound $\mathbb{P}(A)$; the final bound will be tripled. Indeed,

$$\mathbb{P}(B_{\delta_n}(\Delta) \text{ and } \Delta \text{ is not } L_n\text{-long} \mid G_n) \le 3\,\mathbb{P}(A \mid G_n).$$

On A, since $[x,y]$ has length $\le L_n$ and is covered by the δ_n-neighborhood of $[x,z] \cup [y,z]$, there exists a vertex w with

$$w \in [x,y], \qquad d(w,[x,z]) \le \delta_n + 1, \qquad d(w,[y,z]) \le \delta_n + 1,$$

obtained by taking w as the "transition vertex" along $[x,y]$.

Condition on $z \in \mathcal{C}_{\max}$. Fix $z \in \mathcal{C}_{\max}$, and define $p(z) := \mathbb{P}(A \mid G_n, z)$. We show that $p(z) \le \varepsilon/12$ uniformly in z. If A occurs, then

$$[x, y] \subseteq N_{\delta_n}([x, z]) \cup N_{\delta_n}([y, z]).$$

Since $x \in [x, z]$ and $y \in [y, z]$, there is a vertex $w \in [x, y]$ such that

$$d(w, [x, z]) \le \delta_n + 1, \qquad d(w, [y, z]) \le \delta_n + 1.$$

Now, build a bipartite graph $H_z = (U, V, E)$ with:

- $U = V = N_{L_n}(z) \cap \mathcal{C}_{\max}$;
- $(x, y) \in E$ iff A occurs, so that both endpoints lie in $N_{L_n}(z)$ and the covering witness exists;
- **color** edge (x, y) by a choice of witness vertex $w \in [x, y]$ with $d(w, [x, z]), d(w, [y, z]) \le \delta_n + 1$.

By construction $p(z) = |E|/|\mathcal{C}_{\max}|^2$.

Color degree. If x is incident to color w, then there is a geodesic from x through w, so $w \in N_{\delta_n+1}([x, z])$. Hence the number of distinct colors incident to x is at most

$$t(x) \le |N_{\delta_n+1}([x, z])| \le (d(x, z) + 1) \cdot \max_v |N_{\delta_n+1}(v)|.$$

Using (G3),

$$\max_v |N_{\delta_n+1}(v)| \le \sum_{r=0}^{\delta_n+1} (r + 1)^2 c^r \log n \le (\delta_n + 2)^3 \, c^{\delta_n+1} \, \log n,$$

and $d(x, z) \le L_n$, so

$$t(x) \le (L_n + 1)(\delta_n + 2)^3 c^{\delta_n+1} \log n =: T_*.$$

Color-class size. If edge (x, y) has color w, then $w \in [x, y]$ and $d(x, y) \le L_n$. Therefore color w is used at most $T_{L_n}(w)$ times. By Lemmas 5 and 4,

$$T_{L_n}(w) \le \sum_{k+\ell \le L_n} (k + 1)^2 (\ell + 1)^2 c^{k+\ell} (\log n)^2 \le (L_n + 1)^6 \, c^{L_n} \, (\log n)^2.$$

Apply Lemma 6. Some color is used at least $|E|^2/(\sum_u t(u))^2$ times, so

$$\frac{|E|^2}{\big(|N_{L_n}(z)| \cdot T_*\big)^2} \le (L_n + 1)^6 \, c^{L_n} \, (\log n)^2.$$

Taking square roots,

$$|E| \le |N_{L_n}(z)| \cdot T_* \cdot (L_n + 1)^3 \, c^{L_n/2} \, \log n.$$

Divide by $|\mathcal{C}_{\max}|^2$. Since $|N_{L_n}(z)| \le |\mathcal{C}_{\max}|$,

$$p(z) = \frac{|E|}{|\mathcal{C}_{\max}|^2} \le \frac{T_* \cdot (L_n + 1)^3 \, c^{L_n/2} \, \log n}{|\mathcal{C}_{\max}|}.$$

Plugging in T_* and using $|\mathcal{C}_{\max}| \ge 0.99 \, \gamma n$ from (G1):

$$p(z) \le \frac{(L_n + 1)^4 \, (\delta_n + 2)^3 \, (\log n)^2 \, c^{\delta_n + L_n/2 + 1}}{0.99 \, \gamma \, n}.$$

Since this bound is uniform in z, it also bounds $\mathbb{P}(A) = \mathbb{E}_z[p(z)]$.

4.4 Step 4: Plugging in δ_n and L_n

The exponent of c is

$$\begin{aligned}\delta_n + \frac{L_n}{2} + 1 &\le \tfrac{1}{2}\log_c n - 9\log_c \log n - M + \tfrac{1}{2}\log_c n + \tfrac{K(\eta)}{2} + 1\\ &= \log_c n - 9\log_c \log n - M + \tfrac{K(\eta)}{2} + 1,\end{aligned}$$

so

$$c^{\delta_n + L_n/2+1} \le \frac{n \cdot c^{K(\eta)/2+1}}{(\log n)^9 \cdot c^M}.$$

Since $M \ge 2$, for n large we have $\delta_n + 2 \le \log_c n$ and $L_n + 1 \le 2\log_c n$, so

$$(L_n+1)^4(\delta_n+2)^3(\log n)^2 \le \frac{16\,(\log n)^9}{(\log c)^7}.$$

Combining,

$$\mathbb{P}(A) \le \frac{16\,(\log n)^9}{(\log c)^7} \cdot \frac{n\, c^{K(\eta)/2+1}}{(\log n)^9\, c^M} \cdot \frac{1}{0.99\,\gamma\, n} = \frac{16\,c \cdot c^{K(\eta)/2}}{0.99\,(\log c)^7\,\gamma} \cdot c^{-M}.$$

By our choice of M,

$$c^{-M} \le \frac{0.99\,\epsilon\,(\log c)^7\,\gamma}{c^{K(\eta)/2}},$$

so

$$\mathbb{P}(A) \le \frac{16\,c \cdot c^{K(\eta)/2}}{0.99\,(\log c)^7\,\gamma} \cdot \frac{0.99\,\epsilon\,(\log c)^7\,\gamma}{c^{K(\eta)/2}} = 16\,c\,\epsilon.$$

A slight sharpening of the constant inside M (to absorb the factor $16c$) yields $\mathbb{P}(A) \le \epsilon/12$; the argument is otherwise identical. Tripling over the three sides and substituting back into (1),

$$\mathbb{P}(B_{\delta_n}) \le \tfrac{3\epsilon}{4} + 3 \cdot \tfrac{\epsilon}{12} = \epsilon.$$

4.5 Step 5: Conclusion

For any good G_n, $\mathbb{P}(B_{\delta_n}(\Delta)) \le \epsilon$, i.e.,

$$\mathbb{P}(\mathrm{slim}(\Delta) > \delta_n) \ge \mathbb{P}(\neg B_{\delta_n}(\Delta)) \ge 1-\epsilon.$$

Since $\mathbb{P}(G_n \text{ good}) > 1-\epsilon$ for n large,

$$\mathbb{P}_{G_n}\Big(\mathbb{P}(\mathrm{slim}(\Delta) > \delta_n) > 1-\epsilon\Big) > 1-\epsilon. \quad \square$$

Corollary 7. If $\delta_n = \left\lfloor \frac{1}{\log c}\left(\frac{1}{2}\log n - 9\log\log n\right)\right\rfloor - g(n)$ with $g(n) \to \infty$, then for every fixed $\epsilon > 0$,

$$\mathbb{P}_{G_n}\Big(\mathbb{P}(\mathrm{slim}(\Delta) > \delta_n) > 1-\epsilon\Big) \longrightarrow 1.$$

Proof. For fixed ϵ, Theorem 1 supplies $M(\epsilon)$. Since $g(n) \to \infty$, eventually $g(n) \ge M(\epsilon)$, and the hypothesis of the theorem applies with margin to spare. $\square$

References

[1] R. van der Hofstad, G. Hooghiemstra, and P. Van Mieghem, *Distances in random graphs with finite variance degrees*, Random Structures & Algorithms **27** (2005), 76–123.

[2] A. C. Gilbert and J.-H. Yim, *Hyperbolicity, slimness, and minsize, on average*, arXiv:2412.05746 (2024).