Stockholms universitet

Results on The Dynamic Erdős-Rényi Graph — The Critical Case

Sebastian Rosengren

Masteruppsats i matematisk statistik
Master Thesis in Mathematical Statistics

Masteruppsats 2015:11
Matematisk statistik
December 2015

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Mathematical Statistics
Stockholm University
Master Thesis **2015:11**
http://www.math.su.se

Results on The Dynamic Erdős-Rényi Graph — The Critical Case

Sebastian Rosengren*

December 2015

Abstract

The dynamic Erdős-Rényi graph is a natural extension of an Erdős-Rényi graph, in which one starts with n vertices and 0 edges and—independently for each vertex pair—add and remove edges according to a birth-death process. We shall study the critical version of such a model where the birth and death rates are chosen in such a way that the stationary distribution of the dynamic graph equals that of a critical Erdős-Rényi graph.

In studying such a model we present two main results, the first being on how long it takes for the dynamic graph to reach stationarity if it starts with 0 edges. We give an explicit expression for this time, as well as proving that this is the fastest time to reach stationarity.

The second result is regarding how the size of the largest component evolves through time. Mainly we give a lower bound for the probability $P(C(t) > \epsilon \cdot n)$ where $C(t)$ is the size of the largest component in the interval $[0, t]$, and $\epsilon \in (0, 1)$

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: rosengren@math.su.se. Supervisor: Pieter Trapman.

Acknowledgments

I would like to take this opportunity to extend my sincere gratitude towards my supervisor Dr. Pieter Trapman. It is important to know that this simply could not have been done without him. It has been an absolute pleasure working with the problems posed, and I am looking forward to future work in the subject area.

Contents

Contents

1 Introduction

The Erdős-Rényi graph is a well-studied model for random graphs, were one start with n vertices and then either add edges independently between vertex pairs with some probability; or letting the graph have k edges and then choosing uniformly from all possible edge configurations.

In this thesis we study a natural extension of such a model — the dynamic Erdős-Rényi graph. Starting with n vertices and 0 edges we—independently for each vertex pair—add and remove edges according to a birth-death process. We choose the birth and death rates of these processes in such a way that the stationary distribution of the dynamic graph equals that of a critical Erdős-Rényi graph.

The first main result is on how long it takes for the dynamic graph to reach stationarity if it starts with 0 edges. We give an explicit expression for this time, as well as proving that this is the fastest time to reach stationarity. It turns out that this time is distributed as the maximum of a $\binom{n}{2}$ number of independent $\mathrm{Exp}(\beta \cdot (1 + \frac{1}{n-1}))$-distributed random variables, where $\beta > 0$ is the rate at which we remove an edge, if an edge is present between a vertex pair.

Often of interest when studying random graphs is the size of its largest component, since this tend to have interesting practical implications. Mainly we shall study how long it takes for the largest component to reach size $\epsilon \cdot n$ for fixed $\epsilon \in (0, 1)$.
The main result will be a lower bound for $P(C(t) > \epsilon \cdot n)$ where $C(t)$ is the size of the largest component in the interval $[0, t]$. In doing so we shall use the intimate relation between the size of the largest component of an Erdős-Rényi graph and the number of edges present.

2 The Model

In order to define the dynamic Erdős-Rényi graph we must first define its two most basic building block, the two standard versions of the Erdős-Rényi graph.

2.1 The Erdős-Rényi Graph

We begin by stating the mathematical definition if a graph.

Definition 2.1. *A graph is an ordered pair $G = (V, E)$ where V is a vertex set, e.g. $\{1, 2, ..., n\}$, and $E \subset V \times V$ is the set of edges between the vertices.*

Definition 2.2. *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $(S, \mathcal{S})$ be a measurable space, where S is some set of possible graphs on some vertex set V, and $\mathcal{S}$ is some σ-algebra on S.*
A random graph is a random element $\mathcal{G} : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$

We now have the basic framework to define what is meant by a Erdős-Rényi graph.

Definition 2.3. *Let $V = \{1, 2, \ldots, n\}$ be a labeled vertex set.*

1. *$\mathcal{G}(n, M)$ is called an Erdős-Rényi random graph with n vertices and M edges if it is a random graph taking values in S, the set of all graphs with n vertices and M edges, with equal probability. I.e. each graph configuration in this model has the same probability, $p = \frac{1}{\binom{N}{M}}$ where $N = \binom{n}{2}$*

2. *$\mathcal{G}(n, p)$ is called an Erdős-Rényi random graph with n vertices and edge probability p if it is a random graph taking values in S, the set of all possible graphs with n vertices, where the edges between vertices are chosen independently and with probability p.*

We call the graph critical if $M = \frac{n}{2}$ or $p = \frac{1}{n}$.

Remark. *Both these models will come in handy; the term Erdős-Rényi graph is used interchangeably to describe both models, where we use context to make clear which model is referred to.*

Furthermore, the reason that the graph is called critical is because the expected number of edges of a randomly selected vertex has is equal to 1, and this acts as a threshold point for the size of the largest component. Namely if D is the number of edges a randomly selected vertex has, then if $E(D) < 1$ the fraction of vertices in the largest component converges in probability to 0, and if $E(D) > 1$ the fraction of vertices in the largest component converges in probability to a positive number.

2.2 The Dynamic Erdős-Rényi Graph

A Dynamic Erdős-Rényi graph is an Erdős-Rényi graph which evolves dynamically through time, where edges are added and removed according to some specified birth-death process. We will study a specific flavor of this process — the critical version, i.e. we choose the birth-death processes in such a way that the stationary distribution of the dynamic graph equals that of a critical Erdős-Rényi graph.
We begin this section with an informal definition of the process.

Definition 2.4. *Let $G = \{G(t);\ t \in [0, \infty)\}$ be a family of random graphs, i.e. $G(t)$ is a random graph for every $t \in [0, \infty)$. Let $\beta > 0$ and $\alpha > 0$. G is called a Dynamic Erdős-Rényi graph if,*

1. *The number of vertices is fixed at n*
2. *Independently for each vertex pair if no edge is present an edge is added after an Exp($\frac{\beta}{n-1}$)-distributed time.*
3. *Independently for each vertex pair if a edge is present that edge is removed after an Exp(α)-distributed time.*

The dynamic Erdős-Rényi graph is called critical if $\beta = \alpha$.

Remark. *In this report, we start with 0 edges present.*

Definition 2.4 will be the working definition but in order to make this more rigorous we also give the formal definition.

Definition 2.5. *Let $V = \{1, 2, \ldots, n\}$ be a labeled vertex set. Let $E_{ij} = \{E_{ij}(t)\},\ i < j \in V$ be independent birth-death (see Definition 3.2) processes on $\{0, 1\}$, starting in 0, with birth parameter $\lambda = \frac{\beta}{n-1}$ and death parameter $\mu = \alpha$. Furthermore let $G = \{G(t),\ t \in [0, \infty)\}$ be a family of graphs where $G(t) = (V, f(E(t)))$ where $E(t) = (E_{1,2}(t), \ldots, E_{n-1,n}(t))$ and f is a function $\{0, 1\}^n \to V \times V$ with $f(x_{1,2}, \ldots, x_{n-1,n}) = \{(i, j) \in V \times V;\ x_{ij} = 1\}$*
Finally, if $\beta = \alpha$ we call G a critical dynamic Erdős-Rényi graph.
Since f is a 1-to-1 function we may also take $G(t) = (E_{1,2}(t), \ldots, E_{n-1,n}(t))$

Remark. *The fact that we can take $G(t) = (E_{1,2}(t), \ldots, E_{n-1,n}(t))$ will be a very useful one and is well noted.*
Throughout the report we will call $E_{ij}(t)$ the edge processes.

3 Markov Interlude

There is a close connection between the dynamic Erdős-Rényi graph and Markov processes—namely, it is an ergodic Markov process. This fact shall prove very useful, and will be exploited throughout the report.
This section is very technical, and may be skipped if one is willing to accept that the dynamic graph and its underlying edge processes are ergodic Markov chains, and that the stationary distribution of the dynamic graph is that of a critical Erdős-Rényi graph.

We begin by recalling the definition of a Markov process.

Definition 3.1. *Let $X = \{X(t),\ t \geq 0\}$ be a stochastic process, where $X(t) : (\Omega, \mathcal{F}) \to\ (S, \mathcal{S})$ is a random variable for each $t \geq 0$ with respect to the measure $\mathbb{P}$, and S is a countable set with corresponding σ-algebra $\mathcal{S}$.
Secondly—for technical reasons—we assume that $X(\cdot, \omega)$ is right-continuous, in t, for all $\omega \in \Omega$. We say that X is a Markov process if $\forall i, j, i_k, \dots, i_0 \in S$ and $\forall t > s > t_k > \dots t_0 \geq 0$ we have*

$$P(X(t) = j|X(s) = i, X(t_k) = i_k, \dots X(t_0) = i_0) = P(X(t) = j|X(s) = i)$$

whenever these probabilities exists. Furthermore, if we also have

$$P(X(t) = j|X(s) = i) = P(X(t-s) = j|X(0) = i)$$

the process is called homogeneous.

The process is called ergodic if,

1. *The process is irreducible, i.e. starting in state i the probability to eventually reach state j is larger than 0 for all $i, j \in S$.*

2. *The process is positive recurrent, i.e. starting in state i the expected time to return to i is finite.*

Remark. *All standard Markov processes are homogeneous, and it will be assumed throughout that whenever we are talking about Markov processes we are talking about homogeneous Markov processes.*

As previously stated the processes describing if an edge is present between vertex pairs—the edge processes—are birth-death processes. We therefore recall the definition of a birth-death process.

Definition 3.2. *Let X be a Markov process taking values in $S = \{0, 1, \dots, n\}$ or $\mathbb{N}$. Let $\{\lambda_k,\ k \in S\}$ and $\{\mu_k,\ k \in S\}$ be sequences of positive real numbers called the birth and death rates. We say that X is a birth-death process if,*

1. *Each time X enters state k it stays there for an Exp($\lambda_k + \mu_k$)-distributed amount of time.*
2. *The holding times are independent random variables.*
3. *When leaving state $0 < k < n$ the process enters state k+1 with probability $p_{k,k+1} = \frac{\lambda_k}{\lambda_k+\mu_k}$ and state k-1 with probability $p_{k,k-1} = 1 - p_{k,k+1}$ independently of how long it has been in state k.*
4. *When leaving state 0 the process enters state 1 with probability 1, and when leaving state n it enters state n-1 with probability 1.*

In going forward we shall need to know that the underlying edge processes, the dynamic graph, and the process describing the number of edges present all are ergodic Markov chains. This is mainly for existence and uniqueness results. For instance, we will need to calculate the expected value for the hitting time for k edges being present—so we need to know this expectation exists, and ergodicity ensures this.

In order to prove ergodicity we shall first need two auxiliary Lemmas. The first Lemma is well-known, but we shall still state and prove it.

Lemma 3.3. *Let Y be a positive random variable, with $P(Y < \infty) = 1$. Then,*

$$E(Y) = \int_0^\infty P(Y > t)\, dt \tag{3.1}$$

In the sense that if either side exist so does the other, and if either side does not exist i.e. is equal to ∞ so is the other.

Proof. Assume $E(Y) < \infty$.

$$\begin{aligned} E(Y) &= \int Y\, dP = \int y\, dF_Y \overset{1.)}{=} \int_0^\infty \int_0^\infty 1\{y > t\}\, dt\, dF_Y \\ &\overset{2.)}{=} \int_0^\infty \int_0^\infty 1\{y > t\}\, dF_Y\, dt = \int_0^\infty P(Y > t)\, dt. \end{aligned}$$

1.) is valid since Y is positive and $P(Y < \infty) = 1$, and 2.) holds true since $1\{y > t\}$ is positive and we can then interchange integrals by Fubini's theorem, see [1, Ch. 2.15].

Assume $E(Y) = \infty$.
Define $X_n = 1\{Y \leq n\} \cdot Y$. Since $P(Y < \infty) = 1$ we have that,

$$\lim_{n \to \infty} X_n = Y, \ a.s.$$

Also $\{X_n\}$ is an increasing sequence. So by the Lebesgue monotone convergence theorem, see [1, Ch. 2.10],

$$\lim_{n \to \infty} E(X_n) = E(\lim_{n \to \infty} X_n) = E(Y).$$

We further have that $X_n \leq n$ so by previous result,

$$E(X_n) = \int_0^\infty P(X_n > t)\, dt = \int_0^n P(Y > t)\, dt.$$

We can then conclude that,

$$\lim_{n \to \infty} E(X_n) = \lim_{n \to \infty} \int_0^n P(Y > t) = \int_0^\infty P(Y > t) = \infty.$$

Hence, the two sides of (3.1) is always equal. □

Lemma 3.4. *Let X be an irreducible Markov process on a finite state space S. Let $t_0 < t_1 < \cdots < t_k$ be real numbers where $t_i - t_{i-1} = \Delta > 0$, and let B be a measurable set of S. Then for fixed k,*

$$P(X(t_k) \in B, X(t_{k-1}) \in B, \ldots, X(0) \in B) \leq \max_{m \in B} P(X(\Delta) \in B | X(0) = m)^k \tag{3.2}$$

if the conditional probability exists.

Proof. We prove the assertion with induction.
For $k = 1$,

$$\begin{aligned} P(X(t_1) \in B, X(t_0) \in B) &= \sum_{m \in B} P(X(t_1) \in B, X(t_0) = m) \\ &= \sum_{m \in B} P(X(t_1) \in B | X(t_0) = m) \cdot P(X(t_0) = m) \\ &\leq \max_{m \in B} P(X(t_1) \in B | X(t_0) = m) \cdot P(X(t_0) \in B) \\ &\leq \max_{m \in B} P(X(t_1) \in B | X(t_0) = m) = \max_{m \in B} P(X(\Delta) \in B | X(0) = m). \end{aligned}$$

In order to prove it for arbitrary k we shall need an easy consequence of the Markov property, which we state without proof. Namely, $\forall i, j, B_{i_k}, \ldots, B_{i_0} \in \mathcal{S}$ and $\forall t > s > t_k > \ldots t_0 \geq 0$ we have

$$P(X(t) = j | X(s) = i, X(t_k) \in B_{i_k}, \ldots X(t_0) = B_{i_0}) = P(X(t) = j | X(s) = i).$$

Now assume (3.2) holds for arbitrary k.

$$P(X(t_{k+1}) \in B, X(t_k) \in B, \ldots, X(t_0) \in B) = \sum_{m \in B} P(X(t_{k+1}) \in B, X(t_k) = m, \ldots, X(t_0) \in B)$$

$$= \sum_{m \in B} P(X(t_{k+1}) \in B | X(t_k) = m) \cdot P(X(t_k) = m, X(t_{k-1}) \in B, \ldots, X(t_0) \in B)$$

$$\leq \max_{m \in B} P(X(\Delta) \in B | X(0) = m) \sum_{m \in B} P(X(t_k) = m, X(t_{k-1}) \in B, \ldots, X(t_0) \in B)$$

$$= \max_{m \in B} P(X(\Delta) \in B | X(0) = m) P(X(t_k) \in B, X(t_{k-1}) \in B, \ldots, X(t_0) \in B)$$

$$\leq \max_{m \in B} P(X(\Delta) \in B | X(0) = m)^{k+1}.$$

Hence, it holds for all $k \in \mathbb{N}$.

□

The following Lemma shows that irreducible Markov chains on finite state spaces are ergodic. Since the dynamic graph lives on a finite state space, in order to prove ergodicity we need only prove that the graph is Markov and irreducible and then apply the Lemma. But first we recall the definition of a stopping time.

Definition 3.5. *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ be a filtered probability space. Let $T : \Omega \to \mathbb{R}^+ \cup \{\infty\}$ be an extended random variable. Then T is said to be a stopping time relative to the filtration $\{\mathcal{F}_t\}$ if,*

$$\{T \leq t\} \in \mathcal{F}_t, \ \forall t > 0.$$

Furthermore if $\{\mathcal{F}_t\}$ is the natural filtration for some process X, T is said to be a stopping time for X.

Lemma 3.6. *Let X be an irreducible Markov process on a finite state space S. Let $\tau_i(j) = \inf\{t > 0; \ X(t) = j, \ X(0) = i\}$, $i \neq j$, be the time it takes to go from state i to state j if the process starts in state i. Then $\tau_i(j)$ is a stopping time for X starting in state i and,*

$$E(\tau_i(j)) < \infty \text{ and } E(\tau_i(i)) < \infty$$

where $\tau_i(i)$ is the cycle time of i, i.e. the time it takes to go from state i back to i.
Therefore, all irreducible Markov processes on a finite state spaces are ergodic.

Proof. Assume X starts in i.

$$\{\tau_i(j) \leq t\} = \{\bigcup_{s \leq t} \{X(s) = j\}\} = A_t$$

where s runs over all real numbers less than or equal to t. Since by Definition 3.1 $X(\cdot, \omega)$ for all $\omega \in \Omega$ is right-continuous, and since the rationals are dense in $\mathbb{R}$ we must have that,

$$A_t = \{\bigcup_{s \leq t} \{X(s) = j\}\} = \{\bigcup_{r \leq t} \{X(r) = j\}\} = Q_t$$

where r runs over all the rational numbers less than or equal to t.
Now $\{X(r) = j\} \in \mathcal{F}_t$ for $r \leq t$ and since σ-algebras are closed under countable unions this means that $Q_t \in \mathcal{F}_t$ and that $\tau_i(j)$ is a stopping time for X—if X starts in state i—which also of course implies that $\tau_i(j)$ is measurable with regard to the probability space where $X(0) = i$.
Let $t_0 < t_1 < \cdots < t_k$ be real numbers where $t_i - t_{i-1} = \Delta > 0$, and let B be a measurable set. Then by Lemma 3.4,

$$P(X(t_k) \in B, X(t_{k-1}) \in B, \ldots, X(0) \in B) \leq \max_{m \in B} P(X(\Delta) \in B | X(0) = m)^k \tag{3.3}$$

if the conditional probability exists.
Assume $t \geq 1$ and let $t_i = i$ for $i = 0, 1, \ldots, k = \lfloor t \rfloor$ so that $\{t_i\}$ is a partition of $[0, \lfloor t \rfloor] \subset [0, t]$. Let $B = \{j\}^c$. In order to bound $P(\tau_i(j) > t)$ the following inequality is useful, which is true by (3.3)

$$\begin{aligned} P(\tau_i(j) > t) = P(A_t^c) = P(\bigcap_{s \leq t} \{X(s) \neq j\}) \leq P(\bigcap_{n=0}^{k} \{X(t_n) \neq j\}) \\ \leq \max_{m \neq j} P(X(1) \in B | X(0) = m)^k = \max_{m \neq j} P(X(1) \in B | X(0) = m)^{\lfloor t \rfloor}. \end{aligned}$$

Since X is irreducible and the state space is finite we must have that $\max_{m \neq j} P(X(1) \in B | X(0) = m) < 1$.
We also note that the inequality must hold for $t < 1$ since then $\lfloor t \rfloor = 0$, hence it holds for all t. We have,

$$P(\tau_i(j) = \infty) \leq \lim_{t \to \infty} P(\tau_i(j) > t) = 0.$$

We conclude that $P(\tau_i(j) < \infty) = 1$ and we can therefore apply Lemma 3.3.

$$\begin{aligned} E(\tau_i(j)) = \int_0^\infty P(\tau_i(j) > t)\, dt \leq \int_0^\infty \max_{m \neq j} P(X(1) \in B | X(0) = m)^{\lfloor t \rfloor}\, dt = \\ \sum_{k=0}^{\infty} \max_{m \neq j} P(X(1) \in B | X(0) = m)^k < \infty. \end{aligned}$$

Hence, $E(\tau_i(j))$ exists.

In order to show that the expected cycle time of state i exists we note that

$$\tau_i(i) = T_i + \sum_{j \neq i} 1\{i \to j\} \cdot \tau_j(i) \leq T_i + \sum_{j \neq i} \tau_j(i)$$

where T_i is the holding time in state i, and $1\{i \to j\}$ is the indicator variable that the process moves from i to j. We see that $\tau_i(i)$ is measurable, as it is the sum of measurable random variables. The upper bound also implies that $E(\tau_i(i)) < \infty$ since the expectations on the right-hand side exists. This is by the Lebesgue dominated convergence theorem, see [3, Ch. 2]. □

As the dynamic graph is built up by the edge process we shall need some basic results on them. The following Lemma implies that the stationary distribution of the critical dynamical Erdős-Rényi graph equals that of a critical Erdős-Rényi graph.

Lemma 3.7. *The edge processes* $\{E_{ij}(t),\ t \geq 0\}$ *are ergodic Markov processes on* $\{0,1\}$*, with probability transition functions equal to,*

$$\begin{aligned} p_{01}(t) &= \frac{1}{n} \cdot (1 - e^{-\beta \cdot (1+\frac{1}{n+1}) \cdot t}) \\ p_{11}(t) &= \frac{1}{n}(1 + (n-1)e^{-(1+\frac{1}{n-1})\beta t}) \end{aligned}$$

and stationary distribution equal to,

$$\begin{aligned} \pi(1) &= \frac{1}{n} \\ \pi(0) &= 1 - \frac{1}{n} \end{aligned}$$

Proof. By definition $E_{ij}(t)$ is a birth-death process on $\{0,1\}$, with birth rate $\lambda = \frac{\beta}{n-1}$ and death rate $\mu = \beta$. Its is both Markov and irreducible, hence by Lemma 3.6 it is ergodic. Using the Kolmogorov forward equations one can derive, see [2, Ch. 6], that

$$\begin{aligned} p_{01}(t) &= \frac{1}{n}(1 - e^{-(1+\frac{1}{n-1})\beta t}) \\ p_{11}(t) &= \frac{1}{n}(1 + (n-1)e^{-(1+\frac{1}{n-1})\beta t}). \end{aligned}$$

Since both these terms converges towards $\frac{1}{n}$ we know that π is a limiting distribution for $E_{ij}(t)$. For Markov processes limiting distributions are stationary distributions, and for ergodic processes the stationary distribution is unique. □

Lemma 3.8. *The critical dynamic Erdős-Rényi graph* $G = \{G(t),\ t \geq 0\}$ *is an ergodic Markov process, with stationary distribution equal to that of a critical Erdős-Rényi graph.*

Proof. By definition we have that $G(t) = (E_{12}(t), \ldots, E_{n-1,n}(t))$, which lives on a finite state space. By Lemma 3.6 G is ergodic if we can prove that G is irreducible and has the Markov property.
To show that $G = (E_{12}, \ldots, E_{n-1,n})$ has the Markov property we must show that,

$$P(G(t) = i | G(s) = j, G(t_k) = i_k, \ldots, G(t_0) = i_0) = P(G(t-s) = i | G(0) = j)$$

for all $t > s > t_k > \cdots > 0$ and all $i, j, i_k, \ldots, i_0 \in \{0,1\}^{\binom{n}{2}}$ whenever these conditional probabilities exists. This is just a simple—long but tedious—exercise in convolution formulas and an independence argument.

Let $i = (i_{12}, \ldots, i_{n-1,n})$ and $j = (j_{12}, \ldots, j_{n-1,n})$ represent states in $\{0,1\}^{\binom{n}{2}}$. We know the distribution of $P(G(t) = i|G(0) = j)$ as it is completely determined by the edge probability at t for the underlying edge processes and the independence between them.
For instance, $P(G(t) = (1, \ldots, 1)|G(0) = (0, \ldots, 0)) = p_{01}(t)^{\binom{n}{2}}$. We can deduce that $P(G(t) = i|G(0) = j) > 0$ since $p_{01}(t),\ p_{11}(t),\ p_{10}(t),\ p_{00}(t) > 0$. We conclude that G is irreducible, and therefore also ergodic.

Furthermore, since the probability that there is an edge present between a vertex pair converges to $\frac{1}{n}$ we can also conclude that the dynamic graph's stationary distribution is that of a critical Erdős-Rényi graph.

□

Lemma 3.9. *Let $e = \{e(t),\ t \geq 0\}$ be the number of edges present in the critical dynamic Erdős-Rényi graph at time t. Then $e = \{e(t),\ t \geq 0\}$ is an ergodic birth-death process, on $\{0, 1, \ldots, N = \binom{n}{2}\}$, starting in state 0 with birth-rates $\lambda_k = \frac{(N-k)\beta}{n-1}$ and death-rates $\mu_k = k \cdot \beta$.*

Proof. We'll give an informal proof of this, since a formal proof would require invoking the strong Markov property of the underlying processes.
When $e(t)$ enters state k, the underlying processes forgets how long they have been in their current state—this is by the strong Markov property since the time which $e(t)$ enters k is random. When in state k the time until the next death is the minimum of k independent Exp(β)-variables, i.e. an Exp($k \cdot \beta$) time — since the minimum of independent exponential variables is again exponential with new parameter equal to the sum of the parameters; similarly the time until the next birth is an Exp($\frac{N-k}{(n-1)\beta}$)-time. Since there times are independent it follows that $e(t)$ is a birth-death process. □

In closing this section we conclude that, $E_{ij}(t)$ the edge process, $G(t)$ the dynamic graph, and $e(t)$ the number of edges at time t are all ergodic Markov processes. Furthermore, we have seen that the stationary distribution of the dynamic graph is that of a critical Erdős-Rényi graph.

4 The Fastest Time To Stationarity

We have seen that the critical dynamic Erdős-Rényi graph converges in distribution to a critical Erdős-Rényi graph. Natural questions that arise are,

- Does there exist a random time T for which the process is stationary?
- If yes, does it stay in stationarity upon entering it?
- Is there, in some sense, a fastest such time?

In this section we will prove that all of the above questions are answered in the positive, and we call such a time *the fastest time to stationary.*
We shall do so by simply constructing T and then proving that it must be the fastest time to stationarity for a certain class of random variables — randomized stopping times.

In constructing T we shall find the fastest times T_{ij} to stationary for the underlying edge processes, and taking T to be the maximum of those. The rationale being that once an edge process enters stationarity it stays there, therefore waiting until all the edge processes have entered stationarity should ensure that the dynamic graph is in stationarity.

In order to make everything rigorous we shall need the concept of a strong stationary times as well as the separation measure. Roughly speaking, a strong stationary time T for a stochastic process X is a stopping time for X with some extra external randomness such that $X(T)$ has the stationary distribution and is independent of T, see Definition 4.2. The separation $s(t) = \sup_{y \in S}\left(1 - \frac{P(X(t)=y)}{\pi(y)}\right)$, see Definition 4.1, for a stochastic process is a function in time which measures the "distance" between the distribution at time t and its stationary distribution.

Strong stationary times are well-understood for ergodic Markov processes on countable state spaces, and for the interested reader see [6]. The major result of [6] is that for ergodic Markov processes on countable state spaces the following holds,

1. If T is a strong stationary time, then for all $0 \leq t < \infty$

$$s(t) \leq P(T > t)$$

2. If the state space of the process is finite, there exist a strong stationary time T such that 1. holds with equality. We shall call such a time the time to stationary, or the fastest time to stationary.

The proof of this is very technical, mainly because it was made to hold for a large class of processes. Nevertheless, we shall draw upon the results and prove them for our specific situation. Namely, we will give an explicit time T when the critical dynamic Erdős-Rényi graph reaches stationarity and then prove that this time is the smallest time it takes to get there.

4.1 Separation and Strong Stationary Times

One way to measure how close a process is to its stationary distribution at time t is via the separation. There is other more intuitive measures to do this with—e.g. the variation distance—but it turns out that the separation bounds the distribution of the time to stationary very nicely.

Definition 4.1. *Let X be a stochastic process taking values in S, with a stationary distribution π and some arbitrary initial distribution π' Then the separation is defined as,*

$$s(t) = \sup_{y \in S} \left(1 - \frac{P_{\pi'}(X(t) = y)}{\pi(y)} \right)$$

where $P_{\pi'}(X(t) = y)$ is the probability that $X(t) = y$ if $X(0) = x$ is chosen according to the distribution π'.

For convenience we shall also use the notation,

$$a(t) = \inf_{y \in S} \frac{P_{\pi'}(X(t) = y)}{\pi(y)} = 1 - s(t)$$

The separation reflects how close a process is to its stationary distribution at time t, as small values of separation means that the distributions are close.

In order to define a strong stationary time we recall the concept of a stopping time, see Definition 3.5 A stopping time T for the process X is a random variable such that if we know X up to time t we know whether $T \leq t$ or not, i.e. if we should stop or not. Intuitively, it does not seem that a stopping time can be a strong stationary time. Think of how a stopping time for X works in practice. We observe X for some time and at some point we know whether to stop or not. This clearly does not work with stationarity—since we cannot observe when X enters stationarity.

Turns out that intuition is correct, in order to define a strong stationary time one needs to start with a stopping time and then inject some external randomness. This leads of to the definition of a randomized stopping time.

Definition 4.2. *Let* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ *be a filtered probability space. Let* $\mathcal{F}_\infty$ *be the smallest* σ*-algebra containing* $\mathcal{F}_t$ *for all* t*.*
Furthermore, let $\mathcal{G} \subset \mathcal{F}$ *be a sub-*σ*-algebra of* $\mathcal{F}$ *independent of* $\mathcal{F}_\infty$*. We say that* $T : \Omega \to [0, \infty]$ *is a randomized stopping time relative to* $\{\mathcal{F}_t\}$ *if for each* $t \geq 0$*,*

$$\{T \leq t\} \in \sigma(\mathcal{F}_t, \mathcal{G})$$

where $\sigma(\mathcal{F}_t, \mathcal{G})$ *is the smallest* σ*-algebra containing both* $\mathcal{F}_t$ *and* $\mathcal{G}$*.*
If $\mathcal{F}_t = \sigma(\{X_s,\ 0 \leq s \leq t\})$ *is the natural filtration of some process* X *we say that* T *is a randomized stopping time relative to* X

Remark. *We see that a randomized stopping time relative to* $\mathcal{F}_t$ *is just a normal stopping time relative to* $\{\sigma(\mathcal{F}_t, \mathcal{G})\}$*. Also this concept encompasses that of a normal stopping time for a process, since we may just take* $\mathcal{G}$*, the external randomness, to be the trivial* σ*-algebra.*

As the definition is rather technical an example may help to illuminate the concept. Suppose X is a stochastic process and that T is a stopping time for X. Now if Y is a random variable independent of X, then $\min(T, Y)$ would be a randomized stopping time for X.

With the concept of a randomized stopping time in hand we are ready to define what is meant by a strong stationary time.

Definition 4.3. *Let* X *be a stochastic process on* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ *where* $\mathcal{F}_t = \sigma(\{X_s,\ 0 \leq s \leq t\})$ *is the natural filtration of* X*. Assume* X *takes values in some state space* S *and has a unique stationary distribution* π*. Furthermore let* T *be a randomized stopping time relative to* X*. Then,* T *is said to be a strong stationary time if,* $X(T)$ *has the stationary distribution and is independent of* T *conditionally given that* $\{T < \infty\}$*, i.e. if,*

$$P(T \leq t, X(T) = y) = P(T \leq t) \cdot P(X(T) = y) = P(T \leq t) \cdot \pi(y)$$

for all $0 \leq t < \infty$ *and* $y \in S$

We shall state some nice implications of this definition, which is by [6, Prop. 2.4]

Proposition 4.4. *Let* X *be an ergodic Markov chain with right-continuous paths on some finite or countable state space* S*. Then the following are equivalent for a randomized stopping time relative to* X*.*

1. *T is a strong stationary time*

2. $P(T \leq t, X(t) = y) = P(T \leq t) \cdot \pi(y),\ 0 \leq t < \infty$

3. $P(T \leq t, X(u) = y) = P(T \leq t) \cdot \pi(y),\ 0 \leq t < u < \infty$

Proof. See [6, Prop. 2.4] □

Remark. *We see that strong stationary times couples very nicely with ergodic Markov chains on countable state spaces, since by 3. when they reach stationarity, they stay there. Suffice to say that this hinges on the strong Markov property of such chains.*
Definition 4.3 also implies that $P(X(T+s) = y) = \pi(y)$ conditionally on $\{T < \infty\}$.

4.2 The Fastest Time To Stationarity

Our goal is to construct a strong stationary time for the dynamic Erdős-Rényi graph, and as this process is built up by the underlying edges processes it makes sense to study the strong stationary times for these.

Lemma 4.5. *Let $E(t) = E_{ij}(t)$ be the process describing if an edge is present between vertex i and j. Let $T_0 \sim Exp(\frac{\beta}{n-1})$ be the holding time in state 0 and let $T_1 \sim Exp(\beta)$ be distributed as the holding time in state 1, independent of T_0. Then, $T = \min(T_0, T_1) \sim Exp(\beta \cdot (1 + \frac{1}{n-1}))$ is a strong stationary time for E_{ij} starting with 0 edges. Furthermore, T is the fastest time to stationarity.*

Proof. By Lemma 3.7 the edge process E, is an ergodic Markov chain with stationary distribution,

$$\pi(0) = 1 - \frac{1}{n}$$
$$\pi(1) = \frac{1}{n}$$

Secondly we see that $T = \min(T_0, T_1)$ is a randomized stopping time relative to the natural filtration of E and with external randomness $\mathcal{G} = \sigma(T_1)$. Were we have assumed that the underlying probability space is rich enough to support an exponential random variable, independent of T_0. In practice this is always the case, since we can just simulate such a variable.

To clarify we start the process and simultaneously simulate $T_1 \sim \text{Exp}(\beta)$ independently of the process. As the stopping time we take the smallest value of $T_0 \sim \text{Exp}(\frac{\beta}{n-1})$, the time until the first edge occurs, and T_1 an independently simulated random variable.

In order to prove that T is also a strong stationary time for E we need to show that

$$P(T \leq t, E(T) = 1) = P(T \leq t) \cdot \frac{1}{n}$$
$$P(T \leq t, E(T) = 0) = P(T \leq t) \cdot (1 - \frac{1}{n}).$$

We'll prove it for the first case, since calculations are completely analogous for the second case.

Hence we must show that,

$$P(T \leq t, E(T) = 1) = P(T \leq t) \cdot \frac{1}{n}. \tag{4.1}$$

For this purpose we note that $E(T) = 1 \iff T_0 < T_1$ since then an edge has just been added and we have stopped. Using this we get,

$$P(T \leq t, E(T) = 1) = P(T \leq t, T_0 < T_1) = P(T \leq t | T_0 < T_1) \cdot P(T_0 < T_1).$$

With basic probability calculus we get,

$$\begin{aligned} P(T_0 < T_1) &= \frac{\beta/(n-1)}{\beta/(n-1) + \beta} = \frac{1}{n} = \pi(1) \\ P(T \leq t | T_0 < T_1) &= P(\min(T_0, T_1) \leq t | T_0 < T_1) = P(T \leq t). \end{aligned}$$

We can conclude that $P(T \leq t, E(T) = 1) = P(T \leq t) \cdot \frac{1}{n}$ and that T is a strong time to stationary for the edge process E.

We recall that if $Y_1 \sim \text{Exp}(\alpha)$ and $Y_2 \sim \text{Exp}(\beta)$ are independent then
$Z = \min(Y_1, Y_2) \sim \text{Exp}(\alpha + \beta)$
Therefore T is distributed as,

$$P(T > t) = e^{-\beta(1 + \frac{1}{n-1})t}$$

Next we must show that if T' is any other strong stationary time for E we have that,

$$P(T > t) \leq P(T' > t)$$

i.e. that T is the stochastically fastest time to stationary. We begin by calculating the separation for the process E starting with 0 edges. Let $p_{01}(t)$ be the transition probability of starting in state 0 and begin in state 1 at time t. We have seen that, see Lemma 3.7

$$\begin{aligned} p_{01}(t) &= \frac{1}{n}(1 - e^{-\beta(1 + \frac{1}{n-1})t}) \\ p_{00}(t) &= 1 - p_{01}(t) = \frac{1}{n} e^{-\beta(1 + \frac{1}{n-1})t} + \frac{n-1}{n}. \end{aligned}$$

Since $\frac{p_{01}(t)}{\pi(1)} < \frac{p_{00}(t)}{\pi(0)}$ the separation $s(t)$ is as follows,

$$\begin{aligned} s(t) &= \sup_{y \in \{0,1\}} (1 - \frac{P_0(E(t) = y)}{\pi(y)}) = 1 - \frac{p_{01}(t)}{\pi(1)} \\ &= e^{-\beta(1 + \frac{1}{n-1})t} = P(T > t). \end{aligned}$$

Now let T' be any strong stationary time for E. By Proposition 4.4 T' satisfies,

$$P(T' \leq t) = \frac{P(T' \leq t, E(t) = 1)}{\pi(1)} \leq \frac{P(E(t) = 1)}{\pi(1)}$$
$$\iff P(T' > t) \geq 1 - \frac{P(E(t) = 1)}{\pi(1)} = s(t) = P(T > t).$$

We have proven that T is the fastest randomized stopping time to stationarity.

□

Remark. *We just proved that $T = \min(T_0, T_1)$ is the fastest time to stationarity, however we gave no insight on how we arrived at this time — something we will do now.*

Let E_{01} be an edge process starting in 0 and let E_{10} be an independent probabilistic copy of that process, but starting in 1. Then $T = \min(T_0, T_1)$ is the first time the processes meet, where $T_i,\ i = 0, 1$ is the first holding times for the processes. The idea now being that when the processes meet, their initial states does not matter anymore and both processes are in stationarity.

We have seen that an ergodic Markov process stay in stationarity after entering it, see 4.4. It therefore seem very reasonable that if we have two independent ergodic Markov processes they will both be in stationarity if we wait until both have entered it. The following Lemma shows that this is indeed the case.

Lemma 4.6. *Let X and Y be two independent ergodic Markov processes each taking values in finite state spaces S_x and S_y, with stationary distributions π_X and π_Y. Let T_X be the time to stationary for X and T_Y for Y. Then $T = \max(T_X, T_Y)$ is the time to stationary for the process (X, Y)*

Proof. We start be stating that by [6, Th. 1.1] T_X and T_Y exists (also see Theorem stated in beginning of section).

It is straight forward to modify the proof of Lemma 2.4 to show that the process (X, Y) is an ergodic Markov chain, which we will not but instead just state that (X, Y) is an ergodic Markov chain.

The stationary distribution of (X, Y) exists and is given by,

$$\lim_{t \to \infty} P(X(t) = x, Y(t) = x) = \pi_X(x) \cdot \pi_Y(y)$$

for all $(x, y) \in S_x \times S_y$.

First we must argue that $T = \max(T_X, T_Y)$ is a randomized stopping time for the process (X, Y). If we agree we can circumvent a formal proof of this, which would involve defining a new probability space etc. then it becomes fairly easy. If we know $\mathcal{F}_t^X$ and $\mathcal{F}_t^Y$, the natural filtrations, as well as the external sources of randomness for T_X and T_Y, then we know the value of $T = \max(T_X, T_Y)$.

We conclude that T is indeed a randomized stopping time for (X, Y).

Secondly we must prove that T is a strong stationary time for (X, Y) which by Proposition 4.4—since (X, Y) is an ergodic Markov process on a finite state space—is equivalent to showing that,

$$P(T \leq t, X(t) = x, Y(t) = y) = P(T \leq t)\pi_X(x) \cdot \pi_Y(y)$$

for all $0 \leq t < \infty$, $x \in S_x,\ y \in S_y$.
We recall the equality $\{\max(Z_1, \ldots, Z_n) \leq t\} = \{Z_1 \leq t, \ldots Z_n \leq t\}$.

$$\begin{aligned}
&P(T \leq t, X(t) = x, Y(t) = y) = P(T_X \leq t, T_Y \leq t, X(t) = x, Y(t) = y)\\
&\overset{1.)}{=} P(T_X \leq t, X(t) = x) \cdot P(T_Y \leq t, Y(t) = y)\\
&\overset{2.)}{=} P(T_X \leq t) \cdot \pi_X(x) \cdot P(T_Y \leq t) \cdot \pi_Y(y)\\
&= P(T_X \leq t, T_y \leq t) \cdot \pi_X(x) \cdot \pi_Y(y) = P(T \leq t) \cdot \pi_X(x) \cdot \pi_Y(y).
\end{aligned}$$

1.) (X, T_X) and (Y, T_Y) are independent random variables.
2.) We have that T_X is a stationary time for X and this is a implication of Proposition 4.4, same holds for T_Y and Y. We can conclude that T is also a strong stationary time for (X, Y).

Left to show is that $P(T > t) = s(t)$ where $s(t)$ is the separation measure of (X, Y). First we recall the following relation,

$$s(t) = \sup_{(x,y)}\left(1 - \frac{P(X(t) = x) \cdot P(Y(t) = y)}{\pi_X(x) \cdot \pi_Y(y)}\right) = 1 - a(t) = 1 - \inf_{(x,y)} \frac{P(X(t) = x) \cdot P(Y(t) = y)}{\pi_X(x) \cdot \pi_Y(y)}.$$

We can conclude that $P(T \leq t) = a(t) \iff P(T > t) = s(t)$.

$$\begin{aligned}
P(T \leq t) &= P(T_X \leq t) \cdot P(T_Y \leq t) \overset{1.)}{=} a_X(t) \cdot a_Y(t)\\
&= \inf_x \frac{P(X(t) = x)}{\pi_X(x)} \cdot \inf_y \frac{P(Y(t) = y)}{\pi_Y(y)} = \inf_{(x,y)} \frac{P(X(t) = x) \cdot P(Y(t) = y)}{\pi_X(x) \cdot \pi_Y(y)} = a(t).
\end{aligned}$$

1.) We have that T_X and T_Y are the fastest times to stationary for X and Y, and therefore $P(T_X \leq t) = a_X(t)$ and $P(T_Y \leq t) = a_Y(t)$

We can conclude that $P(T > t) = s(t)$, and we recall that this means that T is the fastest time to stationary. Since if T' is any other strong time to stationary we have,

$$P(T' \leq t) = \frac{P(T' \leq t, X(t) = x, Y(t) = y)}{\pi_X(x)\pi_Y(y)} \leq \frac{P(X(t) = x) \cdot P(Y(t) = y)}{\pi_X(x)\pi_Y(y)}$$

since this holds for all states (x, y) we can conclude that $P(T' \leq t) \leq a(t) = P(T \leq t)$, and we see that T is faster. □

We are now ready to present to major result of this section, the fastest time to stationary for the critical dynamic Erdős-Rényi graph.

Proposition 4.7. *Let* $G = \{G(t),\ t \geq 0\} = \{(E_{12}(t), \dots, E_{n-1,n}(t)),\ t \geq 0\}$ *be the critical dynamic Erdős-Rényi graph, where* $E_{ij}(t)$ *are the processes representing if an edge is present or not between vertex* i *and* j*. Let* $T_{ij} \sim$ *Exp*$(\beta \cdot (1 + \frac{1}{n-1}))$ *be the fastest time to stationary for* E_{ij}*. Then,*

$$T = \max_{i,j}\{T_{ij}\}$$

is the fastest time to stationary for G*. Furthermore, its distribution is given by,*

$$P(T < t) = (1 - e^{-\beta(1+\frac{1}{n-1})t})^{\binom{n}{2}}$$

Proof. By Lemma 4.5, $T_{ij} = \min(T_0, T_1) \sim \text{Exp}(\beta \cdot (1 + \frac{1}{n-1}))$ is the fastest time to stationary for E_{ij}. Now apply Lemma 4.6 with an induction argument to conclude that $T = \max_{i,j}\{T_{ij}\}$ is the fastest time to stationary for the critical dynamic Erdős-Rényi graph.
Furthermore,

$$\begin{aligned} P(T < t) &= P(\max(T_{12}, \dots, T_{n-1,n}) < t) \\ &= P(T_{ij} < t)^{\binom{n}{2}} = (1 - e^{-\beta(1+\frac{1}{n-1})t})^{\binom{n}{2}} \end{aligned}$$

□

4.3 Asymptotics

The expression for the fastest time to stationary in Proposition 4.7 is exact but not very intuitive. The following corollary deals with asymptotics of its distribution function—namely we provide two breaking points, depending on n, such that if $t > \bar{f}(n)$ then $P(T < t(n)) \to 1$ and if $t < \underline{f}(n)$ then $P(T < t(n)) \to 0$ as $n \to \infty$. We shall also see that these breaking points are relatively close to each other.

Proposition 4.8. *Let* T *be the fastest time to stationary for the critical dynamic Erdős-Rényi graph. Then for any* $\gamma > 0$ *the following asymptotics holds,*

$$\begin{aligned} &P(T < t(n)) \to 1 \text{ as } n \to \infty \text{ if } t(n) > \frac{(2+\gamma)\cdot\log(n)}{\beta} \\ &P(T < t(n)) \to 0 \text{ as } n \to \infty \text{ if } t(n) < \frac{(2-\gamma)\cdot\log(n)}{\beta}. \end{aligned}$$

Proof. Recall the Taylor expansion of $\log(1-x)$ valid for $|x| < 1$,

$$\log(1-x) = -\sum_{m=1}^{\infty} \frac{x^m}{m}.$$

Let $A(n) = 1 + \frac{1}{n-1}$, and $g(t) = (1 - e^{-\beta A t})^{\frac{n^2}{2}}$
Take $\bar{f}(n) = \frac{-\log(n^{-(2+\gamma)})}{\beta} = \frac{(2+\gamma)\cdot\log(n)}{\beta}$ as a candidate for the upper bound.

$$\begin{aligned} g(\bar{f}(n)) &= (1 - e^{-\beta A \bar{f}(n)})^{\frac{n\cdot(n-1)}{2}} = (1 - n^{-A\cdot(2+\gamma)})^{\frac{n\cdot(n-1)}{2}} \\ &= \exp\{\frac{n\cdot(n-1)}{2} \cdot \log(1 - n^{-A\cdot(2+\gamma)})\} \overset{1.)}{=} \exp\{-\frac{n\cdot(n-1)}{2} \cdot \sum_{m=1}^{\infty} \frac{(n^{-A\cdot(2+\gamma)})^m}{m}\} \\ &\overset{2.)}{=} \exp\{-\frac{n\cdot(n-1)}{2\cdot n^{2A+A\gamma}}\} \cdot \exp\{-O(n^{2-2A(2+\gamma)})\} \to 1 \text{ as } n \to \infty. \end{aligned}$$

1.) This is the Taylor expansion of $\log(1 - n^{-A\cdot(2+\gamma)})$
2.) We have that

$$\sum_{m=1}^{\infty} \frac{(n^{-A(2+\gamma)})^m}{m} = n^{-A(2+\gamma)} + \sum_{m=2}^{\infty} \frac{(n^{-A(2+\gamma)})^m}{m} = n^{-A(2+\gamma)} + O(n^{-2A(2+\gamma)})$$

multiplying everything with $\frac{n\cdot(n-1)}{2}$ and we get that this equals

$$\frac{n\cdot(n-1)}{2\cdot n^{2A+A\gamma}} + O(\frac{n\cdot(n-1)}{2} \cdot n^{-2A(2+\gamma)}) = \frac{n\cdot(n-1)}{2\cdot n^{2A+A\gamma}} + O(n^{2-2A(2+\gamma)}) \to 0 \text{ as } n \to \infty.$$

We can conclude that,

$$P(T < t(n)) \to 1 \text{ as } n \to \infty \text{ if } t(n) > \frac{(2+\gamma)\cdot\log(n)}{\beta}.$$

Now, analogous calculations shows that,

$$P(T < t(n)) \to 0 \text{ as } n \to \infty \text{ if } t(n) < \frac{(2-\gamma)\cdot\log(n)}{\beta}.$$

□

Remark. *As the upper bound* $\frac{(2+\gamma)\cdot\log(n)}{\beta}$ *is relatively small we see that the dynamic graph enters stationarity very quickly, see Figure 1 and 2.*

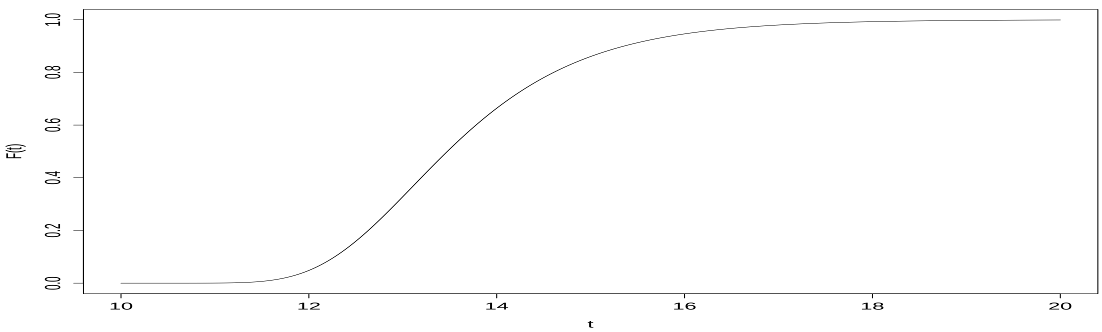

Figure 1: $P(T < t)$ for $n = 1000$ and $\beta = 1$

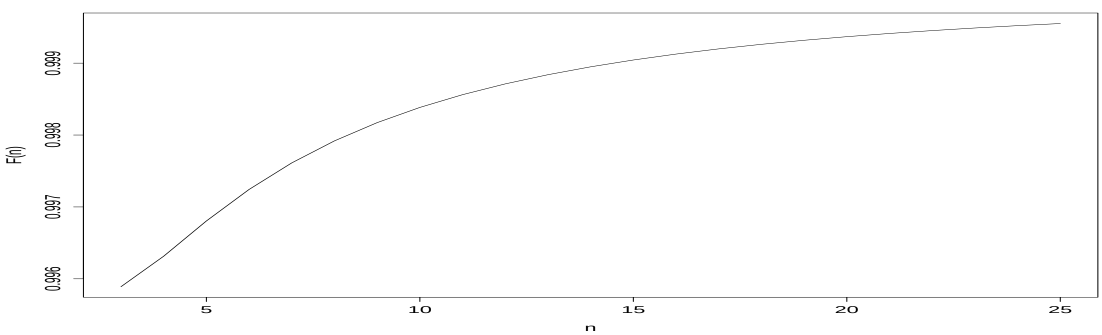

Figure 2: The speed of convergence for $P(T < \frac{(2+\gamma)\cdot\log(n)}{\beta})$, $\gamma = 2$

4.4 Conclusion and Discussion

In closing this section we can conclude that the time it takes for the dynamic critical Erdős-Rényi graph to reach stationarity is given by a randomized stopping time distributed as the maximum of a $N = \binom{n}{2}$ number of independent $\text{Exp}(\beta \cdot (1 + \frac{1}{n-1}))$-variables. Furthermore, this is also the stochastically fastest time it takes to reach stationarity, for a very reasonable class of random variables — the strong stationary times for the dynamic graph. We also conclude that the time to reach stationarity is relatively small.

The usefulness of this result is purely theoretical, as in practice we cannot give a time for which the process is in stationarity.

But in analyzing the long term behavior of the graph the result can indeed be useful. As we have seen the time until the graph reaches stationarity is relatively small—of order $O(\log(n))$— and once it reaches stationarity it stays there. Hence for large n and a large time t we can conclude that in the period $[0, t]$ it spends most of that time in stationarity, and the time is takes to reach stationarity should be negligible compared to t. So in studying certain properties of the critical dynamic graph one may be able to reduce the problem to study properties of the graph when stationary — something that should be easier and more tractable.

Also we mention that the method used in constructing the fastest time to stationarity does not hinge on the critical setting of the random graph, one could easily extend the results to hold for any values of β and α — the tuning parameters of the dynamic graph.

5 The Size of The Largest Component Over Time

Often of interest when studying random graphs is the size of its largest component, since this tend to have interesting practical implications. In this section we will give an asymptotic lower bound for the probability the the critical Erdős-Rényi graph's largest component in the interval $[0, t]$ has exceeded size $\epsilon \cdot n$ for any given $\epsilon \in (0, 1)$. In doing so we shall use the intimate relation between the size of the largest component of an Erdős-Rényi graph and the number of edges present. We begin with some notation.

5.1 Notation and Definitions

Let $G = \{G(t),\ t \geq 0\}$ be a critical dynamic Erdős-Rényi graph, and let $e = \{e(t),\ t \geq 0\}$ be the processes that represents the number of edges present at time t. Furthermore let $c = \{c(t),\ t \geq 0\}$ be the processes that represents the size of the largest component at time t, and let $C(t) = \max_{0 \leq s \leq t} c(s)$ be the size of the largest component in the interval $[0, t]$. Let $C_1(n, m)$ denote the size of the largest component in a Erdős-Rényi graph with n vertices and m edges, and $C_1(n, p)$ is the size of the largest component in a Erdős-Rényi graph with n vertices and edge probability p. Finally let T_k be the time it takes for the size of largest component in the dynamic graph to reach size k, and let t_k be the time it takes for the graph to reach k edges.

To summarize,

- $G(t)$ is the critical dynamic Erdős-Rényi graph
- $e(t)$ is the number of edges at time t
- $c(t)$ is the size of the largest component at time t
- $C_1(n, m)$ is the size of the largest component in a Erdős-Rényi graph with n vertices and m edges
- $C_1(n, p)$ is the size of the largest component in a Erdős-Rényi graph with n vertices and edge probability p.
- t_k is the time it takes for the graph to reach k edges
- $C(t)$ is the size of the largest component in the interval $[0, t]$, i.e. $C(t) = \max_{0 \leq s \leq t} c(s)$
- T_k is the time it takes for the largest component to reach size k

and we will give a lower bound for the probability $P(C(t) > \epsilon \cdot n) = P(T_{\epsilon \cdot n} < t)$.

We know—since the dynamic graph is ergodic and therefore all states will be reached—that the size of largest component will eventually exceed $\epsilon \cdot n$. The question is how it does this. There is basically two ways in which this can occur. Either it happens when the graph is in stationarity with some very unlikely configuration of edges, or it happens after the dynamic graph have reach some specified number of edges for which the size of the largest component *at that time* is very likely to exceed $\epsilon \cdot n$.

In solving the problem we shall opt for the second approach, i.e. in finding a lower bound we shall wait until the dynamic graph have reached some number of edges k for which the size of largest component *at that time* $C_1(n, k)$ is very likely to exceed $\epsilon \cdot n$. The stationary approach is discussed in the Discussion section, and suffice to say it was investigated but the technical difficulties proved insurmountable.

5.2 A Simple Inequality

The following inequality is rather trivial, but will prove useful.

Lemma 5.1. *Let $\epsilon \in (0, 1)$. Then,*

$$P(C(t) > \epsilon \cdot n) > P(C_1(n, k) > \epsilon \cdot n) \cdot P(t_k \leq t)$$
$$> P(C_1(n, k) > \epsilon \cdot n) \cdot (1 - \frac{E(t_k)}{t}).$$

Proof. We have that,

$$P(C(t) > \epsilon \cdot n) = P(C(t) > \epsilon \cdot n | t_k \leq t) \cdot P(t_k \leq t) + P(C(t) > \epsilon \cdot n | t_k > t) \cdot P(t_k > t)$$
$$> P(C(t) > \epsilon \cdot n | t_k \leq t) \cdot P(t_k \leq t) \overset{1.)}{>} P(C_1(n, k) > \epsilon \cdot n) \cdot P(t_k \leq t)$$
$$= P(C_1(n, k) > \epsilon \cdot n) \cdot (1 - P(t_k > t)) \overset{2.)}{>} P(C_1(n, k) > \epsilon \cdot n) \cdot (1 - \frac{E(t_k)}{t}).$$

1.) We are asking for the probability that $C(t) > n \cdot \epsilon$ given that we know that the graph has reached k edges in $[0, t]$, hence this probability must be larger than $P(C_1(n, k) > \epsilon \cdot n)$.
2.) This is just the Markov inequality. □

Remark. *We see that this lower bound is only good for $t > E(t_k)$, since it's actually negative for smaller t.*

Now if we can choose the number of edges k in such away that $P(C_1(n, k) > \epsilon \cdot n) \to 1$ as $n \to \infty$ we see that $P(C(t) > \epsilon \cdot n) = (1 - \delta) \cdot (1 - \frac{E(t_k)}{t})$, for $\delta \in (0, 1)$ and large enough n. In order to do so we need a classical result from Erdős-Rényi. The proof is very technical and is left as a reference.

Theorem 5.2. *Let $\{\mathcal{G}(n, M(n))\}$ be a sequence of Erdős-Rényi graphs with n vertices and $M(n)$ edges. Let $C(n, M(n))$ be the size largest component of $\mathcal{G}(n, M(n))$. Furthermore assume that $\lim_{n\to\infty} \frac{M(n)}{n} = c > \frac{1}{2}$.*

Then, we have for any $\epsilon > 0$,

$$\lim_{n\to\infty} P(|\frac{C(n, M(n))}{n} - A(c)| > \epsilon) = 0$$

where $A(c) = 1 - \frac{x(c)}{2c}$ and $x(c) = \sum_{k=1}^{\infty} \frac{k^{k-1}}{k!}(2ce^{-2c})^k$ is the unique solution satisfying $0 < x(c) < 1$ the equation $x(c)e^{-x(c)} = 2ce^{-2c}$.

Proof. See [5] □

So if the number of edges grows as $c \cdot n$ we have that

$$\frac{C(n, M(n))}{n} \xrightarrow{p} A(c).$$

So c determines the size of $\frac{C(n,M(n))}{n}$ as n grows. This is very useful, but we are in a different situation. Namely, we are asking if there for a given size $A \in (0, 1)$ exist a c such that if $\lim\limits_{n\to\infty} \frac{M(n)}{n} = c$ we have that

$$\frac{C(n, M(n))}{n} \xrightarrow{p} A.$$

This is indeed the case as the following Corollary shows, as well as giving the exact relation between the proportional size of the largest component $A(c)$ and c.

Corollary 5.3. *Let $\{\mathcal{G}(n, M(n))\}$ be a sequence of Erdős-Rényi graphs with n vertices and $M(n)$ edges. Let $C(n, M(n))$ be the size of the largest component of $\mathcal{G}(n, M(n))$.*
Then, for every $0 < \epsilon < 1$ there exist a $c > \frac{1}{2}$ such that $A(c) = \epsilon$, and if $M(n) = \lfloor c \cdot n \rfloor$,

$$\frac{C(n, M(n))}{n} \xrightarrow{p} A(c) = \epsilon.$$

Where $\lfloor x \rfloor$ denotes the largest integer smaller than or equal to x.
Furthermore,

$$c = \frac{-\log(1 - \epsilon)}{2 \cdot \epsilon}.$$

Proof. Pick $0 < \epsilon < 1$. In order to apply Theorem 5.2 we must find a $c > \frac{1}{2}$ such that $A(c) = \epsilon$. Hence c must satisfy

$$1 - \frac{x}{2c} = \epsilon \iff 2c = \frac{x}{1 - \epsilon} \tag{5.1}$$

where $0 < x < 1$ and x satisfies,

$$xe^{-x} = 2ce^{-2c}. \tag{5.2}$$

By Theorem 5.2 if $c > \frac{1}{2}$ equation (5.2) has a unique solution.
Now (5.1) in (5.2) gives us,

$$x \cdot e^{-x} = \frac{x}{1-\epsilon} \cdot e^{\frac{x}{1-\epsilon}} \iff x = \frac{-(1-\epsilon)\cdot \log(1-\epsilon)}{\epsilon}.$$

Using this x in (5.1) gives the following c,

$$c = \frac{-\log(1-\epsilon)}{2\epsilon}.$$

The Taylor approximation of $-\log(1-\epsilon)$ is for $\epsilon \in (0,1)$,

$$-\log(1-\epsilon) = \sum_{n=1}^{\infty} \frac{\epsilon^n}{n} = \epsilon + \sum_{n=2}^{\infty} \frac{\epsilon^n}{n}.$$

Using this we see that $c > \frac{1}{2}$.
So for $c = \frac{-\log(1-\epsilon)}{2\epsilon}$ we have that $A(c) = 1 - \frac{x}{2c} = \epsilon$.
Since $\lim_{n\to\infty} \frac{M(n)}{n} = \lim_{n\to\infty} \frac{\lfloor c\cdot n\rfloor}{n} = c$ we have by Theorem 5.2,

$$\frac{C(n, \lfloor c \cdot n \rfloor)}{n} \xrightarrow{p} \epsilon$$

□

From Corollary 5.3 we can deduce how large the numbers of edges k has to be in order for $\lim_{n\to\infty} P(C_1(n,k(n)) > \epsilon \cdot n) = 1$. If c is such that $A(c) = \epsilon$, the following Proposition shows that in order for $\lim_{n\to\infty} P(C_1(n,k(n)) > \epsilon \cdot n) = 1$ we must choose the number of edges as $k(n) = \lfloor c' \cdot n \rfloor$ where c' is slightly larger than c.

Proposition 5.4. *Let $0 < \epsilon < 1$ and $0 < \eta < 1-\epsilon$. Let c' : $A(c') = \epsilon + \eta$ where $A(c')$ is as in Theorem 5.2, the size of the largest component in an Erdős-Rényi graph with n vertices and $\lfloor c' \cdot n \rfloor$ edges.*

Then,

$$c' = \frac{-\log(1-\epsilon-\eta)}{2\cdot(\epsilon+\eta)}$$

and

$$\lim_{n\to\infty} P(C_1(n, \lfloor c' \cdot n \rfloor) > \epsilon \cdot n) = 1.$$

Proof. The first part of the Proposition was proven in Theorem 5.3. Let $M(n) = \lfloor c' \cdot n \rfloor$. By the same theorem we have that, for given $\delta > 0$

$$\lim_{n\to\infty} P(|\frac{C(n,M(n))}{n} - A(c')| < \delta) = \lim_{n\to\infty} P(A(c') - \delta < \frac{C(n,M(n))}{n} < A(c') + \delta) = 1$$

which implies that,

$$P(\frac{C(n,M(n))}{n} > A(c') - \delta) \geq P(A(c') - \delta < \frac{C(n,M(n))}{n} < A(c') + \delta).$$

This holds for all $\delta > 0$ and therefore it holds for $\delta = \eta$ and since $A(c') = \epsilon + \eta$ we can conclude that,

$$\lim_{n\to\infty} P(\frac{C(n, \lfloor c' \cdot n \rfloor)}{n} > \epsilon) = 1.$$

□

This of course implies that for any $\delta \in (0,1)$ and for large enough n we have that,

$$P(C_1(n, \lfloor c' \cdot n \rfloor) > \epsilon \cdot n) > 1 - \delta$$

where $\epsilon \in (0,1)$ and $c' = \frac{-\log(1-\epsilon-\eta)}{2\cdot(\epsilon+\eta)}$. In our search for a lower bound for $P(C(t) > \epsilon \cdot n)$ we can conclude that we have one part out of two of the inequality needed, see Lemma 5.1, to give such a bound.

The next section will be dedicated to finding the second part needed, i.e. bounds for the expected hitting time of $\lfloor c' \cdot n \rfloor$ edges.

5.3 Hitting Times For The Number of Edges

We begin this section by giving an exact expression for $E(\tau_i(i+1))$ — the expected time it takes for the dynamic graph to go from i to $i+1$ edges. This is an important part in finding bounds for $E(\tau_0(i))$ — the expected time it takes for the dynamic graph to go from 0 to i edges as $E(\tau_0(i)) = \sum_{j=0}^{i-1} E(\tau_j(j+1))$

Proposition 5.1. *Let $e = \{e(t),\ t \geq 0\}$ represent the number of edges at time t in the critical dynamic Erdős-Rényi graph, i,e. a birth-death process on $\{0, 1, \ldots, N = \binom{n}{2}\}$ with birth rates $\lambda_k = \frac{(N-k)\cdot\beta}{n-1}$ and death rates $\mu_k = k \cdot \beta$ for $\beta > 0$. Let T_i be the holding time in state k, i.e $T_k \sim Exp(\lambda_k + \mu_k)$. Define*

$$\tau_i(j) = \inf\{t > 0;\ e(t) = j, e(0) = i\}$$

as the time it takes to go from i edges to j edges.
Then,

$$E(\tau_i(i+1)) = \frac{(n-1)\cdot(N-i-1)!i!}{\beta \cdot N!} \cdot \sum_{k=0}^{i} \binom{N}{i-k} (n-1)^k \tag{5.3}$$

Proof. Recall that by Lemma 3.9 $\{e(t),\ t \geq 0\}$ is an ergodic Markov chain on a finite state space, this ensures that the process has the strong Markov property, see [4, Th. 1.9]. For our purposes we say that a Markov process X has the strong Markov property if for any *a.s. finite* stopping time τ for X we have that $X_\tau = \{X(t+\tau),\ t \geq 0\}$ is a probabilistic copy of X starting in $X(\tau)$, as well as being independent of X up to time τ.

Let $\tau_i^Y(j) = \inf\{t > 0;\ Y(t) = j, Y(0) = i\}$ for some to-be-defined-later stochastic process. We begin by deriving a recursive formula for $E(\tau_i(i+1))$. By Lemma 3.6 $\tau_i(i+1)$ is an a.s. finite stopping time for X starting in i, with $E(\tau_i(i+1)) < \infty$. Now, we derive a recursive formula for $E(\tau_i(i+1))$ by conditioning on the first jump. Let $p(i, i+1)$ is the probability that the process moves from i edges to $i+1$ edges, and let $i \to (i-1)$ indicate such an event. Then,

$$\begin{aligned} E(\tau_i(i+1)) &= E(T_i) + p(i, i-1) \cdot E(\tau_i(i+1) | i \to (i-1)) \\ &\overset{1.)}{=} E(T_i) + p(i, i-1) \cdot E(\tau_{i-1}^Y(i+1)) \\ &\overset{2.)}{=} E(T_i) + p(i, i-1) \cdot E(\tau_{i-1}(i+1)) \\ &\overset{3.)}{=} E(T_i) + p(i, i-1) \cdot (E(\tau_{i-1}(i)) + E(\tau_i(i+1))). \end{aligned}$$

1.) Let $Y(t) = X(t+\tau_i(i-1))$, then we have by the strong Markov property that Y is Markov and independent of $\{X(t), t < \tau_i(i-1)\}$ given that $Y(0) = i-1$ which it is with probability 1. This implies that Y is independent of the first jump.
2.) This is an implication of the strong Markov property, since Y is a probabilistic copy of X their expectations match.
3.) We can write

$$\tau_{i-1}(i+1) = \tau_{i-1}(i) + \tau_i^{X_i}(i+1), \quad X_i(t) = X(t + \tau_{i-1}(i)).$$

We can again apply the strong Markov property and arrive at

$$E(\tau_{i-1}(i+1)) = E(\tau_{i-1}(i) + \tau_i(i+1)).$$

Solving the above equation for $E(\tau_i(i+1))$—which we can since its finite by Lemma 3.6—we arrive at,

$$E(\tau_i(i+1)) = \frac{E(T_i) + p(i, i-1) \cdot E(\tau_{i-1}(i))}{p(i, i+1)} \tag{5.4}$$

To prove (5.3) we use (5.4) together with induction. The following holds for, $\{e(t),\ t \geq 0\}$, our birth-death process,

$$\begin{aligned} &E(T_i) = \frac{1}{\lambda_i + \mu_i} = \frac{n-1}{\beta \cdot (N + i \cdot (n-2)} \\ &p(i, i-1) = 1 - p(i, i+1) = \frac{\mu_i}{\lambda_i + \mu_i} = \frac{i \cdot (n-1)}{N + i \cdot (n-2)} \\ &\frac{E(T_i)}{p(i, i+1)} = \frac{1/(\lambda_i + \mu_i)}{\lambda_i/(\lambda_i + \mu_i)} = \frac{n-1}{(N-i) \cdot \beta} \\ &\frac{p(i, i-1)}{p(i, i+1)} = \frac{\mu_i}{\lambda_i} = \frac{i \cdot (n-1)}{N-i}. \end{aligned}$$

Inserting this in (5.4) we get,

$$E(\tau_i(i+1)) = \frac{n-1}{(N-i)\beta} + \frac{(n-1)\cdot i}{N-i} \cdot E(\tau_{i-1}(i))$$

We will prove (5.3) with induction. For $i = 0$ we have,

$$\begin{aligned} E(\tau_0(0+1)) &= \frac{(n-1)\cdot(N-0-1)!0!}{\beta \cdot N!} \cdot \sum_{k=0}^{0} \binom{N}{i-k} (n-1)^k \\ &= \frac{n-1}{\beta \cdot N} = E(T_0) \end{aligned}$$

which coincides with (5.4).
Assume the (5.3) holds for arbitrary $i+1 < N$.

$$\begin{aligned} &E(\tau_{i+1}(i+2)) = \frac{E(T_{i+1}) + p(i+1,i)E(\tau_i(i+1))}{p(i+1,i+2)} \\ &= \frac{n-1}{(N-i-1)\beta} + \frac{(n-1)\cdot(i+1)}{N-i-1} \cdot (\frac{(n-1)(N-i-1)!i!}{\beta N!} \sum_{k=0}^{i} \binom{N}{i-k} (n-1)^k) \end{aligned}$$

which after standard but tedious algebra equals,

$$\frac{(n-1)(N-i-2)!(i+1)!}{\beta N!} \sum_{k=0}^{i+1} \binom{N}{i+1-k} (n-1)^k$$

which proves equation (5.3). □

The formula in Theorem 5.1 exact, but unfortunately not very intuitive. The following Corollary gives bounds on the expected time it takes for the graph to go from i to $i+1$ edges.

Corollary 5.2. *Let $e = \{e(t),\ t \geq 0\}$ represent the number of edges at time t in the critical dynamic Erdős-Rényi graph, i,e. a birth-death process on $\{0, 1, \ldots, N = \binom{n}{2}\}$ with birth rates $\lambda_k = \frac{(N-k)\cdot\beta}{n-1}$ and death rates $\mu_k = k \cdot \beta$ for $\beta > 0$.*
Define

$$\tau_i(j) = \inf\{t > 0;\ e(t) = j, e(0) = i\}$$

Then for $i < \frac{N}{2}$,

$$\frac{n-1}{\beta \cdot N} \cdot \frac{\rho_l^{i+1} - 1}{\rho_l - 1} \leq E(\tau_i(i+1)) \leq \frac{n-1}{\beta \cdot (N-i)} \cdot \frac{\rho_u^{i+1} - 1}{\rho_u - 1}$$

where,

$$\begin{aligned} \rho_l &= \frac{i \cdot (n-1)}{N} \\ \rho_u &= \frac{i \cdot (n-1)}{N-i} \end{aligned}$$

Proof. By Proposition 5.1 we have that,

$$\begin{aligned} E(\tau_i(i+1)) &= \frac{(n-1)\cdot(N-i-1)!i!}{\beta\cdot N!}\cdot\sum_{k=0}^{i}\binom{N}{i-k}(n-1)^k \\ &= \frac{n-1}{\beta}\sum_{k=0}^{i}\left(\frac{(N-i-1)!}{(n-i+k)!}\right)\cdot\left(\frac{i!}{(i-k)!}\right)\cdot(n-1)^k. \end{aligned}$$

We note that,

$$\frac{i!}{(i-k)!} = \prod_{j=i-k+1}^{i} j = \prod_{j=0}^{k-1}(i-k+1+j)$$

$$\frac{(N-i-1)!}{(N-i+k)!} = \prod_{j=0}^{k}\frac{1}{N-i+j} = \frac{1}{N-i+k}\prod_{j=0}^{k-1}\frac{1}{N-i+j}$$

$$\frac{i!}{(i-k)!}\cdot\frac{(N-i-1)!}{(N-i+k)!} = \frac{1}{N-i+k}\prod_{j=0}^{k-1}\frac{i-k+1+j}{N-i+j} = \frac{1}{N-i+k}\prod_{j=0}^{k-1}f(j)$$

where $f(x) = \frac{i-k+1+x}{N-i+x},\ x\in[0,k-1],\ 1\le k\le i$. We adopt the standard that empty products equals 1. We have that $f'(x) = \frac{N+k-(2i+1)}{(N-i+x)^2} \ge 0$ since $i < \frac{N}{2}$. We can conclude that $f(x)$ is an increasing function and that

$$\begin{aligned} f(j) &\ge f(0) = \frac{i-k+1}{N-i} \overset{1.)}{\ge} \frac{i}{N} \\ f(j) &\le f(k-1) = \frac{i}{N-i+k-1} \overset{2.)}{\le} \frac{i}{N-i}. \end{aligned}$$

1.) Let $g(k) = f(0) = \frac{i-k+1}{N-i}$ and recall that $k \le i$. Then,

$$\begin{aligned} g(k) - \frac{i}{N} &= \frac{i-k+1}{N-i} - \frac{i}{N} = \frac{N(i-k+1)-i(N-i)}{N(N-i)} \\ &= \frac{-kN+N+i^2}{N(N-i)} \ge \frac{i^2-iN+N}{N(N-i)} \\ &= \frac{(i-\frac{N}{2})^2+\frac{3N}{4}}{N(N-i)} \ge 0 \implies g(k) \ge \frac{i}{N} \end{aligned}$$

2.) holds true since $k \ge 1$.
We also note that,

$$\frac{1}{N} \le \frac{1}{N-i+k} \le \frac{1}{N-i}$$

Combine all of the above and we arrive at,

$$
\begin{aligned}
E(\tau_i^X(i+1)) &= \frac{(n-1)\cdot(N-i-1)!i!}{\beta\cdot N!}\cdot\sum_{k=0}^{i}\binom{N}{i-k}(n-1)^k \\
&= \frac{n-1}{\beta}\sum_{k=0}^{i}\frac{(N-i-1)!i!}{(n-i+k)!(i-k)!}(n-1)^k \\
&= \frac{n-1}{\beta}\sum_{k=0}^{i}\frac{(n-1)^k}{N-i+k}\prod_{j=0}^{k-1}f(j) \\
&\leq \frac{n-1}{\beta}\sum_{k=0}^{i}\frac{(n-1)^k}{N-i+k}\cdot f(k-1)^k \\
&\leq \frac{n-1}{\beta}\sum_{k=0}^{i}\frac{(n-1)^k}{N-i}\left(\frac{i}{N-i}\right)^k \\
&= \frac{n-1}{\beta\cdot(N-i)}\sum_{k=0}^{i}\left(\frac{i\cdot(n-1)}{N-i}\right)^k = \frac{n-1}{\beta\cdot(N-i)}\cdot(\frac{\rho_u^{i+1}-1}{\rho_u-1})
\end{aligned}
$$

where $\rho_u = \frac{i\cdot(n-1)}{N-i}$. The last step is just the closed form of a geometric series. Similarly one gets,

$$E(\tau_i(i+1)) \geq \frac{n-1}{\beta\cdot N}\cdot(\frac{\rho_l^{i+1}-1}{\rho_l-1})$$

where $\rho_l = \frac{i\cdot(n-1)}{N}$ □

Remark. *We see that if $\frac{n}{2} < i < \frac{N}{2}$ both the dominating terms in the bounds, $\frac{i(n-1)}{N}$ and $\frac{i(n-1)}{N-i}$ are larger than 1. Hence for these values of i the expected time of just moving up one step is exponentially large.*

We have managed to find bounds for the expected time to move from i to $i+1$ edges, where $i < \frac{N}{2}$. This is an important step, but of interest is to bound the expected time it takes to go from 0 to i edges. The following Corollary gives such bounds, and we shall see that this time is indeed very large. Before stating it we recall that $f(n) = O(g(n))$ if $\exists k, n_0 :\ n > n_0 \implies |f(n)| < k\cdot|g(n)|$

Corollary 5.3. *Let $e = \{e(t),\ t\geq 0\}$ represent the number of edges at time t in the critical dynamic Erdős-Rényi graph, i,e. a birth-death process on $\{0,1,\ldots,N=\binom{n}{2}\}$ with birth rates $\lambda_k = \frac{(N-k)\cdot\beta}{n-1}$ and death rates $\mu_k = k\cdot\beta$ for $\beta > 0$.*

Define

$$\tau_i(j) = \inf\{t > 0;\ e(t) = j, e(0) = i\}.$$

Then for $i < \frac{N}{2}$

$$\frac{n-1}{\beta \cdot N} \cdot \frac{\rho_l^i - 1}{\rho_l - 1} \leq E(\tau_0(i)) \leq \cdot \frac{i \cdot (n-1)}{\beta \cdot (N-i+1)} \cdot \frac{\rho_u^i - 1}{\rho_u - 1} \tag{5.5}$$

where,

$$\rho_l = \frac{(i-1) \cdot (n-1)}{N}$$
$$\rho_u = \frac{(i-1) \cdot (n-1)}{N-i+1}.$$

Furthermore, if $i = \lfloor c \cdot n \rfloor$ *then,*

$$E(\tau_0(i)) = O((2c)^{cn})$$

Proof. We have seen before that by the strong Markov property we can write,

$$E(\tau_0(i)) = \sum_{k=0}^{i-1} E(\tau_k(k+1)).$$

First we note that $E(\tau_0(i)) \geq E(\tau_{i-1}(i))$.

To see that $E(\tau_k(k+1)) \leq E(\tau_{i-1}(i))$ for $k \leq i-1$, without filling in all the details, one simply has to realize that when k edges are present the birth-rate $\lambda_k = \frac{(N-k)\cdot\beta}{n-1}$ is larger than $\lambda_i = \frac{(N-i)\cdot\beta}{n-1}$, as well as the death-rate being smaller. Hence $E(\tau_0(i)) = \sum_{k=0}^{i-1} E(\tau_k(k+1)) \leq i \cdot E(\tau_{i-1}(i))$. Then (5.5) follows from Corollary 5.2.

To prove that $E(\tau_0(i)) = O((2c)^{cn})$—$i = \lfloor c \cdot n \rfloor$—we first note that,

$$\frac{i(n-1)}{\beta(N-i+1)} \cdot \frac{1}{\rho_u - 1} \leq \frac{i(n-1)}{\beta(N-i+1)} \cdot \frac{1}{\rho_u} = \frac{i}{\beta(i-1)} \to \frac{1}{\beta} \text{ as } n \to \infty.$$

Thus we can conclude that $E(\tau_0(i)) = O(\rho_u^i - 1) = O(\rho_u^i)$
Left to show is that $\rho_u^i = O((2c)^i)$
We have that,

$$\begin{aligned}
\rho_u^i &= \left(\frac{(i-1)\cdot(n-1)}{(N-i+1)}\right)^i \leq \left(\frac{i \cdot (n-1)}{N-i}\right)^i \\
&= \left(\frac{2\lfloor c \cdot n \rfloor \cdot (n-1)}{n \cdot (n-1) - 2\lfloor c \cdot n \rfloor}\right)^i \leq \left(\frac{2c \cdot n \cdot (n-1)}{n \cdot (n-1) - 2c \cdot n}\right)^i \\
&= \left(\frac{2c}{1 - \frac{2c}{n-1}}\right)^i = (2c)^i \left(\frac{1}{1 - \frac{2c}{n-1}}\right)^i \overset{1.)}{=} O((2c)^{cn}).
\end{aligned}$$

1.) We have that $\left(\frac{1}{1-\frac{2 \cdot c}{n-1}}\right)^i \to e^{2c^2}$ □

5.4 A Lower Bound

We are finally ready to combine the results of this section to provide a lower bound for $P(C(t) > \epsilon \cdot n)$, one of the major results of this thesis.

Proposition 5.4. *Let $\epsilon, \delta \in (0,1)$ and $\eta \in (0, 1-\epsilon)$. Then for large enough n we have that,*

$$P(C(t) > \epsilon \cdot n) > (1-\delta) \cdot (1 - \frac{\frac{k\cdot(n-1)}{\beta\cdot(N-k+1)} \cdot \frac{\rho_u^k - 1}{\rho_u - 1}}{t}) = (1-\delta)\cdot(1 - \frac{O((2c')^{c'n})}{t})$$

where,

$$c' = \frac{-\log(1-\epsilon-\eta)}{2\cdot(\epsilon+\eta)}$$
$$k = \lfloor c' \cdot n \rfloor = \lfloor \frac{-\log(1-\epsilon-\eta)}{2\cdot(\epsilon+\eta)} \cdot n \rfloor$$
$$\rho_u = \frac{(k-1)\cdot(n-1)}{N-k+1}.$$

The asymptotics is valid when n grows.

Proof. By Proposition 5.1 we have that,

$$P(C(t) > \epsilon \cdot n) > P(C_1(n,k) > \epsilon \cdot n) \cdot (1 - \frac{E(t_k)}{t})$$

and we have seen , by Proposition 5.4, that for $c' : \ A(c') = \epsilon + \eta \iff c' = \frac{-\log(1-\epsilon-\eta)}{2\cdot(\epsilon+\eta)}$ we have,

$$\lim_{n\to\infty} P(C(n, \lfloor c' \cdot n \rfloor) > \epsilon \cdot n) = 1.$$

Which of course implies that for any given $\delta > 0$ we have,

$$\exists n_0 : \ n > n_0 \implies P(C(n, \lfloor c' \cdot n \rfloor) > \epsilon \cdot n) > (1-\delta).$$

Let t_i be the time it takes for the dynamic graph to go from 0 to i edges, then by Corollary 5.3 we have that for $i < \frac{N}{2}$,

$$E(t_i) < \frac{i\cdot(n-1)}{\beta\cdot(N-i+1)} \cdot \frac{\rho_u^i - 1}{\rho_u - 1} \iff 1 - \frac{E(t_i)}{t} > 1 - \frac{\frac{i\cdot(n-1)}{\beta\cdot(N-i+1)} \cdot \frac{\rho_u^i - 1}{\rho_u - 1}}{t}$$

for $i < \frac{N}{2}$. Using this with $k = \lfloor c' \cdot n \rfloor = \lfloor \frac{-\log(1-\epsilon-\eta)}{2\cdot(\epsilon+\eta)} \cdot n \rfloor$ the assertion is proven true, since for large enough n we have that $\lfloor c' \cdot n \rfloor < \frac{N}{2} = \frac{n(n-1)}{4}$

The asymptotics follows from Corollary 5.3 ☐

5.5 Conclusion and Discussion

In closing this section, and this report, we conclude that the time it takes for the size of the largest component in a critical dynamic Erdős-Rényi graph is bounded from below in a manner described explicit in Proposition 5.4.

There is obvious ways in which this bound can be improved. For instance the bound is only good for $t > E(t_{\lfloor c' \cdot n \rfloor})$ which is a consequence of us using the Markov inequality.

In studying the inequality in Lemma 5.1,

$$
\begin{aligned}
&P(C(t) > \epsilon \cdot n) = P(C(t) > \epsilon \cdot n | t_k \leq t) \cdot P(t_k \leq t) + P(C(t) > \epsilon \cdot n | t_k > t) \cdot P(t_k > t) \\
&> P(C(t) > \epsilon \cdot n | t_k \leq t) \cdot P(t_k \leq t) > P(C_1(n,k) > \epsilon \cdot n) \cdot P(t_k \leq t) \\
&= P(C_1(n,k) > \epsilon \cdot n) \cdot (1 - P(t_k > t)) > P(C_1(n,k) > \epsilon \cdot n) \cdot (1 - \frac{E(t_k)}{t})
\end{aligned}
$$

one could spend time in finding a lower bound for $P(t_k < t)$ directly, instead of bounding it by $P(t_k < t) > 1 - \frac{E(t_k)}{t}$ — the way we opted to do this. This is indeed theoretically possible, since it is possible to describe the exact distribution of $P(t_k \leq t)$, see [7, Th. 1.1]. However, this may be practically intractable since it turns out that t_k is distributed as a sum of independent exponential random variables with the positive eigenvalues of $-Q^n$ as parameters, where Q^n is the generator matrix of $e(t)$ up to state n — I say practically intractable since eigenvalues are notoriously difficult to compute. However careful study of these eigenvalues might reveal that there is approximations that can be done in order to improve on the bound given in this thesis.

Another quite glaring drawback of the lower bound is that we only proved it for large enough n, i.e. we did not provide a $n_0 : \ n > n_0 \implies$ (bound is valid for such n). This is however possible, but requires a lot more work — and might also be of little use since is it known that speed for which $\frac{C(n,c \cdot n)}{n} \xrightarrow{p} A(c)$ is fast.

In deriving bounds for the time it takes for the dynamic graph to go from 0 to i edges—$E(\tau_0(i))$, $i < \frac{N}{2}$—we sacrificed tightness for tractability, and shall therefore give an additional heuristic argument that outlines another approach in approximating $\tau_0(i)$, that might lead to better bounds.
As stated earlier—many times—the process representing the number of edges present in the dynamic graph $e(t)$ is ergodic. Its stationary distribution $\{\pi(k),\ k = 0, 1, \ldots, N\}$ is given by,

$$
\pi(k) = \binom{N}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{N-k}.
$$

Since $e(t)$ is ergodic these probabilities represent the long term proportion of times spent in these states. The process is also a regenerative process, meaning

that there exists random times—cycles—for which the process probabilistically restarts itself. Starting with i edges the return time $\tau_i(i)$ is such a time. We shall argue that there is a approximate relation between the cycle time $\tau_i(i)$ and $\tau_0(i)$, if $i > \frac{n}{2}$. Again, we reiterate that this will be an heuristic argument.
Using the key renewal theorem one can derive a exact expression for $\tau_i(i)$ since,

$$\pi(i) = \text{long run proportion of time in state i}$$
$$= \frac{E(\text{proportion of time in state i in an i-i cycle})}{E(\text{length of i-i cycle})}$$

which then is equivalent to, letting T_i be the holding time in state i,

$$\pi(i) = \binom{N}{i}\left(\frac{1}{n}\right)^i\left(1-\frac{1}{n}\right)^{N-i} = \frac{E(T_i)}{E(\tau_i(i))}.$$

Using that $E(T_i) = \frac{n-2}{\beta\cdot(N+i\cdot(n-2))}$ we can solve for $E(\tau_i(i))$ and arrive at,

$$E(\tau_i(i)) = \frac{n-2}{\beta\cdot(N+i\cdot(n-2))\cdot\pi(i)}.$$

We have seen that the dynamic graph enters stationarity relatively fast—of order $\log(n)$ —and then stays there, and the same then holds true for the edge processes $e(t)$, i.e. the majority of time is spent in stationarity.
When the process leaves state $i > \frac{n}{2}$ it therefore—with high probability — quickly returns to a typical configuration of edges, which would be around $\frac{n}{2}$. So intuitively the cycle time $\tau_i(i)$ is dominated by the time $\tau_{n/2}(i)$. Hence, we think the following approximation is reasonable,

$$E(\tau_{n/2}(i)) \approx E(\tau_i(i)).$$

Again, we use the argument that the edge process enters stationarity quickly, and therefore reaches $\approx \frac{n}{2}$ edges quickly — so the hitting time $\tau_0(i)$ should be dominated by $\tau_{n/2}(i)$. We are lead to the following approximation,

$$E(\tau_0(i)) \approx E(\tau_{n/2}(i)) \approx E(\tau_i(i)) \implies E(\tau_0(i)) \approx \frac{n-2}{\beta\cdot(N+i\cdot(n-2))\cdot\pi(i)}.$$

In order to simplify things a bit this we note that,

$$\left(\frac{1}{n}\right)^i\left(1-\frac{1}{n}\right)^{N-i} = \left(1-\frac{1}{n}\right)^N (n-1)^{-i} \sim e^{-\frac{n}{2}}(n-1)^{-i}$$

and therefore suggest the following approximation for $E(\tau_0(i))$,

$$E(\tau_0(i)) \approx \frac{(n-2)\cdot(n-1)^i\cdot e^{\frac{n}{2}}}{\beta\cdot(N+i(n-2))\cdot\binom{N}{i}} \approx \frac{(n-1)^i\cdot e^{\frac{n}{2}}}{\beta\cdot(\frac{n}{2}+i)\cdot\binom{N}{i}}.$$

6 References

[1] Avner Friedman (1982), *Foundations of Modern Analysis.* Dover Publications, New York

[2] Sheldon M. Ross (2007), *Introduction to Probability Models, 9th edition.* Elsevier Inc.

[3] Robert B. Ash (2000), *Probability and Measure Theory, Second edition.* Harcourt Academic Press, San Diego

[4] Wilhelm Huisinga, Eike Meerbach (2005), *Markov Processes for Everyone* `http://compphysiol.mi.fu-berlin.de/teaching/downloads/SS05_MP/MarkovProzessesScript_FuerDieKlausur.pdf` Last accessed 2015-10.

[5] P. Erdös, A. Renyi (1960), *On The Evolution of Random Graphs* `http://www.noar.technion.ac.il/images/attachments/Scitech/erdos-renyi.pdf` Last accessed 2015-10.

[6] James. A. Fill (1991) *Time to Stationarity for a Continuous-Time Markov Chain, Probability in the Engineering and Informational Sciences*, Vol. 5, Issue 1, pp 61 - 76

[7] Yu Gong, Yong-Hua Mao (2010), *Hitting Time Distributions for Denumerable Birth and Death Processes* `http://arxiv.org/pdf/1007.1039.pdf` Last accessed 2015-10.