A Galton Watson Approach to Noise in Random Graphs

Elijah Blum[1], YuiTo Cheng[2], David Fischer[3], Joshua Richland[4], Yuval Peled[5]

[1]Cornell University
[2]National University of Singapore
[3]Princeton University
[4]University of California, Berkeley
[5]Hebrew University of Jerusalem

November 15, 2024

Abstract

An Erdos-Renyi random graph on n vertices contains each edge with a probability p. Many developments in this area use the Galton Watson branching process to model the neighborhood of a given vertex. We research the effect of adding noise to a random graph. To study the relationship between the original and perturbed graph, we develop a modified Galton Watson process, and prove that it locally models the structure of both graphs.

Keywords: Erdős-Rényi random graph, Noise Sensitivity, Giant Component, K-core, Galton-Watson process

Introduction

One of the most striking results in random graph theory is the emergence of a giant connected component. Pittel, Spencer and Wormald [2] found the threshold for the emergence of a non-trivial k-core in the random graph $G(n, \lambda/n)$, and the asymptotic size of the k-core above the threshold. Riordan [3] gave another proof using branching processes, and Molloy [1] gave a proof by analyzing the k-core peeling algorithm, extending the result to hypergraphs. We consider the noise sensitivity of a vertix in the k-core of a random graph, and introduce a new branching model to study noise in random graphs.

Model of Noise

We study the noise sensitivity of a vertex in the giant component of $G(n, p)$ and respectively the k-core in the supercritical regime. We introduce a model for noise in a random graph. Given $G \sim G(n, p)$ we define its noisy graph counterpart by resampling each possible edge with probability ε. Equivalently we can define G^{ε} as follows: given G and G', two independent $G(n, p)$ random graphs, then for each edge in a complete graph on the n edges, we flip an epsilon coin, with probability ε, the edge in G^ε matches the edge from G' and with probability $1-\varepsilon$ the edge in G^ε matches the edge from G. An intuitive and abusive notation expresses this as $G^\varepsilon = (1-\varepsilon)G + \varepsilon G'$.

Galton Watson Epsilon Process

The Galton-Watson process is an important to tool in studying random graphs, approximating the neighborhood of a vertex. We define a modified process to study our noisy model as follows:

Begin with a vertex v. T is a poi(c) Galton Watson branching process. T^ε is a modified version. If v has poi(c) children in T, then each child is retained in T^ε with probability $1-\varepsilon$ and poi(εc) new children are added T^ε. Thus T^ε is also a poi(c) Galton Watson branching process which is correlated with T.

Equivalently we can view these two processes as 1 branching process. Starting with vertex v, it may have three types of children, those which are in $T \cap T^\varepsilon$ denoted by $\bigcirc$, those which are in $T \backslash T^\varepsilon$, denoted by $\square$ and those which are in $T^\varepsilon \backslash T$, denoted by $\triangle$. A $\bigcirc$ has poi($(1-\varepsilon)c$) $\bigcirc$ children, poi(ε) $\square$ children and poi(ε) $\triangle$ children. A $\square$ or $\triangle$ has poi(c) children of its own type.

Note that the number of children of each type is independent, and that $\triangle$ and $\square$ are identically distributed. Further observe that $\triangle$ and $\square$ vertices can be viewed as the start of independent poi(c) Galton-Watson processes, while the $\bigcirc$ vertices can each be viewed as independent processes identical to the root.

Results

theorem 1. *The probability that a vertex is in the giant component of both G and G^ε converges to the probability, $y(\varepsilon)$, that both $|T|$ and $|T^\varepsilon|$ are infinite, and $y(\varepsilon)$ satisfies the implicit formula*

$$y(\varepsilon) = (1-q)^2 \left(e^{cy(1-\varepsilon)-2c\varepsilon} - 1\right) + q^2$$

theorem 2. *For $k > 1$, the probability that a vertex is in the k-core of both G and G^ε converges to the probability, $\tilde{y}_k(\varepsilon)$, that $|T|$ and $|T^\varepsilon|$ contain a k-regular tree containing the root and $\tilde{y}_k(\varepsilon)$ satisfies the implicit formula*

$$\tilde{y}_k(\varepsilon) = 1 + \sum_{j=0}^{k-1} \mathbb{P}[poi((1-\varepsilon)cy_{k-1}) = j] \left(\mathbb{P}[poi(c(q_{k-1} - (1-\varepsilon)y_{k-1})) \geq k-j]^2 - 1\right)$$

Analysis of the Galton Watson Epsilon Process

Let T be a poi(c) Galton-Watson process, and T_ε defined as above. Each vertex of a k-ary tree has k children. Each vertex of a k regular tree has k neighbors. The only difference is the number of children of the first vertex, making quantities for k-ary tree slightly easier to calculate. Define the following

$$q \coloneqq \mathbb{P}[|T| = \infty] = \mathbb{P}[\mathrm{poi}(cq) \geq 1] = 1 - e^{-cq}$$

$$q_k \coloneqq \mathbb{P}[|T| \text{ contains a } k\text{-ary tree containing the root}] = \mathbb{P}[\mathrm{poi}(cq_k) \geq k] = 1 - e^{-cq_k} \sum_{i=0}^{k-1} \frac{(cq_k)^i}{i!}$$

$$\tilde{q}_k \coloneqq \mathbb{P}[|T| \text{ contains a } k\text{-regular tree containing the root}] = \mathbb{P}[\mathrm{poi}(cq_{k-1}) \geq k]$$

Then we calculate the following probabilities

lemma 1. *The probability, $y(\varepsilon, c)$, that both $|T|$ and $|T^\varepsilon|$ are infinite satisfies*

$$y(\varepsilon) \coloneqq \mathbb{P}[|T|, |T^\varepsilon| = \infty] = (1-q)^2 \left(e^{cy(1-\varepsilon)-2c\varepsilon} - 1\right) + q^2$$

Proof. Consider instead the probability, x that both trees are finite.

$$x \coloneqq \mathbb{P}(|T| < \infty, |T^\varepsilon| < \infty)$$

We consider the three types of children, those shared $\bigcirc$, those belonging only to $|T|$, $\square$, and those belonging only to $|T|^\varepsilon$, $\triangle$.

Then x is equal to the probability that all of the processes starting from the children of the root die out eventually. Recall our observations about the process: that the number of children of each type is independent, and that $\triangle$ and $\square$ are identically distributed. Also that $\triangle$ and $\square$ vertices can be viewed as the start of independent poi(c) Galton-Watson processes, while the $\bigcirc$ vertices can each be viewed as independent processes identical to the root. So,

$$\begin{aligned} x &= \mathbb{P}(\text{The } \bigcirc \text{ children of the root `die out'})\mathbb{P}(\text{The } \triangle \text{ children of the root `die out'})^2 \\ &= \mathbb{P}(\mathrm{poi}(c(1-\varepsilon)(1-x)) = 0)\mathbb{P}(\mathrm{poi}(c\varepsilon(1-q)) = 0)^2 \\ &= \exp(-c(1-\varepsilon)(1-x))\exp(-2c\varepsilon(1-q)) \end{aligned}$$

Using our expression for q above. Then consider the possibilities, each entry in the table is a probability:

	$\lvert T^\varepsilon\rvert$ finite	$\lvert T^\varepsilon\rvert = \infty$	sum
$\lvert T\rvert$ finite	x		$1-q$
$\lvert T\rvert = \infty$		y	q
sum	$1-q$	q	1

Then it is clear that

$$y = x - (1 - 2q) \tag{1}$$

Substituting into our previous expression and simplifying, we obtain

$$\begin{aligned} y &= \exp(cy(1-\varepsilon) - 2c\varepsilon - 2cq) - 1 + 2q \\ &= (1-q)^2 \left(e^{cy(1-\varepsilon)-2c\varepsilon} - 1\right) + q^2 \end{aligned}$$

$\square$

lemma 2. *The probability, $y_k(\varepsilon) = y_k$, that $|T|$ and $|T^\varepsilon|$ both contain a k-ary tree containing the root for $k > 1$ satisfies*

$$y_k(\varepsilon) = 1 + \sum_{j=0}^{k-1} \mathbb{P}[poi((1-\varepsilon)cy_k) = j] \left(\mathbb{P}[poi(c(q_k - (1-\varepsilon)y_k)) \geq k-j]^2 - 1\right)$$

Proof. Let $A(T)$ be the indicator on the event T contains a k-ary tree containing the root. Then consider the possibilities

	$A(T^\varepsilon) = 0$	$A(T^\varepsilon) = 1$	sum
$A(T) = 0$	a_{00}	a_{01}	$1 - q_k$
$A(T) = 1$	a_{10}	y_k	q_k
sum	$1 - q_k$	q_k	1

Consider the children of the root which contain a k-ary tree (containing itself) in both T and T^ε. These can only be the $\bigcirc$ children. So there are $j \sim \mathrm{Bin}(\mathrm{poi}((1-\varepsilon)c), y_k)$ such children.

Either j is less than k or at least k. If j is less than k then there is a k-ary tree in both $|T|$ and $|T^\varepsilon|$ containing the root iff there are enough children in each of T and T^ε which each contain a k-ary tree (containing themselves) in only one of the trees. In $|T|$ these children can come from type $\triangle$ or $\bigcirc$. The number of type $\triangle$ is distributed $\mathrm{Bin}(\mathrm{poi}(\varepsilon c), q_k)$ and the number of type $\bigcirc$ is $\mathrm{Bin}(\mathrm{poi}((1-\varepsilon)c), q_k - y_k)$. In total there are $\mathrm{poi}(c(1-\varepsilon)(q_k - y_k)) + \mathrm{poi}((\varepsilon c)q_k) = \mathrm{poi}(c(q_k - (1-\varepsilon)y_k))$. Since T and T^ε are symmetric, it is the same for T^ε. So,

$$y_k(\varepsilon) = \mathbb{P}[\mathrm{poi}((1-\varepsilon)cy_k) \geq k] + \sum_{j=0}^{k-1} \mathbb{P}[\mathrm{poi}((1-\varepsilon)cy_k) = j]\mathbb{P}[\mathrm{poi}(c(q_k - (1-\varepsilon)y_k)) \geq k-j]^2$$

Some alternative forms are listed below

$$= 1 + \sum_{j=0}^{k-1} \mathbb{P}[\mathrm{poi}((1-\varepsilon)cy_k) = j] \left(\mathbb{P}[\mathrm{poi}(c(q_k - (1-\varepsilon)y_k)) \geq k-j]^2 - 1\right)$$

$$= 1 + e^{-(1-\varepsilon)cy_k} \sum_{j=0}^{k-1} \frac{((1-\varepsilon)cy_k)^j}{j} \left[\left(1 - e^{-c(q_k-(1-\varepsilon)y_k)} \sum_{m=0}^{k-j-1} \frac{(c(q_k - (1-\varepsilon)y_k))^m}{m!}\right)^2 + 1\right]$$

$\square$

It is immediate from lemma 2 that

lemma 3. *The probability, $\tilde{y}_k(\varepsilon)$, that $|T|$ and $|T^\varepsilon|$ contain a k-regular tree containing the root for $k > 1$ satisfies*

$$\tilde{y}_k(\varepsilon) = 1 + \sum_{j=0}^{k-1} \mathbb{P}[poi((1-\varepsilon)cy_{k-1}) = j] \left(\mathbb{P}[poi(c(q_{k-1} - (1-\varepsilon)y_{k-1})) \geq k-j]^2 - 1\right)$$

Proof. This calculation is almost the same as the previous one, only now we need to have at least k children each having $k-1$-ary trees. $\square$

A breadth first search model, proof for giant component

Now we will show that our Epsilon Galton-Watson process is a good local model for the neighborhood of a vertex in both the original graph G and the noisy graph G^ε. In particular, we'll prove the first of two convergence theorem, the next section is devoted to the second convergence theorem generalizing this one to the k-core.

theorem 3. *The probability that a vertex is in the giant component of both G and G^ε converges to the probability, $y(\varepsilon)$, that both $|T|$ and $|T^\varepsilon|$ are infinite*

Proof. We define a branching model which is constistent with doing a breadth first search (BFS) on G and G^ε up to minor modifications.

At time zero, we begin with the root vertex v which is of type $\bigcirc$. The following chart is a helpful tool to keep track of the vertices we have explored and those waiting to be explored.

time = 0	$\square$	$\bigcirc$	$\triangle$
explored	0	0	0
unexplored	0	1	0

Between $t = 0$ and $t = 1$ we ask what children v has, we add those children to the queue of unexplored vertices and move v from the unexplored to the explored. We define the helper variable E_t to be the number of explored vertices at time t (the sum of the first row in the table). For example, $E_0 = 0$ and $E_1 = 1$.

v has equal probability of being connected to any other vertex so it has $\mathrm{Bin}(n-1, p)$ children in G, we then flip an ε coin and then a p coin for each of these edges to determine which are in $G \cap G^\varepsilon$ or just in $G \backslash G^\varepsilon$. We then flip an epsilon coin and then a p on the rest of the edges to determine which of these edges are in $G^\varepsilon \backslash G$. Equivalently it has distribution $\mathrm{mult}(n-1, p_\square = p\varepsilon(1-p), p_\bigcirc = p(\varepsilon p + 1 - \varepsilon), p_\triangle = (1-p)\varepsilon p)$.

Now each child has children in BFS order. An arbitrary unexplored vertex may have children among the $n - 1 - E_t$ unexplored vertices since we have not checked the edge between itself and the other unexplored vertices, but it may not have edges between itself and the explored vertices since we already considered this edge when we explored it. In order to simplify the analysis considerably, we'll want to compare this to a BFS search, so if we ever form an edge between two vertices which are already in our BFS tree we immediately halt the process just before this edge is formed. We call this case 'IE' for internal edge.

The distribution of the children of a given vertex is thus as follows:

$$\mathrm{CHILDREN}(\bigcirc, t) \sim \mathrm{mult}(n - 1 - E_t, p_\square = p\varepsilon(1-p), p_\bigcirc = p(\varepsilon p + 1 - \varepsilon), p_\triangle = (1-p)\varepsilon p)$$
$$\mathrm{CHILDREN}(\square, t), \mathrm{CHILDREN}(\triangle, t) \sim \mathrm{Bin}(n - 1 - E_t, p\varepsilon(1-p))$$

Going through each vertex in a BFS manner, we explore vertices until either 'IE' (we are about to connect to a vertex which is already in our tree), or either $\#\square + \#\bigcirc \geq k$ or symmetrically $\#\triangle + \#\bigcirc \geq k$. Wlog assume we reached $\#\square + \#\bigcirc \geq k$ first. Then we stop exploring $\square$ vertices, we relabel the unexplored $\bigcirc$ vertices as unexplored $\triangle$ vertices, and we continue the BFS branching adding one vertex at a time until either $\#\bigcirc + \#\triangle = k$ or we run out of vertices to explore or we create an internal edge (IE).

Note that we will always terminate with less than $2k + \mathrm{Bin}(n, p)$ vertices in our tree.

Thus our BFS can terminate in three ways, (IE) we stop before forming an internal edge, (DIE) we run out of vertices to explore at some point in the process, or (WIN) both G and G^ε have size at least k.

Thus

$$\begin{aligned}\mathbb{P}[C(G,v), C(G^\varepsilon, v) \geq k] =& \mathbb{P}[C(G,v), C(G^\varepsilon, v) \geq k | IE]\mathbb{P}[IE] \\ &+ \mathbb{P}[C(G,v), C(G^\varepsilon, v) \geq k | DIE]\mathbb{P}[DIE] \\ &+ \mathbb{P}[C(G,v), C(G^\varepsilon, v) \geq k | WIN]\mathbb{P}[WIN]\end{aligned}$$

Clearly if the BFS dies then the components are not both at least size k, and if we win then they necessarily are. Hence

$$\mathbb{P}[C(G,v), C(G^\varepsilon, v) \geq k] = \mathbb{P}[C(G,v), C(G^\varepsilon, v) \geq k | IE]\mathbb{P}[IE] + \mathbb{P}[WIN]$$

Next we prove that $\mathbb{P}[IE] \overset{n\to\infty}{\to} 0$

Recall there are less than $2k + \mathrm{Bin}(n, p)$ vertices in our tree by termination, and we explore no more than $2k$ vertices so there are no more than $2k(2k+\mathrm{Bin}(n, p))$ edges. The probability that $\mathrm{Bin}(n, p) > k$ vanishes in the limit, so consider $6k^2$. Then the probability that any of the $6k^2$ edges will be connected in $G \cup G^\varepsilon$ is certainly less than $2p$.

$$\mathbb{P}[IE] \leq \mathbb{P}[\mathrm{Bin}(6k^2, 2p) > 0] + \mathbb{P}[\mathrm{Bin}(n, 2p) > k]$$

By a simple Markov bound

$$\mathbb{P}[\mathrm{Bin}(6k^2, 2p) \geq 1] \leq 12k^2 p$$

Then as long as k is $o(\sqrt{n})$, this converges to 0 (we are in the regime where p is $\Theta(n^{-1})$).

For the latter term, again we can use a simple Markov bound

$$\mathbb{P}[\mathrm{Bin}(n, p) > k] \leq \frac{1}{k} np \tag{2}$$

which will converge to 0 as long as $k \to \infty$ as $n \to \infty$.

Thus we've shown that

$$\mathbb{P}[C(G,v), C(G^\varepsilon, v) \geq k] = \mathbb{P}[WIN] + o(1).$$

Now we will notice that we can slightly modify this branching process to give lower and upper bounds on this probability. Specifically if instead of drawing from the same set of vertices and subtracting the explored vertices at each step, we may draw from a fresh set of $n-1$ vertices each time to get an upper bound, or draw from $n-3k$ vertices for the lower bound. We'll denote the trees from the process with fresh $n-1$ vertices each time by $T_{\mathrm{Bin}(n-1)}, T^{\varepsilon}_{\mathrm{Bin}(n-1)}$, and the trees from the $n-3k$ process similarly. Finally observe that these two processes clearly converge to the $T =: T_{\mathrm{poi}}$, $T^{\varepsilon} =: T^{\varepsilon}_{\mathrm{poi}}$ from the Epsilon Galton Watson process described in the introduction. In symbols:

$$\mathbb{P}[|T_{\mathrm{Bin}(n-3k)}|, |T^{\varepsilon}_{\mathrm{Bin}(n-3k)}| \geq k] \leq \mathbb{P}[WIN] = \mathbb{P}[C(G,v), C(G^{\varepsilon}, v) \geq k] + o(1)$$

and

$$\mathbb{P}[C(G,v), C(G^{\varepsilon}, v) \geq k] + o(1) = \mathbb{P}[WIN] \leq \mathbb{P}[|T_{\mathrm{Bin}(n-1)}|, |T^{\varepsilon}_{\mathrm{Bin}(n-1)}| \geq k]$$

But we also have that

$$\mathbb{P}[|T_{\mathrm{Bin}(n-3k)}|, |T^{\varepsilon}_{\mathrm{Bin}(n-3k)}| \geq k] + o(1) = \mathbb{P}[|T_{\mathrm{poi}}|, |T^{\varepsilon}_{\mathrm{poi}}| \geq k]$$
$$\mathbb{P}[|T_{\mathrm{Bin}(n-1)}|, |T^{\varepsilon}_{\mathrm{Bin}(n-1)}| \geq k] + o(1) = \mathbb{P}[|T_{\mathrm{poi}}|, |T^{\varepsilon}_{\mathrm{poi}}| \geq k]$$

So

$$\mathbb{P}[|T_{\mathrm{poi}}|, |T^{\varepsilon}_{\mathrm{poi}}| \geq k] \leq \mathbb{P}[|C(G,v)|/k, |C(G^{\varepsilon}, v)|/k \geq 1] + o(1) \leq \mathbb{P}[|T_{\mathrm{poi}}|, |T^{\varepsilon}_{\mathrm{poi}}| \geq k]$$

Taking $k = \log n$, and taking the limit we have

$$\mathbb{P}[|C(G,v)|/k, |C(G^{\varepsilon}, v)|/k \geq 1] \to \mathbb{P}[|T_{\mathrm{poi}}|, |T^{\varepsilon}_{\mathrm{poi}}| = \infty]$$

Moreover by the classical result of Erdős and Rényi, the probability of an akward component, i.e. of intermediate size, is vanishing, for example see [4]. So the probability on the left is

$$\mathbb{P}[|C(G,v)|/k, |C(G^{\varepsilon}, v)|/k \geq 1] = o(1) + P[v \in C_1(G), C_1(G^{\varepsilon})]$$

where C_1 is the largest component of G.

□

Proof for k-ary tree

This theorem will generalize the previous one to the case of the k-core.

theorem 4. *For $k > 1$, the probability that a vertex is in the k-core of both G and G^{ε} converges to the probability, $\tilde{y}_k(\varepsilon)$, that $|T|$ and $|T^{\varepsilon}|$ contain a k-regular tree containing the root*

Proof. The proof relies on peeling algorithms for finding the k-core, CORE and CORE2. CORE 'peels' i.e. deletes all vertices with degree less than k and repeats until there are none left. CORE2, introduced by Molloy [1], runs t iterations on a tree, for each $i, 1 \leq i \leq t-1$ it only looks at the vertices at distance t from the root and deletes the ones which have less than $k-1$ remaining children and for $i = t$ it deletes the root iff it has fewer than k

remaining neighbors. It is easy to see that the root survives t peels in CORE2 iff it survives t peels in CORE.

Molloy [1] defines

$$
\begin{aligned}
\rho_0 &= 1 \\
\rho_i &= P(\mathrm{poi}(\rho_{i-1}c) \geq k-1) \\
\lambda_i &= P(\mathrm{poi}(\rho_{i-1}c) \geq k)
\end{aligned}
$$

and shows that for any constant t the probability that a vertex in a $G(n,p)$ graph survives t rounds of peeling is $\lambda_t + o(1)$

We define recursively the probability that a vertex survives in both T and T^{ε} after i peels.

$$
\begin{aligned}
\tilde{\rho}_0 &= 1 \\
\tilde{\rho}_i &= \mathbb{P}[X_{i-1} + \mathrm{poi}((1-\varepsilon)c(\rho_{i-1} - \tilde{\rho}_{i-1})) + \mathrm{poi}(\varepsilon c\rho_{i-1}) \geq k-1, \\
&\quad X_{i-1} + \mathrm{poi}((1-\varepsilon)c(\rho_{i-1} - \tilde{\rho}_{i-1})) + \mathrm{poi}(\varepsilon c\rho_{i-1}) \geq k-1] \\
\tilde{\lambda}_i &= \mathbb{P}[X_{i-1} + \mathrm{poi}((1-\varepsilon)c(\rho_{i-1} - \tilde{\rho}_{i-1})) + \mathrm{poi}(\varepsilon c\rho_{i-1}) \geq k, \\
&\quad X_{i-1} + \mathrm{poi}((1-\varepsilon)c(\rho_{i-1} - \tilde{\rho}_{i-1})) + \mathrm{poi}(\varepsilon c\rho_{i-1}) \geq k]
\end{aligned}
$$

where $X_{i-1} \sim \mathrm{poi}((1-\varepsilon)c\tilde{\rho}_{i-1})$ and the random variables are independent.

X_{i-1} is the number of circle children of the vertex, v, which survive $i-1$ peels in both T and T^{ε}, $\mathrm{poi}((1-\varepsilon)c(\rho_{i-1} - \tilde{\rho}_{i-1}))$ is the number of circle children of v which survive $i-1$ peels in that tree but not the other, and $\mathrm{poi}(\varepsilon c\rho_{i-1})$ is the number of triangle (resp. square) children of v which survive $i-1$ peels in that tree.

lemma 4. *For any constant t, the probability that the root survives after t rounds of peeling in both G and G_{ε} is $\tilde{\lambda}_t + o(1)$*

Proof. As in Molloy, let D_i (resp. D_i^{ε}) be the set of vertices of distance i from v in G (resp. G^{ε}). Let E_2 be the event that fewer than $\log^2 n$ vertices are within distance t of v in each of G and G^{ε} and none of these vertices lie in a cycle of length less than $2t+1$ in the union $G \cup G^{\varepsilon}$. If E_2 holds then the t neighborhood of v is a tree in $G \cup G^{\varepsilon}$, with root v, and size less than $2\log^2 n$.

Recall the CORE2 algorithm, which at each iteration behaves like CORE except it only looks at the vertices in D_{t-i} for $1 \leq i \leq t-1$. If E_2 holds then v remains after t iterations of CORE2 in both G and G_{ε} iff v remains after t iterations of CORE in both G and G^{ε}.

Now we prove by induction on i, for each $0 \leq i \leq t-1$ that a particular vertex $u \in D_{t-i} \cap D_{t-i}^{\varepsilon}$ survives CORE2 in both G and G^{ε} with probability $\tilde{\rho}_i + o(1)$. The base case $i = 0$ is trivial, since no vertices are removed. Assuming the statement is true for $i-1$, then expose $D_0, \ldots, D_{t-i}$, and $D_0^{\varepsilon}, \ldots, D_{t-i}^{\varepsilon}$ and consider $u \in D_{t-i} \cap D_{t-i}^{\varepsilon}$. Suppose $x = |D_0 \cup \cdots \cup D_{t-1}|, x^{\varepsilon} = |D_0^{\varepsilon} \cup \cdots \cup D_{t-1}^{\varepsilon}|$ and $y = |D_0 \cup \cdots \cup D_{t-1} \cup D_0^{\varepsilon} \cup \cdots \cup D_{t-1}^{\varepsilon}|$, and if E_2 holds then $x, x^{\varepsilon} < \log^2 n$ and $y < 2\log^2 n$.

Now we consider the child edges of u which remain after $i-1$ iterations of CORE2 in each of G and G^{ε}. u has children with distribution

$$\mathrm{mult}(n-y, p_{\square} = p\varepsilon(1-p), p_{\bigcirc} = p(\varepsilon p + 1 - \varepsilon), p_{\triangle} = (1-p)\varepsilon p)$$

The probability (up to $o(1)$) that each of these vertices has of surviving the $i-1$ peels in either tree is ρ_{i-1} for a $\triangle$ or $\square$, and $\rho_{i-1} - \tilde{\rho}_{i-1}$ for a $\bigcirc$, and the probability that a circle survives both is $\tilde{\rho}_{i-1}$.

So the surviving children distribution is

$$\begin{gathered}(S_{\square}, S_{\bigcirc,\square}, S_{\bigcirc,\bigcirc}, S_{\bigcirc,\triangle}, S_{\triangle}) \sim \\ \mathrm{mult}(n-y, p_{\square} = \rho_{i-1} p\varepsilon(1-p), p_{\bigcirc,\square} = (\rho_{i-1} - \tilde{\rho}_{i-1}) p(\varepsilon p + 1 - \varepsilon), \\ p_{\bigcirc,\bigcirc} = \tilde{\rho}_{i-1} p(\varepsilon p + 1 - \varepsilon), p_{\bigcirc,\triangle} = p_{\bigcirc,\square}, p_{\triangle} = p_{\square})\end{gathered}$$

Which converges to the following independent poissons,

$$\begin{gathered}(S_{\square}, S_{\bigcirc,\square}, S_{\bigcirc,\bigcirc}, S_{\bigcirc,\triangle}, S_{\triangle}) \to \\ (\mathrm{poi}(\rho_{i-1} c\varepsilon), \mathrm{poi}((\rho_{i-1} - \tilde{\rho}_{i-1}) c(1-\varepsilon)), \mathrm{poi}(\tilde{\rho}_{i-1} c(1-\varepsilon)), \mathrm{poi}((\rho_{i-1} - \tilde{\rho}_{i-1}) c(1-\varepsilon)), \mathrm{poi}(\rho_{i-1} c\varepsilon))\end{gathered}$$

Note that the probability that a particular vertex survives in this asymtotic distribution is exactly $\tilde{\rho}_i$, completing the inductive step. The same analysis shows v survives the final iteration of CORE2 iff it has at least k child edges that remain after $t-1$ iterations. The same analysis shows that this happens with probability $\tilde{\lambda}_i + o(1)$. Since $P(E_2^c) = o(1)$, v survives t rounds of CORE with probability $\tilde{\lambda}_i + o(1)$. $\square$

... $\square$

Analysis of the solutions

We have so far given recursive formulas for probabilities $y(\varepsilon)$ and $\tilde{y}_k(\varepsilon)$. In this section we prove that they define unique solutions, and the surprising fact that the solutions are continuous in ε.

lemma 5. *$y(\varepsilon)$ given by the implicit formula*

$$y(\varepsilon) = (1-q)^2 \left(e^{cy(1-\varepsilon) - 2c\varepsilon} - 1\right) + q^2$$

has a uniqe solution.

lemma 6. *$\tilde{y}_k(\varepsilon)$ given by the implicit formula*

$$\tilde{y}_k(\varepsilon) = 1 + \sum_{j=0}^{k-1} \mathbb{P}[poi((1-\varepsilon)cy_{k-1}) = j] \left(\mathbb{P}[poi(c(q_{k-1} - (1-\varepsilon)y_{k-1})) \geq k-j]^2 - 1\right)$$

has a unique solution

theorem 5. *$y(\varepsilon)$ is continuous in ε*

theorem 6. *$\tilde{y}_k(\varepsilon)$ is continuous in ε*

References

[1] Molloy, M. Cores in random hypergraphs and Boolean formulas. *Wiley InterScience* (2003). doi:10.1002/rsa.20061

[2] Pittel, B., Spencer, J., & Wormald, N. (1996). Sudden emergence of a giant k-core in a random graph. *Journal of Combinatorial Theory*. Series B, 67(1), 111-151. doi.org/10.1006/jctb.1996.0036

[3] Riordan O. The k-Core and Branching Processes. *Combinatorics, Probability and Computing*. 2008;17(1):111-136. doi:10.1017/S0963548307008589

[4] Noga Alon and Joel H. Spencer. 2016. The Probabilistic Method (4th. ed.). Wiley Publishing.