Lower and Upper Bounds for the Survival of Infinite Absorbing Markov Chains

Sanjiban Choudhury

CMU-RI-TR-05-04

February 2015

Robotics Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

© Carnegie Mellon University

Abstract

A wide variety of problems in the domain of directed percolation theory, contact process, voter models rely on the analysis of a infinite absorbing Markov chain. Although results for these problems have been obtained through other approaches, reasoning about the Markov chain gives rise to elegant results that are often *explicit* functions of the chain parameters. This allows results to be re-used across a wide variety of problems. In this paper, we present results for the survival of a class of discrete time Markov processes whose states are finite sets of integers. We present lower bounds using two different approaches as well as an upper bound. We borrow varied techniques and show in detail how they can be used to analyse Markov chains of this nature. The results presented in this paper can be used to derive many results in percolation theory in a completely independent manner.

Contents

1 Problem Statement

1.1 The Markov chain

Consider the discrete time chain A_n on the finite subsets of integers. Let p and q be the parameters of the chain. Given the set A_n, the probability of $\{x \in A_{n+1}\}$ is dependent only on A_n. Let $\mathcal{P}$ be the Markov chain dynamics defined as

$$P(x \in A_{n+1}|A_n) = \begin{cases} q & \text{if } |A_n \cap \{x, x+1\}| = 2 \\ p & \text{if } |A_n \cap \{x, x+1\}| = 1 \\ 0 & \text{if } |A_n \cap \{x, x+1\}| = 0 \end{cases} \quad 0 \le p \le q \le 1 \tag{1}$$

Fig. 1 illustrates the Markov chain.

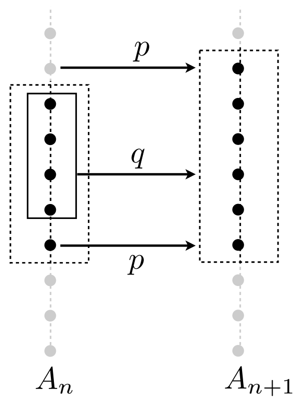

Figure 1: The discrete time Markov chain shown for time steps n and $n+1$. The dotted box shows elements belonging to the current time step. Elements with a consecutive neighbour in the set is transferred with probability q. Elements with only one neighbour is transferred with p.

There are some characteristic features about the chain in (1) that creates difficulty in its analysis. The state space of the Markov chain is a countably infinite dimensional vector s. An element of this vector $s(i) = \{0, 1\}$, where $i \in \mathbb{Z}$ and $\mathbb{Z}$ being the space of integers, indicates if the integer i belongs to the set. Thus at any time n, A_n is a collection of subsets on the integer line.

Before delving into the problem, we try to gain some intuition about this chain.

1. The chain has a state $A_n = \emptyset$ which is *absorbing*. Once the system enters this state, it remains in this state forever. The remaining states are called *transient*.

2. The chain can only accumulate 1 new element in every transition. This occurs with a probability of p.

3. The chain can lose any number of elements in one transition. That is every transient state can enter the absorbing state in one time step with a non-zero probability.

4. For every subset belonging to A_n, all but one end is carried over to the next step with probability q, while one end is carried over with p.

1.2 Survival of the chain

We define survival of the chain as the system never entering the absorbing state.

Definition 1: *The probability of survival of the Markov chain is defined as* $\alpha(p) = \lim_{n\to\infty} P(A_n \neq \emptyset)$

The probability of survival is defined by a relation between q and p, defined as $q(p)$. We define *critical threshold*, $\theta(p)$ as follows

Definition 2: *The probability of survival of the Markov chain is defined as* $\theta(p) = \sup_{\alpha(p)=0} q(p)$

For $q < \theta(p)$, the chain dies else survives. Since directly finding an analytic expression for $\theta(p)$, we will try to find a tight upper and lower bound for

Problem 1: *Find the lower bound* $\mathcal{L}(p)$ *and upper bound* $\mathcal{U}(p)$ *such that*

$$\mathcal{L}(p) \leq \theta(p) \leq \mathcal{U}(p)$$

In this paper, we will prove the following result

Theorem 1. *For the Markov Chain defined in (1)*

(i) If $q < 2(1-p)$*,* A_n *dies.*

(ii) If $q \geq 4p(1-p)$ *and* $\frac{1}{2} < p \leq 1$*,* A_n *survives.*

2 Coupling with the Interval Process

In this section, we will define the *interval process*. This is a simplification of the Markov chain defined in (1) and we will show how coupling to this process will allow the computation of a lower bound.

2.1 The interval process

Let $\mathcal{I}(.)$ be an interval operator defined as follows

$$\mathcal{I}(A) = \arg\max_{a,b} |a-b| \,, \text{s.t } x \in [a,b],\ \forall x \in A \tag{2}$$

An interval process is obtained by applying $\mathcal{I}(.)$ on the Markov chain defined in (1), such that $I_n = \mathcal{I}(A_n)$. Fig. 2 illustrates the process.

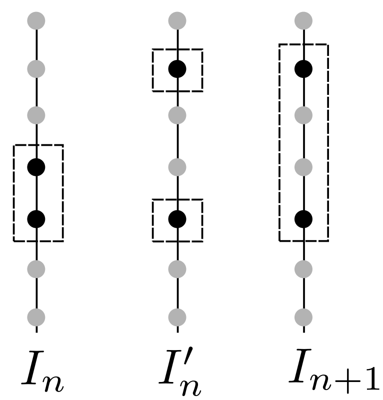

Figure 2: The interval process. The interval is shown by the dotted box. Elements on the number line belonging to a set is shown with a black dot while elements outside the set is shown in a gray dot. On applying the Markov dynamics to I_n, we get a collection of intervals I'_n. On applying the interval operator, we get an interval I_{n+1}.

On applying $\mathcal{P}(.)$ to I_n we get a collection of intervals I'_n. Then on applying $\mathcal{I}(.)$ on I'_n, we get I_{n+1}.

2.2 Coupling between two processes

Theorem 2. *Given the definition of the Markov chain (1) and the interval operator in (2), $A_n \subset I_n$ for all n.*

Proof. At time n, $I_n = \mathcal{I}(A_n)$ which implies $A_n \subset I_n$. On applying $\mathcal{P}$ to both A_n and I_n, we get A_{n+1} and I'_n. Since $I_{n+1} = \mathcal{I}(I'_n)$, which implies $I'_n \subset I_{n+1}$. Hence $A_{n+1} \subset I_{n+1}$. □

Theorem 2 implies that if I_n dies, A_n will die.

2.3 Size of the interval process

Let L_n be the size of I_n at time n. Note that L_n is also a Markov chain. The state $L_n = 0$ is the absorbing state. Let $\theta_I(p)$ be the critical probability

$$\theta_I(p) = \sup_{\lim_{n\to\infty} P(L_n \neq 0)=0} q(p) \tag{3}$$

Then $\theta_I(p) \leq \theta(p)$. Thus a valid lower bound $\mathcal{L}(p) = \theta_I(p)$.

3 Lower Bound using a Random Walk with Negative Drift

In this section we prove the lower bound $\mathcal{L}(p)$ in (1) by proving conditions for *negative drift* of the random walk on L_n to find $\theta_I(p)$ in (3).

3.1 Drift in a random walk

Let $L_0, L_1, \ldots$ be a discrete time random walk on the positive integer line such that $L_i \in \mathbb{Z}_+$. Then drift $\mu(n)$ at time n is defined as follows

Definition 3: *The drift of L_n is defined as $\mu(n) = \mathbb{E}(L_{n+1} - L_n)$*

When $\mu(n) < 0$, the random walk is said to have *negative drift* at time n. For the interval process defined in Section 2, $L_0 = 0$ is an absorbing state. If $\mu(n) < 0$ for all n, the chain is said to die. Thus $\theta_I(p)$ is the critical threshold for which $\mu(n) < 0$ for all n.

3.2 Critical threshold

Theorem 3. *If $q < 2(1-p)$, then the random walk L_n has negative drift $\mu(n) < 0$ for all n.*

Proof. We will first try to define the Markov chain dynamics for L_n. The transition can be divided into 3 cases.

Case I: Both end points of the interval are accepted

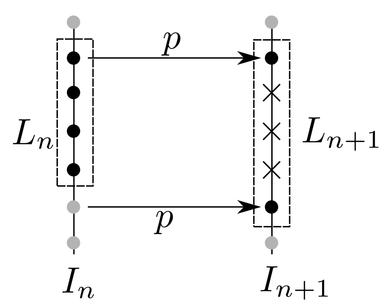

Figure 3: Case I: Both end points of the interval are accepted. The $\times$ denote that the interim elements are irrelevant.

If both ends of the interval are accepted [1], the other points of the interval are irrelevant. Thus the length $L_{n+1} = L_n + 1$ as shown in Fig. 3. This event occurs with probability p^2.

[1]Note that by end points we mean one end being the end of the interval and the other end being the element just outside of the interval. For clarity refer to the Markov chain dynamics (1)

Case II: Only one end point of the interval is accepted

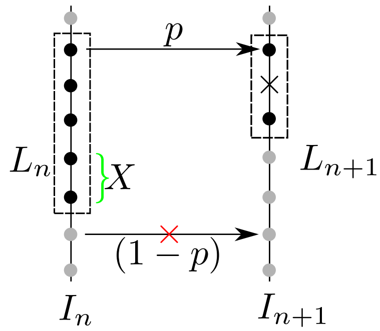

Figure 4: Case II: Only one end point of the interval is accepted. X is a random variable showing elements eliminated from one edge.

One of the endpoint is accepted while the other one is rejected. From the endpoint that is rejected, let X denote the number of elements that is rejected. Thus the length is $L_{n+1} = L_n - X$ as shown in Fig 4. This event occurs with probability $\binom{2}{1}p(1-p) = 2p(1-p)$.

The distribution that X is drawn from is

$$P(X = k) = \begin{cases} q & k = 0 \\ (1-q)q & k = 1 \\ (1-q)^m q & k = m \end{cases} \tag{4}$$

Case III: Neither end points of the interval are accepted

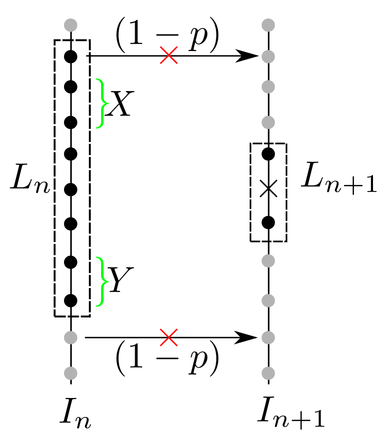

Figure 5: Case III: Neither end pointsof the interval are accepted. X, Y are random variables showing elements eliminated from either ends.

In this case, both end points are rejected. Let X and Y denote the number of elements rejected from either ends. Thus the length is $L_{n+1} = L_n - 1 - X - Y$ as shown in Fig 5. This event occurs with probability $(1-p)^2$. X and Y are drawn from the distribution (4).

The drift in random walk on L_n is $\mathbb{E}(L_{n+1} - L_n)$. Since $\mid \mathbb{E}(L_{n+1} - L_n) \mid$ monotonically increases with n, we want to find the limiting condition of

$$\lim_{n\to\infty} \mathbb{E}(L_{n+1} - L_n) < 0 \tag{5}$$

which implies that $\mathbb{E}(L_{n+1} - L_n)$ for all n.

From the 3 cases and taking the limit $n \to \infty$, we have

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}(L_{n+1} - L_n) = \quad & p^2 + 2p(1-p)[-\mathbb{E}(X)] \\ & + (1-p)^2[-E(X) - E(Y) - 1]\end{aligned}$$

Evaluating $E(X)$ for $n \to \infty$, we have

$$E(X) = E(Y) = \sum_{i=0}^{\infty} iq(1-q)^i = \frac{1-q}{q}$$

Enforcing the condition (5)

$$\begin{aligned} p^2 - (1-p)^2 - [2p(1-p) + 2(1-p)^2]\frac{1-q}{q} &< 0 \\ -1 + 2p - [2-2p]\frac{1-q}{q} &< 0 \\ q &< 2(1-p) \end{aligned}$$

Thus the lower bound $\mathcal{L}(p) = \theta_I(p) = 2(1-p)$. $\square$

4 Lower Bound using Generating Functions

In this section, we provide an alternate proof for the lower bound $\mathcal{L}(p)$ in (1) by enforcing the conditions for the Markov chain to not enter an *absorbing* state. The probability of being in a *transient* state is derived and then solved for using generating functions. We will show that for a nonzero probability of remaining in a transient state constrains a generating function to have radius of convergence as 1, which is then used to solve for $\theta_I(p)$ in (3).

4.1 Probability of being in a transient state

We briefly review the theory stated by Feller [2]. The interval process has two states, the *absorbing* and *transient*. The absorbing states are closed irreducible states, i.e, once the system enters these states it does not leave. The remaining states are transient as they have a non zero probability of entering the absorbing states.

The transition probability P of the interval process can be expressed as

$$\begin{bmatrix} 1 & 0 \\ \underbrace{R}_{\substack{\text{transition to}\\ \text{absorbing states}}} & \underbrace{Q}_{\substack{\text{transition to}\\ \text{transient states}}} \end{bmatrix}$$

Consider the substochastic matrix Q. Q is obtained by deleting all rows and columns corresponding to absorbing states. Each element of Q, $q_{ij} \geq 0$ and $\sum_i q_{ij} \leq 1$.

Let σ_i^n be the sum of the i^{th} row of Q^n. Then σ_i^{n+1} is

Lemma 1: *The sum of rows σ_i^n satisfies*

$$\sigma_i^{n+1} = \sum_v q_{iv}\sigma_v^n \tag{6}$$

Proof.

$$\begin{aligned} q_{ik}^{n+1} &= \sum_v q_{iv}q_{vk}^n \\ \sum_k q_{ik}^{n+1} &= \sum_k \sum_v q_{iv}q_{vk}^n \\ \sigma_i^{n+1} &= \sum_v q_{iv}\sigma_v^n \end{aligned}$$

□

This relation remains valid for $n = 0$ as $\sigma_v^0 = 1$ for all v. Since Q is substochastic, this means $\sigma_i^1 \leq \sigma_i^0$. This gives rise to the following lemma

Lemma 2: $\sigma_i^{n+1} \leq \sigma_i^n$. *Thus for a fixed i, the sequence $\{\sigma_i^n\}$ decreases monotonically to a limit $\sigma_i \geq 0$.*

Proof. We will prove this by induction. Let $\sigma_i^n \leq \sigma_i^{n-1}$. Then

$$\begin{aligned} \sigma_i^n \leq & \quad \sigma_i^{n-1} \\ \sum_v q_{iv}\sigma_i^n \leq & \quad \sum_v q_{iv}\sigma_i^{n-1} \\ \sigma_i^{n+1} \leq & \quad \sigma_i^n \end{aligned}$$

Since $\sigma_i^1 \leq \sigma_i^0$, this holds for all n. □

σ_i satisfies the following relation

$$\sigma_i = \sum_v q_{iv}\sigma_v$$

Note that by definition σ_i is the probability of a system that started at state i staying in transient state forever. Thus the following lemma can be stated.

Lemma 3: *For a system to stay in a transient state forever, there exists i such that $\sigma_i \neq 0$.*

Now consider the solution of the following system of equations

$$x_i = \sum_v q_{iv}x_v \tag{7}$$

where we are only interested in solution where $0 \leq x_i \leq 1,\ \forall i$. Since $\sigma_i^0 = 1$, this can be re-written as $0 \leq x_i \leq \sigma_i^0$. By induction, this implies $x_i \leq \sigma_i^n$ for all n. Hence $x_i \leq \sigma_i \leq 1$. Thus the maximal solution of (7) is $\{\sigma_i\}$. For survival (staying in transient state forever), the maximal solutions of (7) must not all be zero. This is stated in the following theorem

Theorem 4. *The probability that the system stays in transient state forever is given by maximal solutions of $x_i = \sum_v q_{iv}x_v$ that satisfy $0 \leq x_i \leq 1,\ \forall i$ such that there exists i, $x_i \neq 0$.*

4.2 Finding substochastic matrix Q

To find Q, we need to write P.

$$\begin{aligned} P(0,0) = & \quad 1 & & \text{if dead, remain dead} \\ P(0,i) = & \quad 0 & & \text{if dead, become alive} \\ P(i,0) = & \quad (1-p)^2(1-q)^{i-1} & & \text{from i to dead} \\ P(i,1) = & \quad \underbrace{(1-p)^2}_{\substack{\text{both ends}\\ \text{rejected}}} \underbrace{\binom{i-1}{1}}_{\substack{\text{pick 1 site}\\ \text{from i-1 to}\\ \text{survive}}} \underbrace{q}_{\substack{\text{survival}\\ \text{of site}}} \underbrace{(1-q)^{i-2}}_{\substack{\text{death of}\\ \text{other sites}}} & & + \\ & \quad \underbrace{\binom{2}{1}}_{\substack{\text{pick 1 end}\\ \text{to survive}}} \underbrace{p}_{\substack{\text{survival of}\\ \text{one end}}} \underbrace{(1-p)}_{\substack{\text{death of}\\ \text{other end}}} \underbrace{(1-q)^{i-1}}_{\substack{\text{death of}\\ \text{other sites}}} & & \end{aligned}$$

When $i > j+1$

$$P(i,j) = \underbrace{(1-p)^2}_{\substack{\text{both ends}\\ \text{rejected}}} \underbrace{\binom{i-j}{1}}_{\substack{\text{pick 1 edge}\\ \text{other edge}\\ \text{is j}}} \underbrace{q^2}_{\text{end points}} \underbrace{(1-q)^{i-j-1}}_{\substack{\text{reject points}\\ \text{outside}}} +$$
$$\underbrace{\binom{2}{1}}_{\substack{\text{pick 1 end}\\ \text{to survive}}} \underbrace{p}_{\substack{\text{survival of}\\ \text{one extreme}}} \underbrace{(1-p)}_{\substack{\text{death of}\\ \text{other extreme}}} \underbrace{q}_{\substack{\text{one}\\ \text{limit}}} \underbrace{(1-q)^{i-j}}_{\substack{\text{reject}\\ \text{others}}}$$

When $i = j+1$, both end points selected

$$P(i,j) = p^2$$

Note that for this Q, $\sigma_{i+1} \geq \sigma_i$, i.e., the longer the interval, the more likely it will remain in transient state.

4.3 Generating Functions

We will now invoke a tool, generating functions, to apply to (7). Generating functions are defined as follows

Definition 4: *The generating function for a sequence $\sigma_1, \ldots$ is defined as* $\Sigma(z) = \sum_{k-1}^{\infty} \sigma_k z^k$

For the non-decreasing sequence σ_i, the generating function $\Sigma(z)$ will only converge for $\mid z \mid < 1$.

Generating functions are often used in recurrence relation problems to get explicit solutions. Bishir [1] used this technique to derive lower bounds for a different Markov chain.

4.4 Critical threshold

Theorem 5. *If $q < 2(1-p)$, then $x_i = \sum_v q_{iv} x_v$ has no solutions satisfy $0 \leq x_i \leq 1$, $\forall i$.*

Proof. Let $\Sigma(z)$ be the generating function on the sequence σ_i. Applying $\Sigma(z) = Q\Sigma(z)$ we have

$$\begin{bmatrix} z & z^2 & \ldots \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} z & z^2 & \ldots \end{bmatrix} Q \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \end{bmatrix}$$

Let $\begin{bmatrix} Q_1(z) & Q_2(z) & \ldots \end{bmatrix} = \begin{bmatrix} z & z^2 & \ldots \end{bmatrix} Q$.

$$
\begin{aligned}
Q_1(z) = & \sum_{i=1}^{\infty} z^i P(i,1) \\
= & \sum_{i=1}^{\infty} z^i \left[(i-1)q(1-p)^2(1-q)^{i-2} + 2p(1-p)(1-q)^{i-1}\right] \\
= & \frac{q(1-p)^2}{(z(1-q)-1)^2} - \frac{2pz(1-p)}{(z(1-q)-1)}
\end{aligned}
$$

$$
\begin{aligned}
Q_j(z) = & \sum_{i=1}^{\infty} z^i P(i,j) \\
= & \sum_{i=j}^{\infty} z^i P(i,j) + \underbrace{z^{j-1}p^2}_{i=j-1} \\
= & \sum_{i=j}^{\infty}(i-j)q^2(1-p)^2(1-q)^{i-j-1}z^i + 2p(1-p)q\sum_{i=j}^{\infty}(1-q)^{i-j}z^i + z^{j-1}p^2 \\
= & \frac{q^2(1-p)^2}{1-q}\sum_{i=j}^{\infty}(i-j)((1-q)z)^{i-j} + \frac{2pq(1-p)z^j}{1-z(1-q)} + z^{j-1}p^2 \\
= & z^j\frac{q^2(1-p)^2z}{(z(1-q)-1)^2} - z^j\frac{2pq(1-p)}{z(1-q)-1} + z^j\frac{p^2}{2} \\
= & z^j\left[\frac{q^2(1-p)^2z}{(z(1-q)-1)^2} - \frac{2pq(1-p)}{z(1-q)-1} + \frac{p^2}{2}\right]
\end{aligned}
$$

$$
\begin{aligned}
Q\Sigma(z) = & K\sum_{j=1}^{\infty} z^j\sigma_j + (Q_1(z) - Kz\sigma_1) \\
= & K\Sigma(z) + (Q_1(z) - Kz\sigma_1)
\end{aligned}
$$

Plugging in $\Sigma(z) = Q\Sigma(z)$

$$
\begin{aligned}
\Sigma(z) = & K\Sigma(z) + (Q_1(z) - Kz\sigma_1) \\
= & \frac{Q_1(z) - Kz\sigma_1}{1-K}
\end{aligned}
$$

For radius of convergence to be $\mid z \mid = 1$, the poles of $\Sigma(z)$ should be more than 1. Hence the zeros of $1 - K(z)$ should be more than 1. Let $N(z)$ be the numerator of $1 - K(z)$.

Plugging in $\Sigma(z) = Q\Sigma(z)$

$$
\begin{aligned}
N(z) = & z(z(1-q)-1)^2 - (q^2(1-p)^2z^2 - 2zpq(1-p)(z(1-q)-1) + p^2(z(1-q)^2-1)) \\
= & (z-1)(p^2z - p^2 - 2pqz - q^2z^2 + 2qz^2 - z^2 + z) \\
= & (z-1)(z^2(-q^2+2q-1) + z(p^2-2pq+1) - p^2)
\end{aligned}
$$

The roots are $z = 1$ and

$$\frac{(p^2 - 2pq + 1) \pm \sqrt{((p^2 - 2pq + 1)^2 - 4p^2(q-1)^2)}}{2(q-1)^2} < \quad 1$$

Multiplying and re-arranging

$$\begin{aligned} 4(q-1)^2 q(2p + q - 2) <& \quad 0 \\ q <& \quad 2(1-p) \end{aligned}$$

Thus $q < 2(1-p)$ that there is no maximal solution to (7) such that $0 \leq x_i \leq 1$. □

Thus $\theta_I(p) = 2(1-p)$.

5 Upper Bound using

In this section we prove an upper bound $\mathcal{U}(p)$. This result has been proved by Liggett [3]. We provide a detailed explanation here along with background material omitted by Liggett for brevity.

5.1 Proof Outline

We first outline the general structure of the proof. The intuition is that analyzing the dynamics of A_n in (1) is intractable. This is because if A_n contains m elements, atleast 2^m possibilities must be reasoned about. Instead we find a function $H : \{0,1\}^{\mathbb{Z}} \to \mathbb{R}$ that maps any collection of subsets on the integer line to a scalar value. This function should be such that instead of reasoning about the dynamics of A_n directly, the dynamics of the function is reasoned about. We expand on this below.

Let A be a collection of finite subsets of integers. The objective is to find a function $H(A)$ which satisfies the following properties

Step I: Value of the empty set

$$H(\emptyset) = 1$$

Step II: All non empty set has a smaller value

$$H(A) < 1 \quad \forall\, A \neq \emptyset$$

Note the strict inequality. This in conjunction with Step I means that as long as $H(.)$ can be shown to be less than 1, it implies that the underlying set is not empty.

Step III: $H(.)$ is a martingale for an interval

Let $A = \{1, 2, \dots, n\}$ be an interval on the integer line. Let A' be the set obtained by a 1 step simulation of the Markov dynamics $\mathcal{P}$ (1) starting with A. If p and q satisfy the relation stated in Theorem 1, then $H(A)$ is a martingale

$$\mathbb{E}_{P(A'|A)}(H(A')) = H(A) \quad \forall\, \text{intervals}\, A \tag{8}$$

The objective of this step is to show that under a 1 step simulation, $H(.)$ is a martingale when starting with an interval. Note that this is not the same as the interval process approximation. Even though A is an interval, A' is not. This implies that the expected value of H does not increase (increasing implies the set converges towards the empty set).

Step IV: $H(.)$ is a supermartingale for finite sets

Let A be a collection of finite sets of integers. Let A' be the set obtained by a 1 step simulation of the Markov dynamics $\mathcal{P}$ (1) starting with A. If (8) is true, then $H(A)$ is a supermartingale

$$\mathbb{E}_{P(A'|A)}(H(A')) \leq H(A) \quad \forall \, \text{finitesets}\, A \tag{9}$$

The objective of this step is to show that compared to an interval, starting with a set can lead to a decrease in the function value.

Let A be a collection of finite sets of integers. Let A_n be the n step simulation of the Markov chain (**??**) starting with A. Then from (9) we have

$$\mathbb{E}(H(A_n)) \leq H(A) < 1 \tag{10}$$

Since non-survival implies $\lim_{n\to\infty} H(A_n) = H(\emptyset) = 1$, it contradicts with (10). Thus on satisfying the conditions stated in Theorem 1, A_n survives.

In order to describe $H(.)$, we need to review some theory of discrete renewal process.

5.2 Stationary Renewal Process

Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of positive integer-valued random variables (independent and identically distributed), $X_n > 0$ for $n \geq 1$. Let $(S_n)_{n\in\mathbb{N}}$ be the sequence of partial sums

$$S_n = X_1 + \cdots + X_n \tag{11}$$

and $S_0 = 0$.

An interpretation of X_n is that it is the lifetime of the n^{th} object. After an object's lifetime expires, a new one is installed. This installation is termed as *renewal*. Thus S_n is the time of the n^{th} renewal.

Let $(Y_n)_{n\in\mathbb{N}}$ be a sequence defined as

$$Y_n = n - \max\{S_j : S_j \leq n\} \tag{12}$$

$(Y_n)_{n\in\mathbb{N}}$ can be interpreted as the age process associated with the renewal process. Y_n increases by 1 at every unit of time and is reset to 0 every time a renewal occurs. Thus Y_n represents the age of the current object in service and $Y_n = 0$ corresponds to a renewal event. An illustration of this is shown in Fig. 6.

Let $\{f(n) : n > 1\}$ be a probability density on positive integers such that $\sum_{k=1}^{\infty} f(k) = 1$. The random variables X_n are drawn from the distribution $f(n)$ such that

$$P(X_k = n) = f(n) \tag{13}$$

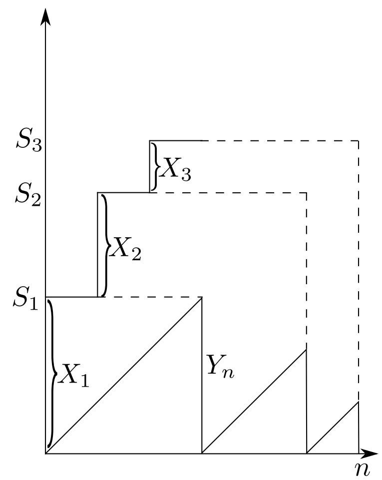

Figure 6: Illustration of a renewal process.

Let $F(n)$ be the tail probabilities $F(n) = \sum_{k=n}^{\infty} f(k)$. Then

$$P(X_k \geq n) = F(n) \tag{14}$$

Now we will show that Y_n is infact a Markov chain whose transition probabilities can be expressed in terms of $f(n)$ and $F(n)$. Let $Y_n = k$. Y_{n+1} can take on two values, $k+1$ or 0. $Y_n = k$ implies the following from (12): $\exists j, S_j = n-k$ and $S_{j+1} \geq n+1$.

If $Y_{n+1} = k+1$, it implies from (12) $S_{j+1} \geq n+2$. From (11), we have $X_{j+1} = S_{j+1} - S_j$. Thus $X_{j+1} \geq k+2$. This implies the following conditional probability from (14)

$$P(Y_{n+1} = k+1 \mid Y_n = k) = P(X_j \geq k+2) = F(k+2)$$

If $Y_{n+1} = 0$, it implies from (12) $S_{j+1} = n+1$. Thus $X_{j+1} = k+1$. This implies the following conditional probability from (13)

$$P(Y_{n+1} = 0 \mid Y_n = k) = P(X_j = k+1) = f(k+1)$$

Thus Y_n is a Markov chain described by

$$P(Y_{n+1} \mid Y_n = k) = \begin{cases} f(k+1) & \text{if } Y_{n+1} = 0 \\ F(k+2) & \text{if } Y_{n+1} = k+1 \end{cases} \tag{15}$$

Note that the transition probabilities are independent of n. Thus this is a *stationary renewal process*.

Let η be a stationary sequence on the integer line such that $\eta(i) \in \{0,1\}, i \in \mathbb{Z}$. Let η be associated with the renewal chain (15) such that $\eta(i) = 1$ if $Y_i = 0$ and $\eta(i) = 0$ otherwise.

Let ν be a *stationary renewal measure* on $\{0,1\}^{\mathbb{Z}}$, i.e., $\nu(\eta) = P(\eta)$. For example, $\eta(x) = 1$ for some x is a sequence and the measure $\nu(\eta : \eta(x) = 1)$ is independent of x due to translation invariance. The sequence can be written as

$$\eta = \{\cdots \times \times 1 \times \times \dots\}$$

where $\times$ indicates the element can take any value from $\{0,1\}$. We will omit the trailing $\times$ and represent this sequence as $\eta = \{1\}$. The measure of this sequence is written as $\nu(1)$.

We can use (15) to define $\nu(11)$ as follows

$$\begin{aligned}\nu(11) &= \nu(11 \mid 1)\nu(1) \\ &= P(Y_{n+1} = 0 \mid Y_n = 0)\nu(1) \\ &= f(1)\nu(1)\end{aligned}$$

Similarly, $\nu(10)$ is

$$\begin{aligned}\nu(10) &= \nu(10 \mid 1)\nu(1) \\ &= P(Y_{n+1} = 1 \mid Y_n = 0)\nu(1) \\ &= F(2)\nu(1)\end{aligned}$$

The same principle can be applied to get

$$\nu(1\underbrace{0\dots0}_{n-1}) = F(n)\nu(1)$$

$$\nu(1\underbrace{0\dots0}_{n-1}1) = f(n)\nu(1)$$

Some results that will be used are given below.

The sequence $\{00\}$ is can be derived from subtracting from 1 all other possibilities

$$1 - \nu(00) = \nu(01) + \nu(10) + \nu(11) \tag{16}$$

The difference between two sequence of different lengths can also be derived. Let $\nu(0^n) = \nu(\times\underbrace{0\dots0}_{n}\times)$.

$$\begin{aligned}\nu(0^{n-1}) - \nu(0^n) &= \nu(\times\underbrace{0\dots0}_{n-1}\times) - \nu(\times\underbrace{0\dots0}_{n}\times) \\ &= \nu(\times1\underbrace{0\dots0}_{n-1}\times) + \nu(\times0\underbrace{0\dots0}_{n-1}\times) - \nu(\times0\underbrace{0\dots0}_{n-1}\times) \\ &= \nu(\times1\underbrace{0\dots0}_{n-1}\times) \\ &= F(n)\nu(1)\end{aligned}$$

Finally, the probability density of a number is just the difference of two tail probabilities.

$$f(n) = F(n) - F(n+1)$$

5.3 Martingale conditions

We will now define $H(A)$ in relation to $\nu(.)$ such that it satisfies the properties stated in subsection 5.1.

$H(A)$ is defined as

$$H(A) = \nu\{\eta : \eta(x) = 0, \forall x \in A\}$$

Thus if $A = \{1, 2, 5, 6\}$, then $H(A) = \nu(\cdots \times 00 \times \times 00 \times \dots)$. If $A = \emptyset$, $H(A) = \nu(\times) = 1$. This is because the sequence $\eta = \{\times\}$ is not constrained at any point and has probability 1.Step I and Step II are satisfied as $\nu(.)$ is a probability measure. The more elements A has, the less the probability, i.e., the smaller $\nu(.)$.

We will now prove the following theorem

Theorem 6. *If $q \geq 4p(1-p)$ and $\frac{1}{2} < p \leq 1$, then* $\mathbb{E}_{P(A'|A)}(H(A')) = H(A)$, $\forall$ intervals A

Proof. Even though the set A is an interval, on simulating through $\mathcal{P}$ it will produce A' which will not be an interval. If A contains m terms, the expectation of A' will require reasoning about atleast 2^m terms if done naively. However, by using certain properties of the probability measure, this can be avoided. To do this we first need to define a *thinning* operation.

5.3.1 Thinning Process

Let ν^* be a measure obtained by thinning ν, i.e., the distribution of

$$\{\eta(i) \wedge \zeta(i)\, i \in \mathbb{Z}\}$$

where η is distributed according to ν and $\zeta(i)$ is i.i.d with $P(\zeta(i) = 1) = q$. Note that ν^* is also a renewal measure. The sequence is said to have thinned by $1-q$ because elements in $\eta(i)$ irrespective of their value is converted to 0 in the new sequence with probability $1-q$. Note that thinning of a sequence makes it more unlikely than the original sequence, hence $\nu^*(\eta) \leq \nu(\eta)$. The reason for introducing ν^* will be clear after we introduce the next tool

5.3.2 Stochastic edge set

We will now define an operation on A to create A^* which is a set with stochastic edges. If $A = [l+1, m]$, then A^* is created by carrying over elements $[l+1, m-1]$ with probability 1 and elements l and m with probability $r = \frac{p}{q}$. Note that this set is or the most part deterministic with only the endpoints having a probability of inclusion. Thus an expectation evaluation from A to A^* only has to deal with 3 terms.

In conjunction with the thinning process, we can write

$$A' = \{k \in A^* : \zeta(k) = 1\}$$

Now we can re-write the expectation of $H(A')$ as follows

$$\begin{aligned} \mathbb{E}H(A') = &\ \mathbb{E}\nu\{\eta : \eta(i) = 0 \quad \forall i \in A'\} \\ = &\ P\left[\eta(i) = 0 \quad \forall i \in A^* \ni \zeta(i) = 1\right] \\ = &\ P\left[\eta(i) \wedge \zeta(i) = 0 \quad \forall i \in A^*\right] \\ = &\ \mathbb{E}\nu^*\left[\eta : \eta(i) = 0 \quad \forall i \in A^*\right] \end{aligned} \tag{17}$$

Now we see that the problem has been transformed to reasoning about the measure ν^*. The set $A*$ allows the expectation to have only 3 terms. The objective is to establish relations between ν^* and ν to enforce the martingale condition.

For $n = 1$, enforcing $H(A) = \mathbb{E}H(A')$ and using (17)

$$\nu(0) = (1-r)^2\nu^*(0) + 2r(1-r)\nu^*(0) + r^2\nu^*(00)$$

Subtracting from 1 we have

$$\begin{aligned} \nu(1) &= 1 - [(1-r)^2\nu^*(0) + 2r(1-r)\nu^*(0) + r^2\nu^*(00)] \\ &= 1 - [(1-r)^2 + 2r(1-r)(1-\nu^*(0)) \\ &+ r^2(1 - \nu^*(10) - \nu^*(01) - \nu^*(11)) \qquad \because \text{using}(16) \\ &= 2r(1-r)\nu^*(1) + r^2(\nu^*(11) + \nu^*(10) + \nu^*(01)) \end{aligned}$$

Now we will use some results proved in the previous subsection. Using the following

1. $\nu^*(11) = \nu^*(1)f^*(1)$
2. $F^*(2) = 1 - f^*(1)$
3. $\nu^*(10) = \nu^*(1)F^*(2)$
4. $\nu^*(1) = q\nu(1)$

we have

$$F^*(2) = \frac{1 - 2qr + qr^2}{qr^2} \tag{18}$$

For n we have

$$\nu(0^n) = (1-r)^2\nu^*(0^{n-1}) + 2r(1-r)\nu^*(0^n) + r^2\nu^*(0^{n+1}) \tag{19}$$

Applying (19) $n-1$ and n and subtracting we have

$$\begin{aligned} \nu(0^{n-1}) - \nu(0^n) = &\ (1-r)^2(\nu^*(0^{n-2}) - \nu^*(0^{n-1})) \\ &+ 2r(1-r)(\nu^*(0^{n-1}) - \nu^*(0^n)) + r^2(\nu^*(0^n) - \nu^*(0^{n+1})) \\ F(n)\nu(1) = &\ \nu^*(1)((1-r)^2F^*(n-1) + 2r(1-r)F^*(n) + r^2F^*(n+1)) \\ F(n) = &\ q((1-r)^2F^*(n-1) + 2r(1-r)F^*(n) + r^2F^*(n+1)) \end{aligned} \tag{20}$$

Since (20) contains both $F(n)$ and $F^*(n)$ terms, a relationship is required to be established from properties of the renewal measure.

$$
\begin{aligned}
\nu^*(1)F^*(n) = \quad & \nu^*(1\underbrace{0\ldots0}_{n-1}) \\
= \quad & \nu(\eta:\{1\times\}\wedge\xi:\{1\times\},\eta(i)\wedge\xi(i)=0,\ \forall\,1\le k<n) \\
= \quad & \nu(\eta:\{1\underbrace{0\ldots0}_{n-1}\}\wedge\xi:\{1\times\}) \\
& +\nu(\eta:\{1\underbrace{0\ldots0}_{n-2}1\}\wedge\xi:\{1\underbrace{\times\cdots\times}_{n-2}0\}) \\
& +\nu(\eta:\{1\underbrace{0\ldots0}_{n-3}1\}\wedge\xi:\{1\underbrace{\times\cdots\times}_{n-3}0\}),\eta(i)\wedge\xi(i)=0,\ \forall\,n-2<k<n) \\
& +\ldots \\
& +\nu(\eta:\{11\}\wedge\xi:\{10\}),\eta(i)\wedge\xi(i)=0,\ \forall\,1<k<n) \\
= \quad & \nu(1)F(n)q \\
& +\nu(1)f(n-1)q(1-q) \\
& +\nu(1)f(n-2)q(1-q)F^*(1) \\
& +\ldots \\
& +\nu(1)f(1)q(1-q)F^*(n-1)
\end{aligned}
$$

Collecting all terms on the right hand side

$$
\begin{aligned}
\nu^*(1)F^*(n) = \quad & \nu^*(1)F(n)+\nu^*(1)(1-q)\sum_{k=1}^{n-1}f(k)F^*(n-k) \\
F^*(n) = \quad & F(n)+(1-q)\sum_{k=1}^{n-1}f(k)F^*(n-k)
\end{aligned}
$$

Substituting in (20)

$$
\begin{aligned}
F(n) = \quad & qr^2(F(n+1)+(1-q)\sum_{k=1}^{n}f(k)F^*(n+1-k)) \\
& +2qr(1-r)(F(n)+(1-q)\sum_{k=1}^{n-1}f(k)F^*(n-k)) \\
& +q(1-r)^2(F(n-1)+(1-q)\sum_{k=1}^{n-2}f(k)F^*(n-1-k))
\end{aligned}
$$

We still havent got rid of F^*. We will use (20) to convert

$$\begin{aligned}
F(n) = \quad & qr^2F(n+1) + 2qr(1-r)F(n) + q(1-r)^2F(n-1) \\
& +(1-q)\sum_{k=1}^{n-2} f(k)(qr^2F^*(n+1-k) + 2qrF^*(n-k) + q(1-r)^2F^*(n-1-k)) \\
& +qr^2(1-q)f(n)F^*(1) + qr^2(1-q)f(n-1)F^*(2) + 2qr(1-r)(1-q)f(n-1)) \\
= \quad & qr^2F(n+1) + 2qr(1-r)F(n) + q(1-r)^2F(n-1) \\
& +(1-q)\sum_{k=1}^{n-2} f(k)F(n-k) \\
& +qr^2(1-q)f(n)F^*(1) + qr^2(1-q)f(n-1)F^*(2) + 2qr(1-r)(1-q)f(n-1))
\end{aligned}$$

$F^*(1) = 1$ and $F^*(2) = \frac{1-2qr+qr^2}{qr^2}$. Using $q = pr$ and simplifying one of the terms

$$\begin{aligned}
& qr^2(1-q)f(n-1)F^*(2) \\
= \quad & qr^2(1-q)f(n-1)\frac{1-2qr+qr^2}{qr^2} \\
= \quad & (1-q)(1-2p+pr)f(n-1)
\end{aligned}$$

Then

$$\begin{aligned}
& qr^2(1-q)f(n-1)F^*(2) + 2qr(1-r)(1-q)f(n-1) \\
= \quad & (1-q)(1-2p+pr)f(n-1) + 2p(1-r)(1-q)f(n-1) \\
= \quad & (1-q)f(n-1)(1-pr) \\
= \quad & (1-q)f(n-1) - pr(1-q)f(n-1) \\
= \quad & (1-q)f(n-1) - pr(1-q)(F(n-1) - F(n))
\end{aligned}$$

Also

$$\begin{aligned}
& qr^2(1-q)f(n)F^*(1) \\
= \quad & qr^2(1-q)(F(n) - F(n+1)) \\
= \quad & pr(1-q)(F(n) - F(n+1))
\end{aligned}$$

Now we collect all terms corresponding to $F(n)$

$$\begin{aligned}
& F(n)(1 - 2qr(1-r) - pr(1-q) - pr(1-q)) \\
= \quad & F(n)(1 - 2p + 2p^2)
\end{aligned}$$

Now we collect all terms corresponding to $F(n-1)$

$$\begin{aligned}
& F(n-1)(-q(1-r)^2 + pr(1-q)) \\
= \quad & F(n-1)(q - p(2-p))
\end{aligned}$$

Now we collect all terms corresponding to $F(n+1)$

$$\begin{aligned}
& F(n+1)(-qr^2 + pr(1-q)) \\
= \quad & F(n+1)(-p^2)
\end{aligned}$$

Putting it all together

$$(1-q)\sum_{k=1}^{n-1} f(k)F(n-k) = \quad (1-2p+2p^2)F(n) \\ +(q-p(2-p))F(n-1) - p^2F(n+1)$$

To eliminate $f(k) = F(k) - F(k+1)$

$$(1-q)\sum_{k=1}^{n-1}(F(k)-F(k+1))F(n-k) = \quad (1-2p+2p^2)F(n) \\ +(q-p(2-p))F(n-1) - p^2F(n+1)$$

Shifting around terms left and right - objective is to prove an invariance

$$\begin{aligned} & (1-q)\sum_{k=1}^{n-1} F(k)F(n-k) + (q-p(2-p))F(n-1) - p^2F(n) \\ = \quad & (1-q)\sum_{k=1}^{n-1} F(k+1)F(n-k) + (1-2p+p^2)F(n) - p^2F(n+1) \end{aligned}$$

Simplifying RHS

$$\begin{aligned} & (1-q)\sum_{k=1}^{n-1} F(k+1)F(n-k) + (1-2p+p^2)F(n) - p^2F(n+1) \\ = \quad & (1-q)\sum_{k=1}^{n} F(k)F(n+1-k) - (1-q)F(n) + (1-2p+p^2)F(n) - p^2F(n+1) \\ = \quad & (1-q)\sum_{k=1}^{n} F(k)F(n+1-k) - (1-q)F(n) + (1-2p+p^2)F(n) - p^2F(n+1) \\ = \quad & (1-q)\sum_{k=1}^{n} F(k)F(n+1-k) + (q-p(2-p))F(n) - p^2F(n+1) \end{aligned}$$

Let $\Gamma(n+1) = (1-q)\sum_{k=1}^{n} F(k)F(n+1-k) + (q-p(2-p))F(n) - p^2F(n+1)$. Then

$$\Gamma(n) = \Gamma(n+1) \tag{21}$$

For $\Gamma(2)$

$$\begin{aligned} \Gamma(2) = \quad & (1-q)F^2(1) + (q-p(2-p))F(1) - p^2F(2) \\ = \quad & (1-q) + (q-p(2-p)) - p^2F(2) \\ = \quad & 1-2p+p^2-p^2F(2) \end{aligned} \tag{22}$$

To evaluate $F(2)$ we use the fact that

$$\begin{aligned} F^*(2) = &\quad F(2) + (1-q)f(1)F*(1) \\ = &\quad F(2) + (1-q)(F(1) - F(2)) \\ = &\quad qF(2) + (1-q) \end{aligned}$$

Substituting in (18), we get $F(2) = \frac{(1-p)^2}{p^2}$. Substituting this in (22) we get $\Gamma(2) = 0$.

Thus the recursive relation (21) results in $\Gamma(n) = 0, \ \forall n$

Rearranging the expression $\Gamma(n+1) = 0$

$$\begin{aligned} p^2F(n+1) = &\quad (1-q)\sum_{k=1}^{n} F(k)F(n+1-k) + (q - p(2-p))F(n) \\ = &\quad (1-q)\sum_{k=1}^{n-1} F(k)F(n+1-k) + (1 - q + q - p(2-p))F(n) \\ = &\quad (1-q)\sum_{k=1}^{n-1} F(k)F(n+1-k) + (1-p)^2F(n) \end{aligned}$$

$$p^2F(n+1) = (1-q)\sum_{k=1}^{n-1} F(k)F(n+1-k) + (1-p)^2F(n) \tag{23}$$

From (23), we have $p^2F(n+1) \geq (1-p)^2F(n)$, which

$$\frac{(1-p)^2}{p^2} \leq \frac{F(n+1)}{F(n)} \leq 1$$

leads to $p \geq \frac{1}{2}$.

From the property of $F(n)$, we have

$$M = \sum_{n=1}^{\infty} F(n) < \infty$$

Summing both sides of (23) from $n = 1$ to ∞ we have

$$p^2(M-1) = (1-q)M^2 + (q - 2p + p^2)M$$

The discriminant of the quadratic equation is $q[q - 4p(1-p)]$ which must be ≥ 0. This leads to $q \geq 4p(1-p)$. □

5.4 Supermartingale proof

(Sketch) Let the finite set be B. We first decompose B into a set of disjoint intervals such that $B = \bigcup_{i=1}^{l} A_i$. For all A_i, A_j which are a distance of atleast 1 apart, the dynamics of the Markov chain do not interact. Hence the interval analysis can be applied independently to each such interval and the martingale equation holds.

We will now show the result for two sets $A_1 = \{i, \ldots, j\}$ and $A_2 = \{j + 1, \ldots, k\}$ that are a distance of 1 element apart. Let A_1^* and A_2^* be the random set transformation when applied to A_1 and A_2 separately. Let A_{12}^* be the random set transformation applied to $A_1 \cup A_2$. Then $P(j \in A_{12}^*) = 1$ while $P(j \in A_1^*) = r$. Thus the expected volume of A_{12}^* is more than when the Markov chain is applied independently. Since the stationary measure $\nu\{\eta : \eta(x) = 0, x \in A\}$ decreases with increasing A.

$$\mathbb{E}_{P(A'|A)}(H(A')) \leq H(A) \quad \forall \, \text{finitesets} \, A \tag{24}$$

For a rigorous proof, refer to Liggett [3].

References

[1] J Bishir. A lower bound for the critical probability in the one-quadrant oriented-atom percolation process. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 401–404, 1963.

[2] Vilim Feller. *An Introduction to Probability Theory and Its Applications: Volume One*. John Wiley & Sons, 1950.

[3] Thomas M Liggett. Survival of discrete time growth models, with applications to oriented percolation. *The Annals of Applied Probability*, pages 613–636, 1995.