Random Walks With Reinforcement: As an Application of Exchangeability

Paul Mollema

Senior Seminar Paper
May 1, 2005
Division of Science and Mathematics
University of Minnesota Morris
moll0088@morris.umn.edu

Abstract: In an effort to cover the topic of Random Walks With Reinforcement, we will first cover the basics, go over some simulations and my findings, then walk through the proof of partial exchangeability.

1. Introduction

Random Walk With Reinforcement. In 1987, D. Coppersmith and P. Diaconis introduced a new way to explore a new city as a simple model. Consider a person walking the streets of a new city. At first, each street is as unfamiliar as any of the others and the walker must randomly choose between them. As the exploration goes on , the more often a street has been traversed in the past, the more familiar it becomes and thus, is more likely to be chosen again. This was using the idea of homesickness and was their motivation for further exploration.

Now to put this in terms of graph theory. Let G be a finite connected graph. Give each edge an initial weight which is generally a positive integer. Furthermore, the initial weights are generally set to 1 for the sake of simplicity. Start the *walker* at some initial vertex, v_0. In each step, the random walker chooses which edge to traverse based on its weight with respect to the weights of the other edges incident to vertex v. For example, let (e_1, e_2, e_3) be the edges incident to v with respective weights of (a, b, c). The probability that we choose the first edge is:

$$P(e_1) = \frac{a}{a+b+c}.$$

Each time that we travel across an edge, we add 1 to its weight. As time goes on, the process prefers to continue crossing edges that it has crossed in the past [8].

In this paper we will be working with Markov chains which basically represent the random motions of an object [4]. More formally, a stochastic process $\{X_n : n = 0, 1, ...\}$ with a finite or countably infinite state space S is said to be a Markov chain, if for all i, j, i_0, ..., $i_{n-1} \in S$, and n = 0, 1, 2, 3,...,

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, ..., X_0 = i_0) = P(X_{n+1} = j | X_n = i) \text{ [6]}.$$

This is simply stating that the probability of our next choice is dependent only on our previous choice. The *state space* is essentially the set of all the places the object can go. As an example, the state space for a particular Markov chain may be $S = \{1, 2, 3\}$ or S could be the set of all positive integers. It is not required that all of the elements of the state space are integers or numbers at all [4]. A *stochastic process* is just a process that proceeds randomly over time.

Furthermore, we can use a transition probability matrix to represent Markov chains. Each entry in such a matrix represents the probability that we transition from state i to state j which will be denoted as $p_{ij} = P(X_{n+1} = j|X_n = i)$. Such a Markov matrix may appear as follows:

$$\mathbf{P} = \begin{pmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ p_{20} & p_{21} & p_{22} & \cdots \\ \vdots & & & \end{pmatrix}$$

Note that for all $i \in S$, the probability of a transition *from* state i is $\sum_{j=0}^{\infty} p_{ij}$. Therefore we must have $\sum_{j=0}^{\infty} p_{ij} = 1$; that is the sum of the elements in each row of the transition probability matrix must be 1. It is, however, not necessary that the column sums be 1. Such a matrix can be used to construct a sample space, associate probabilities to all events of the sample space, and then define a Markov chain over the sample space in such a way that its transition probability matrix is P.

Exchangeability. Consider a random sequence of quantities $\{x_i : i = 1,2,3,...\}$ with associated probability $p(x_1,x_2,...,x_n)$. The sequence is said to be exchangeable if the following property holds true for all finite subsets of the sequence:

$$p(x_1,x_2,...,x_n) = p(x_{\pi(1)},x_{\pi(2)},...,x_{\pi(n)})$$

for all permutations π defined on the set $(1,2,...,n)$. [1]

Let e be and edge with endpoints u and v. We will arbitrarily assign one endpoint as *positive*, say v, and the other, u, as *negative*. We will call the movement along $e$from vertex u to vertex v *positively oriented*, denoted as e^+, and the movement along e from v to u to be *negatively oriented*, denoted as e^-. We will want to keep track of the number of times that each of these happens. To do this we will denote $k_n(e^+)$ as the number of times that the walk traverses e^+, $k_n(e^-)$ is denoted similarly up to time n. Further, we can set $k_n^+ = k_n(e^+; e \in E)$ and $k_n^- = k_n(e^-; e \in E)$

k-path: Consider a sequence $p = (u_0,e_1,u_1,e_2,...,e_n,u_n)$ with n $\geq$1, where $u_i \in V$, $e_i \in E$, and the vertices of e_i are u_{i-1} and u_i for all $i \in \{$1,....,n$\}$. Note that an edge can occur more than once in the sequence. Consider and edge e, clearly the number of times that it has been traversed, k_e, is $k_n^+ + k_n^-$. We call p a (k^+,k^-)-*path* or a *k-path*.

This brings us to the concept of *Partial Exchangeability*. Consider a process $Z_n := (X_1,X_2,...,X_n)$ and consider two admissible paths π and π'. These paths are considered to be equivalent as long as the have the same starting point and satisfy $k_e(\pi) = k_e(\pi')$ for all $e \in E$. We define P to be *partially exchangeable* if $P(Z_n = \pi) = P(Z_n = \pi')$ for any equivalent paths π and π' of length n. Where P is a set of stochastic matrices.

As an example, consider a Markov chain with state space $S = \{0,1\}$. Two chains are called equivalent if the begin with the same symbol and have the same number of transitions between the symbols(0 to 0, 0 to 1, 1 to 0, and 1 to 1). As an example, consider the following sequence: 0101101011. We can easily break up this chain by declaring the end of each block to be the transition from 1 to 0. In terms of block, our chain will now appear as 01|011|01|011. By switching the blocks around we can obtain the following sequences that would be considered equivalent:

0110101011 and 0110110101 and 0101011011

as the all satisfy the following transition matrix:

$$\begin{array}{cc} & \text{to} \\ & \begin{array}{cc} 0 & 1 \end{array} \\ \text{from} \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} 0 & 4 \\ 3 & 2 \end{pmatrix}. \end{array}$$

Note that the first sequence was obtained by switching the first and second blocks of the original sequence. The probability on such binary sequences is considered to be *partially exchangeable* if it assigns the same probability to equivalent strings [2].

In **Section 2**, I will introduce the density function for random walks. In **Section 3**, I will show you how the density functions appear for certain graphs. Finally, in **Section 4**, I will walk you through the proof for *partial exchangeability*.

2. The Density Function

Assume $G = (V, E)$ is a finite connect graph where V and E represent the vertex and edge sets, respectively. Also, let l and m be the respective cardinalities of V and E. For the sake of simplicity, we will assume that G has no loops. However, parallel edges will be allowed. Each edge e is give positive weights and we will denote its set of endpoints as $\bar{e}$. Before the walk begins we will assign non-random weights to each edge $e \in E$ denoted as $a_e > 0$. At time n the weight of edge e will denoted by $w_n(e)$ and the sum of the edges incident to vertex v as $w_n(v)$ [9].

We can define a random walk with reinforcement starting at $v_0 \in V$ as a sequence $X_0, Y_1, X_1, Y_2, X_2, \ldots$ where each X_i takes on values in V, each Y_i takes on values in E, also $\bar{Y_i}$ = $\{X_{i-1}, X_i\}$ for all $i \in$ {1, 2,...}. It further holds that $P(X_0 = v_0)$ = 1. The probability that $Y_{n+1} = e$ and $X_{n+1} = v$ given that we have taken n steps and are currently at Y_n and X_n, is the weight,$w_n(e)$, of edge e divided by the sum of all the edges incident to vertex, X_n that we are currently on. This is demonstrated by the following equation:

$$P(Y_{n+1} \quad = \quad e, X_{n+n1} = v | X_0, Y_1, X_1, \ldots, Y_n, X_n) = \begin{cases} \frac{w_n(e)}{w_n(X_n)} & \text{if } \bar{e} = \{X_n, v\} \\ 0 & \text{otherwise.} \end{cases}$$

Also, the weights satisfy

$$w_0(e) \quad = \quad a_e$$

and

$$w_{n+1}(e) \quad = \quad \begin{cases} w_n(e) + 1 & \text{if } Y_{n+1} = e \\ w_n(e) & \text{otherwise.} \end{cases}$$

We can denote the number of times that edge e was traversed during the walk up to time n as:

$$k_n(e) = |\{i \in \{1, 2, ..., \text{n}\}: Y_i = e\}|$$

We can denote the proportion of the weight of e up to time n as $a_n(e) := \frac{k_n(e)}{n}$. It is clear to see that all of the coordinates of a_n are nonnegative and sum to 1 Further they take on values in the $(m-1)$-dimensional simplex

$$\Delta := \left\{ x = (x_e; e \in E) \in [0,1]^E : \sum_{e \in E} x_e = 1 \right\}.$$

The sum of the initial weights of the edges incident to vertex v is denoted by a_v, and the sum of the proportional weights of the edges incident to v as x_v, where:

$$a_v = \sum_{\{e: v \in \bar{e}\}} a_e \quad (1)$$

$$x_v = \sum_{\{e: v \in \bar{e}\}} x_e \quad (2)$$

let $c1, c2..., c_{m-l+1}$ be a fundamental set of cycles of G and orient them in an arbitrary way. Where a fundamental set of cycles is a set of cycles that can be used to represent all the cycles in the graph. We can define for $x \in \Delta$ an $(m-l+1) \times (m-l+1)$-dimensional matrix $A(x) = (a_{i,j}(x))$ by

$$a_{i,i}(x) = \sum_{e \in c_i} \frac{1}{x_e}, \quad a_{i,j}(x) = \sum_{e \in c_i \cap c_j} \pm \frac{1}{x_e} \quad \text{for} \quad i \neq j.$$

Theorem 2.1 [2] *For random walks with reinforcement on a finite graph, $G = (V, E)$:*
(a) the sequence a_n converges almost surely to a limit with respect to the surface measure on Δ
(b) the limit is random with absolutely continuous distribution having density:

$$\phi(x) = C \frac{2^{1-l+\sum_{e \in E} a_e}}{(m-1)! \pi^{\frac{m}{2}}} \frac{\prod_{e \in E} x_e^{a_e - \frac{1}{2}}}{x_{v_0}^{\frac{a_{v_0}}{2}} \prod_{v \in V/\{v_0\}} x_v^{\frac{a_v+1}{2}}} \sqrt{\det(A(x))} \quad (3)$$

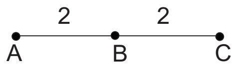
where

$$C = \frac{\Gamma(\frac{a_{v_0}}{2}) \prod_{v \in V/\{v_0\}} (\Gamma(\frac{a_v+1}{2}))}{\prod_{e \in E} \Gamma(a_e)}$$

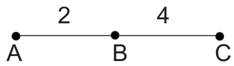
The above equation of course uses our definitions in (Eq. 1) and (Eq. 2).

3. Densities on Selected Graphs

3.1 Line segment with 2 edges and 3 vertices:
Consider a simple graph with all initial edge weights equal to 2:



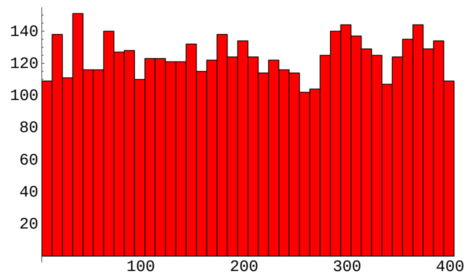
Let the random walk start at point B, so the *walker* must choose between vertices A and C with probability 1/2. Say that the *walker* choose vertex C. Now the only choice that the choice to be made is to return back to vertex B. Now the graph will appear as follows:



Now at vertex B the probability of choosing A is 1/3 and the probability of choosing vertex C is 2/3.

Taking a closer look at this particular problem, the edge weights on this graph will behave the same as the balls in *Polya's Urn*. In the urn is two colored balls, say black and white. With each turn we select one ball, then replace that ball into the urn along with another of the same color. As the number of balls drawn grows, the density approaches a uniform distribution. For this reason we should expect to obtain a uniform distribution when from (Eq. 3) for this particular graph.

For this simulation I took a total of 400 steps for each simulation. Then I ran 5000 simulations and collected the end weights for each of the edges. After I had obtained all of the data I made a histogram of the data on the first edge to take a little at how the weights are distributed:



As you can see, the distribution looks fairly uniform, however not exactly uniform this simulation does not run off to infinity.

Now we will take a look at (Eq. 3) to see how it behaves. First, $C = \frac{\Gamma(\frac{2+2}{2})}{\Gamma(2)\cdot\Gamma(2)} = 1$. After plugging in all of the appropriate numbers into (Eq. 3) we find that the density is $\frac{4}{\pi}$ for all x.

We will now take a look at applying (Eq. 3) to this example. Let x and y be the respective proportional weights of A,B and B,C at time n. The initial weights are all 2, so $a_e = 2$ for both edges. Also, $x_1 = x$ and $x_2 = y$ in this case. Now we will look at the various components of $\phi(x)$:

$$2^{1-l+\sum_{e\in E} a_e} = 2^{1-3+(2+2)} = 4$$

$$\prod_{e\in E} x_e^{a_e - \frac{1}{2}} = x^{3/2} y^{3/2}$$

$$x_{v_0}^{\frac{a_{v_0}}{2}} = (x+y)^2$$

and

$$\prod_{v \in V/\{v_0\}} x_v^{\frac{a_v+1}{2}} = x^{3/2}y^{3/2}$$

Plugging all of this into (Eq. 3) we get

$$\begin{aligned}
\phi(x) &= C\frac{2^{1-l+\sum_{e\in E} a_e}}{(m-1)!\pi^{\frac{m}{2}}} \frac{\prod_{e\in E} x_e^{a_e-\frac{1}{2}}}{x_{v_0}^{\frac{a_{v_0}}{2}} \prod_{v\in V/\{v_0\}} x_v^{\frac{a_v+1}{2}}} \sqrt{\det(A(x))} \\
&= \frac{4}{\pi}\frac{x^{3/2}y^{3/2}}{(x+y)^2x^{3/2}y^{3/2}} \\
&= \frac{4}{\pi(x+y)^2} \\
&= \frac{4}{\pi} \quad \text{because } (x+y) = 1 \text{ by definition}
\end{aligned}$$

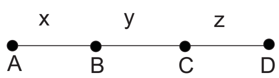
Let a and b be the weights of edges $\{$A,B$\}$ and $\{$B,C$\}$ respectively. Here our state space is $\{$A,B,C$\}$ and our Markov matrix will appear as follows:

$$\begin{array}{cc} & \text{to} \\ & \begin{array}{ccc} A & B & C \end{array} \\ \text{from} \begin{array}{c} A \\ B \\ C \end{array} & \begin{pmatrix} 0 & 1 & 0 \\ \frac{a}{a+b} & 0 & \frac{b}{a+b} \\ 0 & 1 & 0 \end{pmatrix}. \end{array}$$

The application of the density function of the following examples will be in **Appendix A**.

3.2 Line segment with 3 edges and 4 vertices:

Consider a simple graph with all initial edge weights equal to one ($a_e = 1\ \forall e \in E$) . At time n, let x, y, and z be the proportional edge weights on the edges such that $x+y+z = 1$. Here is a picture of the graph to help:



For this simulation we will start at Vertex B, take 400 steps, and run the simulation 2000 times. After the data was collected, it was put into tables and the plotted so that we could take a look at the distribution of the random walks. Below are two such plots:

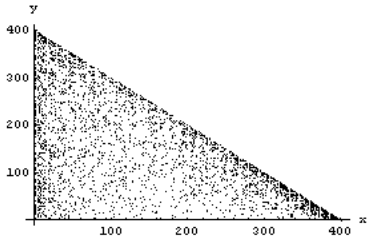


Figure 3.1

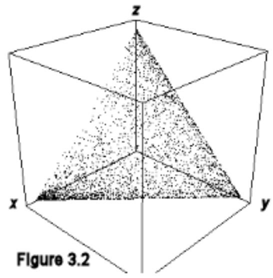


Figure 3.2

As you can see in figure 3.1, there appears to be a lot of clustering around x, y, and the line $x+y=1$ This is because we are start from vertex B for every simulation. We always have to choose x or y first. This brings up the probability of choosing that edge up and it has a tendency to continue to do so. Now taking a look there also seems to some clustering along the line $y+z=1$. It appears that there is a higher tendency to remain on y and z, once y has been chosen, than to return to x. To further illustrate this, here is a three-dimensional histogram of the same distribution:

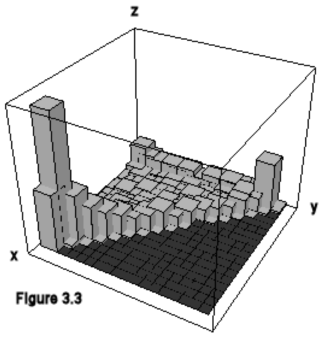


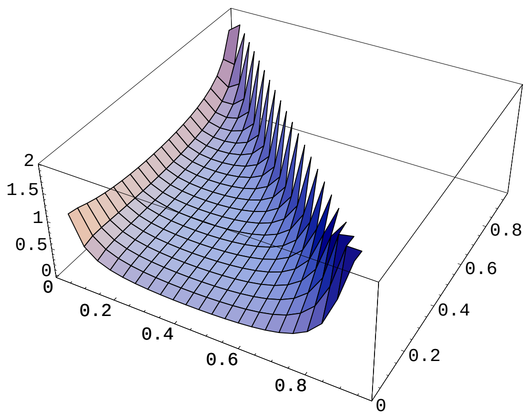
Figure 3.3

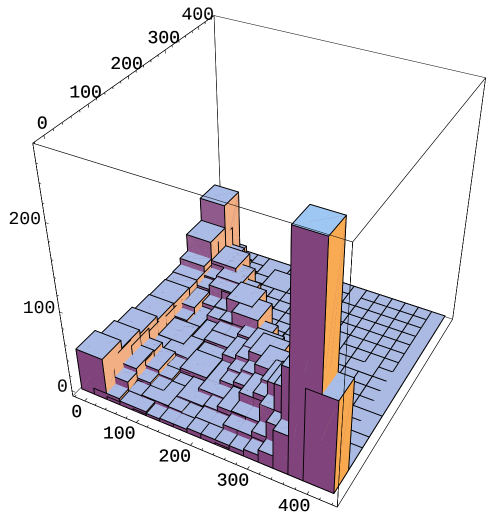
Also we notice from this three-dimensional histogram that we have a much higher tendency to stay on x than what we have to move on to y.

Now we will take a look at how the density function for such a random walk will look like. We will use the values x, y and z as defined above on our graphical representation. First $C = \frac{\Gamma(1)\Gamma(1)\Gamma(\frac{3}{2})\Gamma(1)}{\Gamma(1)\Gamma(2)\Gamma(2)\Gamma(1)} = \Gamma(\frac{3}{2})$. It follows that the density function from (Eq. 3) for this particular graph looks like:

$$\phi(x) \quad = \quad \Gamma(3/2)\frac{1}{2!\pi^{\frac{3}{2}}}\frac{y}{(x+y)(y+z)^{\frac{3}{2}}\sqrt{xyz}}$$

From here I wanted to take a look at what the limiting density function would look like. Also, I wished to compare the limiting density function with the 3-dimensional histogram that I obtained after running my simulations:





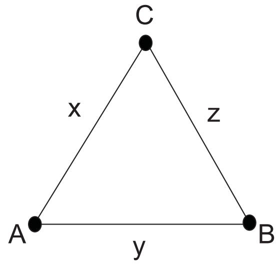
As you can see, there are some similarities between these two graphs. The overall shape of the two are the same. However, it is apparent that they are different by since there was a lot of clustering about the first edge (bottom right corner of the histogram). Included with the similarities is the noticeably raised ridges along the diagonal as well as the left side of each of these graphs. This represents high number of traversals for both the first and second edges (the diagonal) or a high number of traversals for the second and third edges (the left side) but not the other. What is really interesting is the noticeably lower part along the bottom of the graph. This shows that there are not a high number of traversals for the two end edges with a low number of traversals for the middle.

Let a, b, and c be the edge weights of {A,B}, {B,C}, and {C,D} respectively. We can form a probability transition matrix as follows:

$$\begin{array}{cc} & \text{to} \\ & \begin{array}{cccc} \text{A} & \text{B} & \text{C} & \text{D} \end{array} \\ \text{from} \begin{array}{c} \text{A} \\ \text{B} \\ \text{C} \\ \text{D} \end{array} & \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{a}{a+b} & 0 & \frac{b}{a+b} & 0 \\ 0 & \frac{b}{b+c} & 0 & \frac{c}{b+c} \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array}.$$

3.3 A Triangle:

The most popular and most looked at random walk with reinforcement is that on a triangle. First we will take a look at what happens when we run the same type of simulations as before on a triangle. Here we will let $\{A,B,C\} \in V$ and $\{x,y,z\}$ be the set of proportional weights after n steps. We will let all of the initial weights be one. Below is a graphical representation of where the weights are with respect to the vertices:



For this simulation we will start at Vertex B, take 400 steps, and run the simulation 2000 times. After the data was collected, it was put into tables and the plotted so that we could take a look at the distribution of the random walks. Below are three such plots:

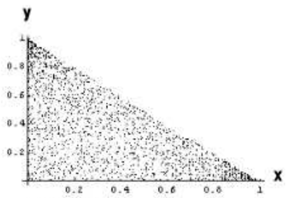


Figure 3.4

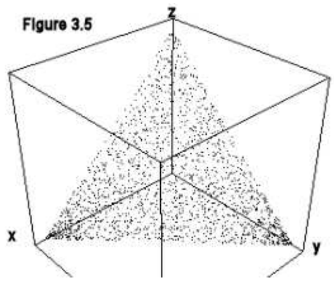


Figure 3.5

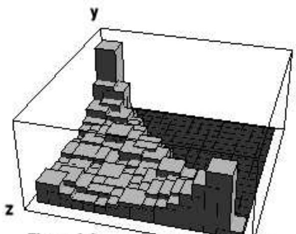


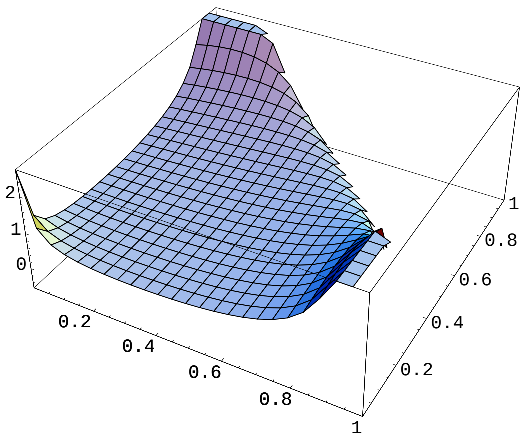
Figure 3.6

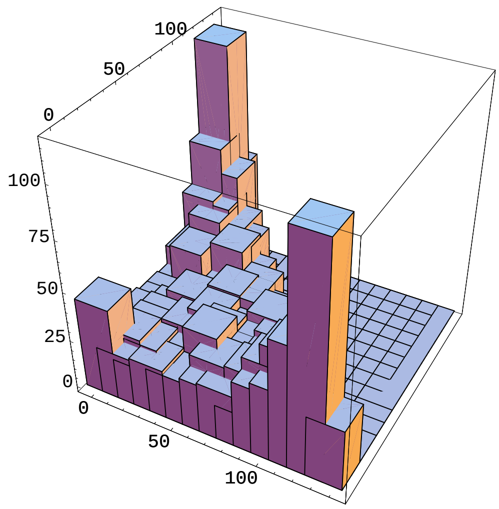
Taking a look at Figure 3.4 we can see that there is some clustering about x and y. This is because we are starting at Vertex A and x or y most be chosen first and thus has a higher tendency to oscillate on that edge than on z. This is also confirmed in Figure 3.5 as you can see the same type of clustering. Now taking a look at Figure 3.6 you can see that the distribution for these are really much higher.

Now we will take a look at how the density function for such a random walk will look like. We will use the values x, y and z as defined above on our graphical representation. First, $C = \frac{\Gamma(1)\Gamma(\frac{3}{2})\Gamma(\frac{3}{2})}{\Gamma(2)\Gamma(2)\Gamma(2)} = (\Gamma(\frac{3}{2}))^2$.It follows that the density function from (Eq. 3) for this particular graph looks like:

$$\phi(x) = \frac{(\Gamma(\frac{3}{2}))^2}{\pi} \frac{\sqrt{xy+xz+yz}}{(x+y)(y+z)^{\frac{3}{2}}(x+z)^{\frac{3}{2}}}$$

Taking a closer look at this density function, it appears to be quite symmetrical. To help illustrate this, I will provide a the limiting density plot along with the 3-dimensional histogram plot for analysis:





Taking a look at graph on the left, you can see that it is truly symmetrical. The most noticeable feature of this graph is the way it curves along the diagonal. This does is also noticed in the histogram to the right, even though it is not as apparent. Another, noticeable attribute that can be seen in both of these plots is the way they both seem to build up while approaching the diagonal. The major difference between the two lies with the height in the left corner as compared with the other two.

Let a, b, and c be the weights of $\{$A,C$\}$, $\{$A,C$\}$, and $\{$B,C$\}$ respectively. The Markov matrix at time n will look like:

$$\begin{array}{cc} & \text{to} \\ & \begin{array}{ccc} \text{A} & \text{B} & \text{C} \end{array} \\ \text{from} \begin{array}{c} \text{A} \\ \text{B} \\ \text{C} \end{array} & \begin{pmatrix} 0 & \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{b}{b+c} & 0 & \frac{c}{b+c} \\ \frac{a}{a+c} & \frac{c}{a+c} & 0 \end{pmatrix}. \end{array}$$

4. Partial Exchangeability in Random Walks with Reinforcement

In this section I will actually show that random walks can be written as a mixture of Markov chains. This result will lead to prove that random walks are partially exchangeable. The majority of this section will come from M.S. Keane and S.W.W. Rolles' paper "Edge-reinforced random walk on finite graphs [8]." I, however, have expanded the proofs to help make them more understandable.

Proposition 4.1. [8] *The reinforced random walk traversed almost surely every edge infinitely often in both directions*

Proof. Let $G=(V,E)$ be a graph and we'll choose an arbitrary edge $e \in E$ with $\bar{e} = \{u,v\}$ and $e^- = e(u,v)$. Let τ_i be the time of the i^{th} visit vertex v to where $\tau_i = \infty$ if v is visited at most $i-1$ times. So,

$$P(v \text{ is visited infinitely often, } e^- \text{ is traversed at most finitely often})$$

$$= \lim_{i_0\to\infty}\lim_{I\to\infty} P(\bigcap_{i_0\le i\le I}\{\tau_i<\infty, X_{\tau_i+1}\ne u\})$$

Now between two successive visits to v, the sum of the weights of the edges incident to v increase by two. Therefore $w_{\tau_i} \le a_v + 2i$. Therefore, given the past time up to time τ_i, the probability of not traversing e^- at time $\tau_i - 1$ is equal to $1 - \frac{w_{\tau_i}(e)}{w_{\tau_i}(v)} \le 1 - \frac{a_e}{a_v+2i} \le \exp(-\frac{a_e}{a_v+2i})$ By Induction we get,

$$P(\bigcap_{i_0\le i\le I}\{\tau_i<\infty, X_{\tau_i+1}\ne u\}) \le \prod_{i_0\le i\le I}\exp(-\frac{a_e}{a_v+2i})$$
$$= \exp(-\sum_{i_0\le i\le I}\frac{a_e}{a_v+2i}).$$

Now, when we take the limits as $I\to\infty$ and $i_0\to\infty$ in the last inequality we see that P(v is visited infinitely often, e^- is traversed at most finitely often) = 0. Since the graph is finite, there exists at least one vertex that is visited infinitely often. Thus P(an edge is traversed at most finitely often) = 0. Starting at v, we know that it has been visited an infinite number of times. Each edge incident to v has been traversed an infinite number of times in each direction. Thus all of the vertices that share an with v have been traversed an infinite number of times. Since our graph is connected, we can extend this to all of the edges in our graph. ■

*Mixture of Markov chains:*Let I be a finite set. Consider the set P of stochastic matrices on $I\times I$ with the topology of coordinate convergence. The set $I\times P$ is compact. We denote the coordinates of $p\in P$ by $p(i,j)$. A stochastic process $\{Z_n\}$ with state space I is called a *mixture of Markov chains* if for each $i\in I$ there exists a probability measure $\mu(i,\cdot)$ on the Borel sets of P such that

$$P(Z_j = i, for\ j=0,1,...,n) = \int_P \prod_{j=0}^{n-1} p(i_j,i_{j+1})\mu(i_0,\mathrm{dp}). \quad [8]$$

Theorem 4.2. [8]*Let $\{Z_n;\ n\in N\}$be a stochastic process with finite state space, and suppose that $P(Z_n = Z_0$ for infinitely many n) = 1. If Z is partially exchangeable, the Z is a mixture of Markov chains.*

Results: Let $p = (u_0, e_1, u_1, e_2, ..., e_n, u_n)$ be a path, and let $v\in V$. Denote the number of departures from v in the path p as $n_v(p)$ and the number of arrivals to v as $\bar{n}_v(p)$. We set

$$n_v = |\{i \in \{0,1,...,n-1\} : n_i = v\}|, \quad \bar{n}_v = |\{i\in\{0,1,...,n\} : n_i = v\}|.$$

It's clear that the number of departures and arrivals is determined by k^+ and k^-, and

$$n_v + \bar{v}_v = \sum_{\{e:v\in\bar{e}\}} k_e \quad (4)$$

If we denote v_0 as the starting point and v_1 as the endpoint, the we have

$$n_v + \delta_{v_1}(v) = \bar{n}_v + \delta_{v_0}(v), \quad (5)$$

where $\delta_u(v)$ is either 0 or 1 depending on whether $u\ne v$ or $u=v$ respectively. From (Eq. 4) and (Eq. 5) we see that

$$n_v = \frac{1}{2}(\delta_{v_0}(v) - \delta_{v_1}(v) + \sum_{\{e:v\in\bar{e}\}} k_e) \quad (6)$$

and

$$\bar{n}_v = \frac{1}{2}(\delta_{v_1}(v) - \delta_{v_0}(v) + \sum_{\{e:v\in\bar{e}\}} k_e)$$

Taking a closer look at the last equation, we can see that (k_e mod 2; $e \in E$) together with the starting point v_0 uniquely determines the endpoint of a *k-path*: If $\sum_{e:v\in\bar{e}} k_e$ is ever for all vertices, then the endpoint is also v_0. If there exists a vertex $v_1 \neq v_0$ such that $\sum_{e:v\in\bar{e}} k_e$ is even for $v \in V/\{v_0, v_1\}$ and odd for $v \in \{v_0, v_1\}$, then the endpoint is v_1. There is no *k-path* possible in all other cases.

Similar findings are in M.S. Keane's paper *Solution to problem 288* with his work on the triangle. Returning to our triangle, after n steps for a reinforced random walk let

$$\begin{aligned} a &= w_n(AC) \\ b &= w_n(AB) \\ c &= w_n(BC). \end{aligned}$$

Assuming that the random walk starts at vertex A, we can determine which vertex the process has stopped on at time n as follows:

If	Then
$c \equiv a \equiv b \ mod 2$	$v_n = A$
$a \not\equiv c \equiv b \ mod 2$	$v_n = C$
$b \not\equiv c \equiv a \ mod 2$	$v_n = B$
$c \equiv a \not\equiv b \ mod 2$	$v_n =$ no path possible [7]

Lemma 4.3 [8] *Two k-paths of reinforced random walks with the same starting point have the same probability. In particular, reinforced random walks are partially exchangeable.*

Proof: Here we will compute the probability of a fixed *k-path* of length n with starting point v_0. The probability will be the product of n factors as we will be taking n steps, or making n transitions. This comes from the independence principal which states that the probability of two independent events, A and B, is equal to the product of the probabilities. So, $P(AB) = P(A)P(B)$. This is used due to the fact tat we are using the property of Markov chains. So we can break up the probability of our random walk as follows:

$$P(X_0, X_1, ..., X_n) = P(X_0) * P(X_1|X_0) * \ldots * P(X_n|X_{n-1})$$

Since we are now just multiplying all of these probabilities together, we can use the commutative property of multiplication to arrange the numerator and denominator in crafty ways.

We will first start with the numerator. Each time we choose and edge, the weight at the time that we choose it is in the denominator. The very first time that we traverse an edge, we will use its initial weight and so on. Multiplying these together for each traversal of the edge we can create the following product:

$$\prod_{i=0}^{k_e-1}(a_e + i).$$

Since we can do this for all of the edges, we can just multiply them together so that our numerator will appear as follows:

$$\prod_{e\in E}\prod_{i=0}^{k_e-1}(a_e+i).$$

Now we will consider the denominator. Each time that we leave a vertex, we use the sum of all of the weights of the edges incident to the vertex as our denominator. We will collect these sums in a similar way to the way that we collected the weights for the numerator. We will need to separate the initial vertex, v_0, from the rest of the vertices. With the first step we will use just the sum of the initial weights. The next time that we leave v_0 we will have departed from and arrived to the vertex, thus increasing the sum by two. Following this reasoning we can group these sums for v_0 as follows:

$$\prod_{i=0}^{n_{v_0}-1}(a_{v_0}+2i).$$

Now looking at the other vertices, we follow the same logic for including the $2i$ term. However, for all vertices other than the initial vertex, they will have arrived once before they could leave, thus requiring us to add one. Grouping them together for each of these vertices, and then multiplying those products together we can obtain a final denominator of:

$$\prod_{i=0}^{n_{v_0}-1}(a_{v_0}+2i)\prod_{v\in V/\{v_0\}}\prod_{i=0}^{n_v-1}(a_v+1+2i).$$

Putting all of this together, we have the probability of a fixed *k-path* of length n with starting point v_0 is:

$$\frac{\prod_{e\in E}\prod_{i=0}^{k_e-1}(a_e+i)}{\prod_{i=0}^{n_{v_0}-1}(a_{v_0}+2i)\prod_{v\in V/\{v_0\}}\prod_{i=0}^{n_v-1}(a_v+1+2i)}$$

Since the number of crossings k and the starting point v_0 determine the unique endpoint, we can see that the probability only depends on k and v_0. Thus any path that starts from the same vertex, v_0, takes n steps, and ends with all of the same edge weights for respective edges will have the same probability. Thus satisfying the property of partial exchangeability. ■

5. Conclusions

We have taken a look at the general density function that is applicable to all random walks with reinforcement. We have gone over some random walks on specific graphs and dipped into the Markov matrix representations of each graph. Finally we walked through the fact that random walks with reinforcement have the property of partial exchangeability.

Appendix A.

Applying the Density Function to the Special Graphs

Line with where $m = 3$, and $l = 4$:

As I stated in the paper, $C = \Gamma(3/2)$. Now we will work through the components of (Eq. 3) and show the result. As a refresher, here the pictures of what is going on:

Here we will let x, y, and z be the proportional weights at time n, such that they satisfy $x+y+z=1$. The initial weights are all 1, so $a_e = 1$ for all edges. Also, $x_1 = x$, $x_2 = y$, $x_3 = z$ in this case. Now we will look at the various components of $\phi(x)$:

$$2^{1-l+\sum_{e\in E} a_e} = 2^{1-4+(1+1+1)} = 1$$

$$\prod_{e\in E} x_e^{a_e - \frac{1}{2}} = x^{1/2}y^{1/2}z^{1/2}$$

$$x_{v_0}^{\frac{a_{v_0}}{2}} = (x+y)$$

and

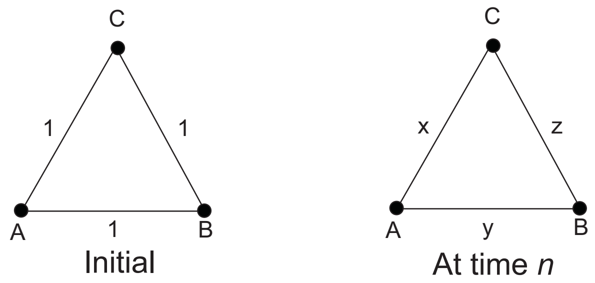
$$\prod_{v\in V/\{v_0\}} x_v^{\frac{a_v+1}{2}} = x(y+z)^{3/2}z$$

Plugging all of this into (Eq. 3) we get

$$\begin{aligned}
\phi(x) &= C\frac{2^{1-l+\sum_{e\in E} a_e}}{(m-1)!\pi^{\frac{m}{2}}}\frac{\prod_{e\in E} x_e^{a_e-\frac{1}{2}}}{x_{v_0}^{\frac{a_{v_0}}{2}}\prod_{v\in V/\{v_0\}} x_v^{\frac{a_v+1}{2}}}\sqrt{\det(A(x))} \\
&= \Gamma(3/2)\frac{1}{2!\pi}\frac{x^{1/2}y^{1/2}z^{1/2}}{(x+y)x(y+z)^{3/2}z} \\
&= \Gamma(3/2)\frac{1}{2!\pi^{\frac{3}{2}}}\frac{y}{(x+y)(y+z)^{\frac{3}{2}}\sqrt{xyz}}
\end{aligned}$$

The Triangle ($m = 3$ and $l = 3$):

As I stated in the paper, $C = (\Gamma(\frac{3}{2}))^2$. Now we will work through the components of (Eq. 3) and show the result. As a refresher, here the pictures of what is going on:



Here we will let x, y, and z be the proportional weights at time n, such that they satisfy $x+y+z=1$. The initial weights are all 1, so $a_e = 1$ for all edges. Also, $x_1 = x$, $x_2 = y$, $x_3 = z$ in this case. Now we will look at the various components of $\phi(x)$:

$$2^{1-l+\sum_{e\in E} a_e} = 2^{1-3+(1+1+1)} = 2$$

$$\prod_{e\in E} x_e^{a_e - \frac{1}{2}} = x^{1/2}y^{1/2}z^{1/2}$$

$$x_{v_0}^{\frac{a_{v_0}}{2}} = (x+y)$$

$$\prod_{v\in V/\{v_0\}} x_v^{\frac{a_v+1}{2}} = (y+z)^{3/2}(x+z)^{3/2})$$

and

$$\sqrt{\det(A(x))} = \sqrt{1/x+1/y+1/z}$$

Plugging all of this into (Eq. 3) we get

$$\begin{aligned}
\phi(x) &= C\frac{2^{1-l+\sum_{e\in E} a_e}}{(m-1)!\pi^{\frac{m}{2}}} \frac{\prod_{e\in E} x_e^{a_e-\frac{1}{2}}}{x_{v_0}^{\frac{a_{v_0}}{2}} \prod_{v\in V/\{v_0\}} x_v^{\frac{a_v+1}{2}}} \sqrt{\det(A(x))} \\
&= \frac{(\Gamma(\frac{3}{2}))^2}{\pi} \frac{x^{1/2}y^{1/2}z^{1/2}}{(x+y)(y+z)^{3/2}(x+z)^{3/2}} \sqrt{1/x+1/y+1/z} \\
&= \frac{(\Gamma(\frac{3}{2}))^2}{\pi} \frac{\sqrt{xy+xz+yz}}{(x+y)(y+z)^{\frac{3}{2}}(x+z)^{\frac{3}{2}}}
\end{aligned}$$

Appendix B.

Mathematica Code for Random Walks

Line with where $m = 2$, and $l = 3$:

The following code sets up the edge set e, and also sets up the data that I will be collecting:

```
<< Graphics‘Graphics‘
e = Table[2, {i, 1, 2}]
xdata := List[]
ydata := List[]
zdata := List[]
p := 1
```

The following code runs the random walk through a loop so that I can obtain many walks. It also collects the edge weights at the end of each walk as a list of data:

```
time := 100
Do[{e = Table[2, {i, 1, 2}], p = 1,
For[i = 0, i < time, i++,
fwd = e[[p + 1]];
back = e[[p]];
choice = If[Random[] < back/(fwd + back), -1, +1];
e[[p + (choice + 1)/2]] = e[[p + (choice + 1)/2]] + 2;
p = 1],
xdata = Join[xdata, {e[[1]]}],
ydata = Join[ydata, {e[[2]]}],
m = n},
{n, 1, 1000}]
```

The following code gives me a histogram for the data collected on one edge:

```
Histogram[xdata, Frequency → True, HistogramCategories → {4, 14, 24, 34, 44, 54, 64, 74, 84, 94, 104, 114,
124, 134, 144, 154, 164, 174, 184, 194, 204}]
```

Line with where $m = 3$, and $l = 4$:

The following code sets up the edge set e, and also sets up the data that I will be collecting as well as loads the packages that I will be needing later:

```
<< Realtime3D‘
<< Graphics‘Graphics3D‘
e = Table[1, {i, 1, 3}]
xdata := List[]
ydata := List[]
```

```
zdata := List[]
p := 1
```

The following code runs the random walk through a loop so that I may obtain data from many walks of the same length. there are many if statements in this one due to the fact that at the endpoints I must force the next step. I also collect the weights of the edges into their respective lists of data.

```
time := 400
Do[{e = Table[1, {i, 1, 3}], p = 1,
For[i = 0, i < time, i++,
fwd = e[[p + 1]];
back = e[[p]];
choice = If[Random[] ¡ back/(fwd + back), -1, +1];
e[[p + (choice + 1)/2]]++;
p = p + choice;
If[p < 1, i++, i = i];
If[p > 2, i++, i = i];
If[(p < 1) && (i < time), e[[1]]++, e[[1]] = e[[1]]];
If[(p > 2) && (i < time), e[[3]]++, e[[3]] = e[[3]]];
If[p < 1, p = p + 1, p = p];
If[p > 2, p = p - 1, p = p];],
xdata = Join[xdata, {e[[1]]}],
ydata = Join[ydata, {e[[2]]}],
zdata = Join[zdata, {e[[3]]}],
m = n},
{n, 1, 4000}]
```

The following code arranges the data in desired ways and then produces the plots that are used in the paper.

```
xydata := Table[{xdata[[i]], ydata[[i]]}, {i, 1, m}]
hdata := Table[{xdata[[i]]/2, ydata[[i]]/2}, {i, 1, m}]
alldata := Table[{xdata[[i]], ydata[[i]], zdata[[i]]}, {i, 1, m}]
ListPlot[xzdata]
ScatterPlot3D[alldata, AxesLabel→ {"xdata", "ydata" , "zdata"}]
Histogram3D[hdata, FrequencyData→False]
```

Triangle ($m = 3$ **and** $l = 3$)**:**

The following code sets up the edge set e, and also sets up the data that I will be collecting as well as loads the packages that I will be needing later:

```
<< Realtime3D`
<< Graphics`Graphics3D`
```

```
<< Graphics`Graphics`
e = Table[1, {i, 1, 3}]
xdata := List[]
ydata := List[]
zdata := List[]
p := 1
```

The following code runs the random walks through a loop.

```
time := 300
Do[{e = Table[1, {i, 1, 3}], p = 1,
For[i = 0, i<time, i++,
fwd = e[[Mod[p, 3] + 1]];
back = e[[p]];
choice = If[Random[]<back/(fwd + back), -1, +1];
e[[Mod[p + (choice + 1)/2 - 1, 3] + 1]]++;
p = Mod[p + choice - 1, 3] + 1],
xdata = Join[xdata, {e[[1]]}],
ydata = Join[ydata, {e[[2]]}],
zdata = Join[zdata, {e[[3]]}],
m = n},
{n, 1, 2000}]
```

The following code arranges the data in desired ways and then produces the plots that are used in the paper.

```
xydata := Table[(xdata[[i]] - 1)/(time), (ydata[[i]] - 1)/(time)}, {i, 1, m}]
ListPlot[xydata, AxesLabel → {"xdata", "ydata"}]
alldata:= Table[{xdata[[i]]/(time + 3), ydata[[i]]/(time + 3), zdata[[i]]/(time + 3)}, {i, 1, m}]
ScatterPlot3D[alldata, AxesLabel → {"xdata", "ydata" , "zdata"}]
hxydata := Table[{xdata[[i]]/2, ydata[[i]]/2}, {i, 1, m}]
Histogram3D[hxydata, FrequencyData → False]
```

References

[1] Bernardo, Jose M. ”The Concept of Exchangeability and its applications.” *Far East J. Math. Sci*. 1996, Special Volume, Part II, 111-121.

[2] Diaconis, P. ”Recent progress on de Finetti’s notions of exchangeability.” In *Bayesian Statistics*, volume 3, pages 111-125. Oxford University Press, 1988.

[3] Diaconis, P. and Rolles, S.W.W. ”Baysian analysis for reversible Markov chains.” *The Annals of Statistics*, Vol. 34, No. 3 June 2005,

[4] Evans, Michael and Rosenthal, Jeffry. Probability and Statistics: The Science of Uncertainty. New York: W.H. Freeman and Company. 2003.

[5] Freedman, David A. ”Mixtures of Markov Processes.” *The Annals of Mathematical Statistics*, Vol. 33, No. 1 (Mar., 1962), 114-118.

[6] Ghahramani, Saeed. Fundamentals of Probability with Stochastic Processes New Jersey: Peasron Education, Inc. 2005.

[7] Keane, M.S. Solution to problem 288. *Statistica Neerlandica,* 44(2):95-100, 1990

[8] Keane, M.S. and Rolles, S.W.W. ”Edge-reinforced random walk on finite graphs.” *Infinite dimensional stochastic analysis*, pages 217-234. Verh. Afd. Natuurkd. 1. Reeks. K. Ned. Akad. Wet., 52, *R. Neth. Acad. Arts Sci., Amsterdam, 2000.*

[9] Rolles S.W.W. ”How edge-reinforced random walks arise naturally. *Probab. Theory related Fields*, 126(v):143-260, 2003