

Lemmas 4 and 5 and Theorem 1 used general ℓ -vertex rules, but these cannot guarantee uniqueness of the giant component: consider a rule that only adds an edge if all ℓ points come from different components, and it will merge the two smallest components in this case. Then $\ell - 1$ large components are able to grow, and they will never be merged together.

The underlying issue here is that the probability of our large components merging is 0, an issue that is remedied by forcing our rule to be merging.

Definition 1. Suppose C_1 and C_2 are distinct components with

$$|C_1|, |C_2| \geq \varepsilon n.$$

An ℓ -vertex rule is **merging** if the probability of C_1 and C_2 merging after a single step is at least ε^ℓ .

So merging rules have a lower bound on this merging probability that grows based on the size of the components, ensuring that they will always be likely to merge. We will show in Theorem 2 that merging rules have unique giant components. But before we get there, we'll need two technical lemmas about merging rules. The first is a standard half life bound.

Lemma 1. Suppose $|C_1|, |C_2| \geq \varepsilon n$. Using a merging ℓ -vertex rule, the probability that C_1 and C_2 are *not* merged after m steps is at least $e^{-\varepsilon^\ell m}$.

Corollary 1. If $|C_1|, |C_2| \geq \frac{\varepsilon n}{2\ell}$, then the probability they don't merge after $\Delta/2$ steps is at least

$$\exp\left(-\frac{\varepsilon n}{\ell^\ell k}\right).$$

Proof. By the previous lemma,

$$\mathbb{P}(C_i, C_j \text{ don't merge}) \leq \exp\left(-\left(\frac{\varepsilon}{2\ell}\right)^\ell \frac{\Delta}{2}\right) = \exp\left(-\frac{\varepsilon^\ell}{2\ell^\ell} \left\lceil \frac{2^\ell n}{\varepsilon^{\ell-1}k} \right\rceil\right) \leq \exp\left(-\frac{\varepsilon n}{\ell^\ell k}\right).$$

□

We can now prove an analogue of Lemma 4 for merging rules. Let $V_{\geq k}(m)$ be the union of all components with k or more vertices, so $|V_{\geq k}(m)| = N_{\geq k}(m)$.

Lemma 6. Suppose we have a merging ℓ -vertex rule, and set $\Delta = 2 \left\lceil \frac{2^\ell n}{\varepsilon^{\ell-1}k} \right\rceil$. With probability at least $1 - \ell e^{-cn/k}$, there is a component of $G(m + \Delta)$ with at least $N_{\geq k}(m) - \varepsilon n$ vertices from $V_{\geq k}(m)$, where $c > 0$ is a function of ε and ℓ .

Proof. Let $W = V_{\geq k}(m)$. Assume $|W| \geq \varepsilon n$, and let $\alpha = |W|/n \geq \varepsilon$. Suppose we haven't yet gotten to the point where $\ell - 1$ components together contain at least $(\alpha - \varepsilon/2)n$ vertices from W , then the probability that we choose ℓ vertices all in W and all in different components is at least $\alpha(\varepsilon/2)^{\ell-1}$:

1. Pick some v_1 , say in component C_1 . The probability that $v_1 \in W$ is α .
2. Pick some v_2 . If $\mathbb{P}(v_2 \in W - C_1) < \varepsilon/2$, then $|C_1| \geq (\alpha - \varepsilon/2)n$, so by contradiction, $\mathbb{P}(v_2 \in W - C_1) \geq \varepsilon/2$.
3. Continue picking vertices until we get to v_ℓ . If $\mathbb{P}(v_\ell \in W - C_1 - \dots - C_{\ell-1}) < \varepsilon/2$, then $|C_1 \cup \dots \cup C_{\ell-1}| \geq (\alpha - \varepsilon/2)n$, so by contradiction, $\mathbb{P}(v_\ell \in W - C_1) \geq \varepsilon/2$.

Thus the probability of picking ℓ vertices all in W and all in different components is at least $\alpha(\varepsilon/2)^{\ell-1}$. In this case, since all the components of the chosen vertices are distinct, at least 2 components must merge, so the number of components intersecting W must decrease by at least 1.

Appeal to Lemma 4.

Thus after $\Delta/2$ steps, we have (with high probability?) components $C_1, \dots, C_{\ell-1}$ that together contain at least $(\alpha - \varepsilon/2)n$ vertices of W .

Ignore any C_i with fewer than $\frac{\varepsilon n}{2\ell}$ vertices in W , then by the previous corollary, the probability that any two sufficiently large C_i merge after $\Delta/2$ steps is at least $\exp\left(-\frac{\varepsilon n}{\ell^2 k}\right)$.

Need to take this statement about *some* pair of components and make it a bound for *all* the component pairs, which the paper coincidentally doesn't do. This will show why that all components merge, giving one large component of the proper size. There's also the issue that we have two bounds: one for the first $\Delta/2$ steps and one for the last $\Delta/2$ steps). Combining these bounds should give the desired $1 - \ell \exp\left(-\frac{\varepsilon n}{\ell^2 k}\right)$.

To make this fit with the statement of the lemma, simply set $c(\varepsilon, \ell) = \varepsilon/\ell^\ell$. \square

This lemma lets us prove Theorem 2 from the paper, which essentially says that with high probability, there will be *no* time throughout our graph construction process where we have lots of vertices in equally large, yet still separate, components. Thus the bad example that I started these notes with can't happen with merging rules.

Theorem 2. Suppose we have a merging ℓ -vertex rule for some $\ell \geq 2$ that generates a random sequence of graphs $\{G(m)\}_{m \geq 0}$. For each $\varepsilon > 0$, there is a $K = K(\varepsilon, \ell)$ such that:

$$\mathbb{P}(\text{for all } m, N_{\geq K}(m) < L_1(m) + \varepsilon n) \rightarrow 1.$$

as $n \rightarrow \infty$.

Proof. I haven't worked through this entirely yet. \square