ASYMMETRIC DIFFUSION PROCESSES

LISA RUAN, ANNELISE STEELE, AND MICHAEL TRAUB

1. Introduction

The focus of this paper will be systems of particles on a line that move with asymmetric random motion. In asymmetric motion, one direction is favored over another. Suppose we have a system of N positions and k particles. Let each particle have the ability to move one position either to the left or right. We also impose the following two conditions: at most one particle can move in each time step, and a particle can only move to a position that is not occupied. These rules lead us to characterize the system's dynamics as transitions between states, rather than transitions for each particle independently. This introduced asymmetry has unknown effects on the equilibrium distribution of particles and the speed of the particles in equilibrium. In this paper, we study these phenomena in two such asymmetric systems.

For all systems, we introduce the asymmetry parameter $q \in [0, 1]$. If the parameter q equals one, it is totally symmetric and every particle has an equal probability of moving right or left. However, if the parameter q equals 0, the particles in the system will always move to the right.

The first system is referred to as the periodic boundary. It is closed and periodic, which means that the number of particles in our system cannot change from some initial value k. This system S and the system's transitions can be visualized in

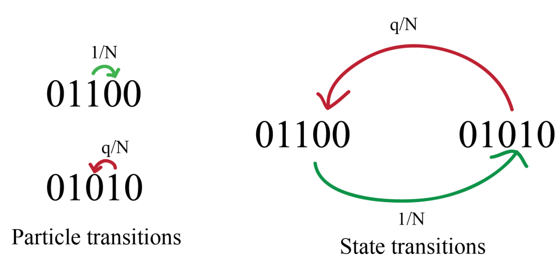

Figure 1. In this illustration of system S, State 1 corresponds to "01100" and State 2 corresponds to "01010". State 1 and State 2 highlight how particle transitions correspond to transitions between states when $N = 5$ and $k = 2$. A 1 represents a position occupied by a particle and a 0 represents an empty position. If a particle in State 1 transitions one position to the right, as demonstrated in State 2, it moves with probability $1/N$. If a particle in State 2 transitions one position to the left, as demonstrated in State 1, it moves with probability q/N.

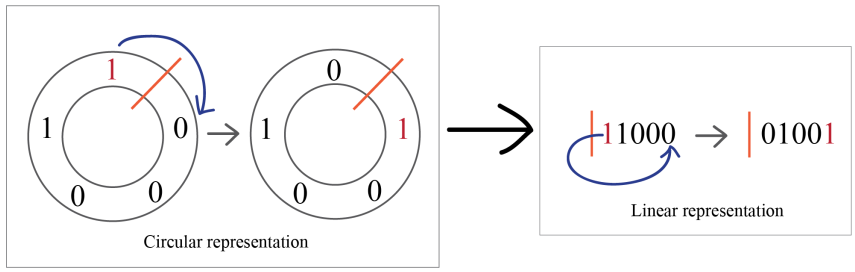

FIGURE 2. This circular movement demonstrates how particles can transition from one end of the system to the opposite end.

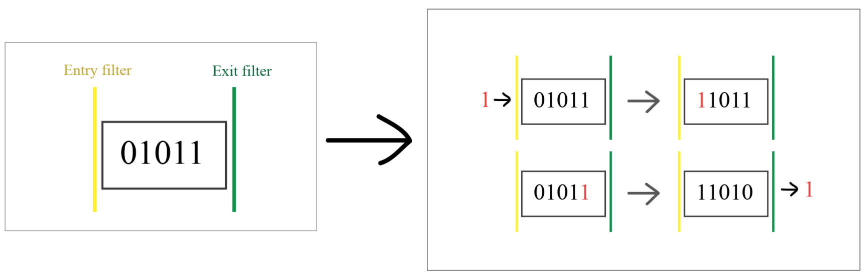

FIGURE 3. The open boundary can be visualized using a pipe. The two ends of the pipe represent entry and exit filters, because particles have the ability to enter from the left and exit from the right.

Figure 1. This implies that particles can transition from one end of the linear representation to the opposite end of our system. The easiest way to visualize this is to consider our system as a circle as can be seen in Figure 2.

The second system that we consider is referred to as the open boundary. We can consider a system of N positions with endpoints where particles have the ability to enter from the left and exit from the right. Because there are a variable number of particles for a system at any time, there exist 2^N possible states for a given N. We will assume in our system that a particle can never enter the system if the leftmost position is already occupied, and a particle can never exit the system from the left. A particle can only exit the system if the rightmost position is occupied, and no particle can ever enter the system from the right. We introduce the boundary parameters $a, b \in [0, 1]$. Particles can enter the system from the left with probability $a/(N+1)$. On the opposite side of the system, particles can exit the system from the right with probability $b/(N+1)$. All other particles transition to the right with probability $1/(N+1)$ and transition to the left with probability $q/(N+1)$. We can visualize this process in Figure 3.

For the Periodic System, we characterized the equilibrium distribution of particles and flow by obtaining closed-form equations for the steady state equilibrium and average speed. For the open boundary, we constructed simulations to understand how the asymmetry of the motion of each particle interacted with the asymmetry of the boundary conditions.

1.1. **Markov Process Steady State.** We treat the periodic boundary and open boundary as Markov processes. We can make this statement, because the probability that we transition to a new state depends only on the current state. Our goal is to characterize the steady state distribution for the aforementioned asymmetric diffusion processes. To do this, we build a transition matrix P of size M x M, where M is the number of possible states. Let P_{ij} represent the probability that the state s_i transitions to state s_j. Therefore, $P_{ij} = \mathbb{P}\ [\ x_{(t+1)} = s_i \mid x_t = s_j\ ]$. Given some initial distribution over states $y^{(0)}$ where $y_i^{(0)} = \mathbb{P}[x^{(0)} = s_i]$, the distribution over states at time t is $y^{(t)} = P^t y^{(0)}$. Since the distribution $\mathbb{P}[x^{(t)}]$ is normalized, we also know that $y^{(t)}$ sums to 1 for all t.

Since our transition matrix is a probability matrix where columns sum to 1, P is a Markov matrix. If follows that P always has an eigenvalue equal to one. This tells us that the eigenvalue of one has a corresponding eigenvector y_{eq} that represents the steady state of the system. This equilibrium occurs when $y_{eq} = Py_{eq}$. Our goal is to characterize this y_{eq} for each of the aforementioned boundary conditions in terms of the system's parameters.

2. Periodic Boundary

In this section we will show that for the periodic system, in equilibrium the distribution expected number of particles at each position is uniform, but the flux of particles in the system follows a non-trivial distribution, that as $N \to \infty$, is tightly concentrated around $(1-q)\frac{k}{N}(1-\frac{k}{N})$. We will explore other properties of this distribution in the limit.

2.1. **Process Definition.** Here we introduce the asymmetric diffusion process with periodic boundary conditions more formally, a Let S_N be the set of length N bitstrings $\{0,1\}^N$. For each $s \in S_N$, let $L(s)$ be the set of strings where exactly one instance of the substring 01 has been replaced by 10. Also, if s starts with a 1 and ends with a 0, then include the state where these are switched, corresponding to the first particle moving and looping around to the end. For example, $L(10110) = \{11010, 00111\}$. Define $R(s)$ to be the set of strings such that for every $s' \in L(s)$, $s \in R(s')$, so that $R(s)$ are the states after some particle in s moves right. This implicitly uses the fact that a particle can always move back to where it came from on the next time step.

Definition 2.1 (Asymmetric Diffusion Process)**.** Let the asymmetric diffusion process with closed boundary defined by the 3-tuple $\{N, q, x^{(0)}\}$, where $q \in [0,1]$ and $x^{(0)} \in S_{N,k}$ where $S_{N,k}$ are the length N bitstrings with k ones. For a state of the system $x^{(t)} \in S_N$, the distribution over $x^{(t+1)}$ is given by:

$$\mathbb{P}\ [\ x^{(t+1)} = s' \mid x^{(t)} = s\ ] = \begin{cases} 1/N & \text{if } s' \in R(s) \\ q/N & \text{if } s' \in L(s) \\ 1 - \frac{|R(s)|+q|L(s)|}{N} & \text{if } s' = s \\ 0 & \text{otherwise} \end{cases}, \tag{2.1}$$

and $x^{(t+1)}$ is drawn from this distribution.

2.2. **Equilibrium Position.** Below we prove that for the periodic system, the equilibrium distribution over states is uniform, and it will immediately follow that the expected number of particles in each position is equal to k/N. While the system has asymmetric transitions to the left and right for any particle, the periodic boundary condition introduces a new symmetry that gives us the uniform equilibrium. Specifically, the periodic boundary means that $|L(s)| = |R(s)|$ for all states. As before, $L(s)$ are the states where one particle in s moves left and $R(s)$ are the states where one particle in s moves right.

First we will prove $|L(s)| = |R(s)|$, then we will show how this gives us a uniform steady state because the rows of the transition matrix P sum to 1.

Lemma 2.2. *Let $L(s)$ be the number of particles that can move to the left when the system is in state s with a periodic boundary condition, and similarly let $R(s)$ be the particles that can move to the right. For all states s, $L(s) = R(s)$.*

Proof. For any maximal sequence of 0s in the string, then there is a 1 at the beginning of this sequence that can move right and a 1 at the end that can move left. These 1s may be the same. Each of these correspond to some element in $R(s)$ and $L(s)$ respectively. Since the size of each of these sets is equal to the number of longest sequences of 0, their sizes are equal. $\square$

Lemma 2.3. *Given some $N \times M$ matrix A such that the rows of A sum to 1, i.e. $\Sigma_j A_{ij} = 1$, then $\mathbb{1}$, the length M vector of all 1s, is a right eigenvector of A with eigenvalue 1.*

Proof. Let A be defined as above. $(A\mathbb{1})_j = \sum_{i=1}^{M} A_{ij} = 1$, so $A\mathbb{1} = \mathbb{1}$, proving that $\mathbb{1}$ is an eigenvector of A with eigenvalue 1. $\square$

Theorem 2.4. *Let P be the $\binom{N}{k} \times \binom{N}{k}$ transition matrix of the asymmetric diffusion process with a periodic boundary condition with parameters N, k, and q. The all-ones vector $\mathbb{1}$ is an eigenvector of P corresponding to eigenvalue 1.*

Proof. Using Lemma 2.3, it is sufficient to show that the rows of P sum to 1. Let $M = \binom{N}{k}$. Using the fact that $s \in R(s')$ for every $s' \in L(s)$, we know that $|R(s')| = |L(s)|$ for every s, s' such that $s' \in L(s)$. We use this fact to get Equation 2.2c. With Lemma 2.2 and the fact that probabilities sum to 1, we complete the proof as shown below:

$$\sum_{j=1}^{M} P_{ij} = \sum_{j=1}^{M} \mathbb{P}\,[\ s_{t+1} = s_i \mid s_t = s_j\ ] \tag{2.2a}$$

$$= \mathbb{P}\,[\ s_{t+1} = s_i \mid s_t = s_i\ ] + |R(s_j)|\frac{1}{N} + |L(s_j)|\frac{q}{N} \tag{2.2b}$$

$$= \mathbb{P}\,[\ s_{t+1} = s_i \mid s_t = s_i\ ] + |L(s_i)|\frac{1}{N} + |R(s_i)|\frac{q}{N} \tag{2.2c}$$

$$= \mathbb{P}\,[\ s_{t+1} = s_i \mid s_t = s_i\ ] + |R(s_i)|\frac{1}{N} + |L(s_i)|\frac{q}{N} \tag{2.2d}$$

$$= \sum_{j=1}^{M} \mathbb{P}\,[\ s_{t+1} = s_j \mid s_t = s_i\ ] \tag{2.2e}$$

$$= 1. \tag{2.2f}$$

□

Corollary 2.5. *When the asymmetric diffusion process with a periodic boundary is in equilibrium, the expected number of particles in each position is k/N.*

Proof. When the system is in equilibrium, according to Theorem 2.4, the distribution over states is uniform. The expected number of particles at position j in the length N bit-string at equilibrium is just the fraction of states where position j has a particle, which is $\binom{N-1}{k-1}/\binom{N}{k} = k/N$. □

2.3. **Equilibrium Flow.** While the equilibirum distribution over states is uniform, this does not mean that the particles do not have any net motion in equilibrium. In this section we define what speed of a state means in terms of its expectation of any particle moving left or right. This is linearly related to the number of particles that can move right. We show that even in equilibrium, the distribution over speeds is not uniform, and show that for fixed k/N, the expected speed converges to a fixed value as N is taken larger towards infinity.

We define the speed as the expected number of particles that move right minus the expected number of particles that move left. This gives us the following definition:

Definition 2.6 (State speed in periodic system)**.**

$$\mathbb{E}[v(s)] = \frac{1-q}{N}\ |R(s)|, \tag{2.3}$$

where $R(s)$ is the number of possible transitions to the right in state s.

Where $P[R]$ is the probability $R = c$ given N and k.

2.3.1. *Counting States with Fixed Open Spots to Right.*

Lemma 2.7. *Given y identical boxes and x identical objects, the number of ways to distribute x objects in y boxes such that all boxes are nonempty is $\binom{x-1}{y-1}$.*

Theorem 2.8. *Given an asymmetric diffusion process with a periodic boundary condition of size N with k particles, let the states such that $|R(s)| = R$ be S_R. It holds that*

$$|S_R| = \frac{N}{R}\binom{N-k-1}{R-1}\binom{k-1}{R-1}. \tag{2.4}$$

Proof. For each string $s \in S_R$, we can form some circular string $\sigma(s)$ by wrapping s around a circle clockwise. Since there are exactly R possible transitions to the right in each string, each one of these $\sigma(s)$ will have R boxes of 1s alternating with R boxes of 0s. First we find the size of set of all circular strings $\Sigma_R = \{\sigma(s) \mid s \in S_R\}$. To do this, we count the ways in which we can place the k ones into these R boxes such that all boxes are nonempty, and similarly count the ways we can place the $N-k$ zeros into the boxes. This division of 0s and 1s into boxes on the circle is illustrated in Figure 4. Using Lemma 2.7, we get $|\Sigma_R| = \binom{N-k-1}{R-1}\binom{k-1}{R-1}$.

Notice that each $\sigma \in \Sigma_R$ can be mapped to a set of strings $s \in S_R$ by unraveling $\sigma \in \Sigma_R$ at different indices. However, some $\sigma \in \Sigma_R$ produce the same strings as illustrated in Figure 5. To finish the proof, we need to count the number of unique $s \in S_R$ that are mapped to the same $\sigma \in \Sigma_R$, and multiply this times $|\Sigma_R|$. For every $\sigma \in \Sigma_R$, let us fix a symmetry index C such that σ can be divided into C

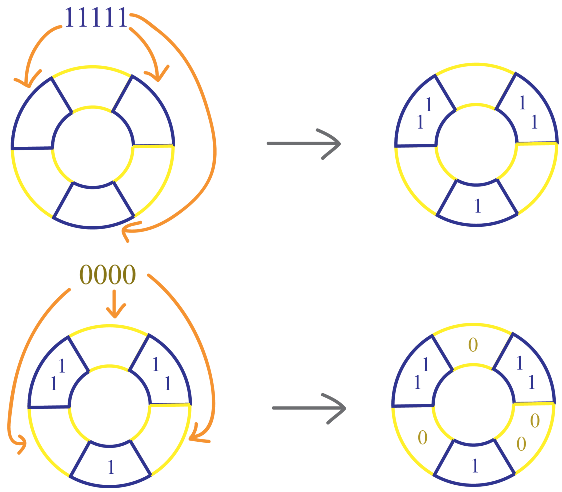

FIGURE 4. Illustration of putting 5 ones into the R boxes, followed by putting 4 zeros into R boxes.

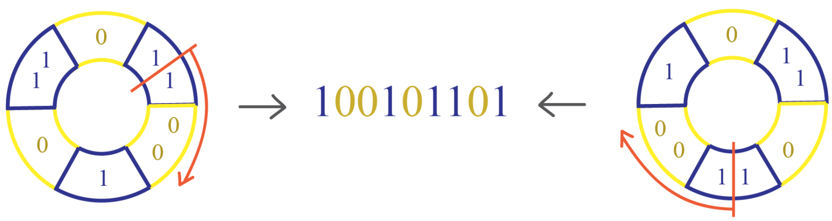

FIGURE 5. Nonunique strings created by rotationally symmetric circles

chunks that are rotationally symmetric. For example, 010101 has $C = 3$ and 00101 has $C = 1$.

For each rotationally symmetric chunk in σ we generate N/C unique strings of length N. There are R/C ways to rotate each symmetric chunk to produce the same set of strings. This is illustrated in Figure 6. Taken together, there are $\frac{N}{C}\frac{C}{R} = \frac{N}{R}$ elements in S_R that correspond to one element in Σ_R. This completes the proof. $\square$

With this result, we can use the uniformity of the equilibrium to exactly define the distribution over $|R(s)|$ in equilibrium.

2.4. **Distribution of Flow.**

Corollary 2.9. *Given an asymmetric diffusion process with a periodic boundary condition of length N with k particles, in equilibrium the distribution over $|R(s)|$ for all states s in is given by*

$$\mathbb{P}[R] = \frac{N}{R}\frac{\binom{k-1}{R-1}\binom{N-k-1}{R-1}}{\binom{N}{k}}. \tag{2.5}$$

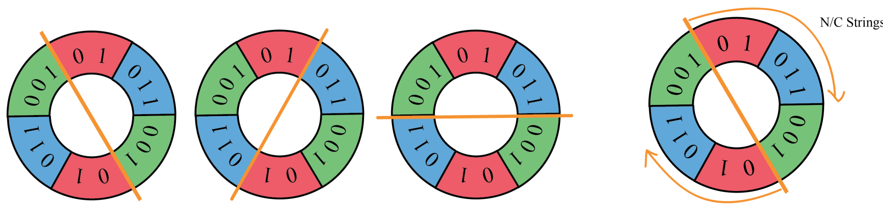

FIGURE 6. The circular string pictured above has $N = 16$, $k = 8$, and $R = 6$. It has 2-symmetry, illustrated by the orange dividing line. We can see there are R/C ways to align the divider, and for a given divider we can generate N/C strings by picking a starting symbol in a divided section and unwinding the string clockwise.

Proof. By Theorem 2.4, the equilibrium distribution over states is uniform. This means $\mathbb{P}[R] = \sum_s \mathbb{P}[x_{eq} = s]\mathbb{1}\{|R(s)| = R\} = \frac{1}{\binom{N}{k}} \sum_s \mathbb{1}\{|R(s)| = R\}$. Using Theorem 2.8, we prove the result. □

Theorem 2.10. *Let the distribution over R be given as in Equation 2.5. Let $\alpha = k/N$. The expected value of R is $\frac{N^2}{N-1}\alpha(1-\alpha)$.*

Proof.

$$\begin{aligned}
\mathbb{E}[R] &= \frac{N}{\binom{N}{k}} \sum_R \binom{k-1}{R-1}\binom{N-k-1}{R-1} \\
&= \frac{N}{\binom{N}{k}} \binom{N-2}{k-1} \\
&= \frac{k(N-k)}{N-1} \\
&= \frac{N^2}{N-1}\alpha(1-\alpha)
\end{aligned}$$

□

Corollary 2.11. *The expected flow in an asymmetric diffusion process with periodic boundary conditions and parameters N, k, and q is $\frac{N}{N-1}(1-q)\alpha(1-\alpha)$ where $\alpha = k/N$. In the limit $N \to \infty$ with α held fixed, this gives a limiting expected flow of $(1-q)\alpha(1-\alpha)$.*

3. Open Boundary

3.1. **Introduction.** For the system with an open boundary, we consider a chain of N positions with endpoints where particles have the ability to enter from the left and exit from the right. This means that the number of particles in a system is variable, so there exist 2^N possible states for a given N. In the open boundary case, there are three asymmetric parameters: q, a, and $b \in [0, 1]$.

In this section, we analyze the behavior of the steady state eigenvector and the expected number of particles at each position in the steady state. As described in the introduction, we treat the open boundary as a Markov process, because the probability that we transition to a new state depends only on the current state.

Using the same reasoning as in Section 1.1, we know that there is a steady state distribution over the states, which is given by the eigenvector corresponding to the eigenvalue of one in the transition matrix. Ultimately, we constructed simulations to understand how the asymmetry of the motion of each particle interacts with the asymmetry of the boundary conditions. The periodic boundary had a symmetric boundary condition, which was the reason for its uniform equilibrium. However, the open boundary exhibited non-uniform steady states as a result of the asymmetric boundaries—particles cannot leave from left or enter from right. This often lead to a non-linear steady state behavior, which we explore in this section.

3.2. **Steady State Eigenvector and Behavior.** We ran numerous simulations for different values of N, q, a, and b, in order to determine the behavior of a system in the steady state. Because there are 2^N possible states for a given N, our simulations slowed down significantly for values larger than $N = 10$. However, from these simulations, we were able to understand how the parameters interacted with each other to affect the steady state behavior. In this section, we will explain the process of how we created these simulations and analyze three cases: one where we change N and keep all other parameters constant, one where we change a and b and keep all other parameters constant, and another where we change q and keep all other parameters constant.

3.2.1. *Process.* We constructed line plots to represent the expected number of particles in each position for a given N in the steady state. We constructed eigenvectors for the different values of N using the following method:

Let y_{eq} be the $2^N \times 1$ steady state probability vector that is a probability distribution on the set $\{1, ..., 2^N\}$ such that $Py_{eq} = y_{eq}$ where P is the $2^N \times 2^N$ transition matrix. Rather than finding the full set of eigenvectors, which has asymptotic complexity of $O(8^N)$, we only found the eigenspace corresponding to the steady state. In the steps below, we find y_{eq} much faster by finding the nullspace of $P - I$, which runs $O(4^N)$ steps.

$$\begin{aligned} Py_{eq} &= y_{eq} \\ Py_{eq} &= Iy_{eq} \\ (P-I)y_{eq} &= 0 \\ y_{eq} &\in \text{Nullspace}(P-I) \end{aligned}$$

We performed the operations above both numerically and symbolically. The symbolic matrix operations were much slower than the numerical algorithms.

Let us look at a steady state example for a single choice of parameters. Define the matrix S, which is an N x 2^N matrix where each column of S represents a state expressed in binary. In order to find the expected number of particles in each position in the steady state, we multiplied Sy_{eq}, where y_{eq} was normalized. For example, when $N = 2$, $q = 0$, $a = 1$, and $b = 1$:

$$Sy_{eq} = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1/5 \\ 1/5 \\ 2/5 \\ 1/5 \end{bmatrix} = \begin{bmatrix} 3/5 \\ 2/5 \end{bmatrix}$$

The resulting matrix is an N x 1 vector, which has corresponding probabilities for the expected number of particles within each position for a given N. In this

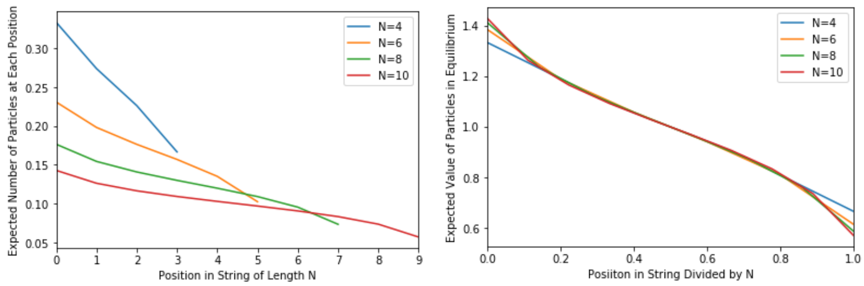

FIGURE 7. Suppose $q = 0, a = 1, b = 1$, and N takes on the values of 4, 6, 8, and 10. The first graph represents the expected number of particles at each position in a string of length N. In the second graph, we scale to adjust for the different values of N. Each N has a different eigenvector that is normalized over a set of probabilities for the number of expected particles within each state. On the y-axis, we are essentially undoing this normalization of the probabilities by multiplying each eigenvector by their respective N. On the x-axis, we break up the different N values by dividing each line's index by its respective N value. This allows for us to compare all of the lines equally.

example where $N = 2$, the expected number of particles at position 0 is 3/5 and the expected number of particles at position 1 is 2/5.

3.2.2. *Analyzing the Steady State Behavior.* The aforementioned process allowed us to create the simulations below for different values of N, a, b, and q.

In the second graph of Figure 7, the lines for N are the equations that we are trying to describe for the eigenvectors. When $a = b$, we see that the lines are mirrored across the center of the graph. As we increase N, the line begins to take on the shape of a decreasing spline function. It also appears that as N increases, the degree of curvature increases. This means the function is flatter around the inflection point and the tail-ends have a greater slope. We assume that this pattern would continue for larger values of N given $q = 0$. Therefore, we hypothesize that in the continuum limit, the density near the middle of the system is completely insensitive to the boundary conditions. As a result, the steady state behavior is non-linear in this system. Furthermore, this plot shows how increasing N exaggerates the asymmetric motion of the particles.

In Figure 8, we can see that as we increase the values of a and b, the line goes from having a positive slope to a negative slope. It is clear that for smaller values of a and b, particles tend to accumulate on the right. This confirms our intuition that as the boundary permits fewer particles to pass, the preference towards moving to the right causes the particles to pile up. Furthermore, for larger values of a and b, the expected number of particles on the left half tend to be higher than the right half. This is not as intuitive. However, it reflects how the asymmetry of the boundary conditions amplifies the asymmetric motion to pump particles out from the right side of the string faster than the particles can diffuse across the string. For the value of $a = b = .5$, the line is flat, but this does not hold for all q. This shows that even with the open boundary, we can manipulate a and b such that we

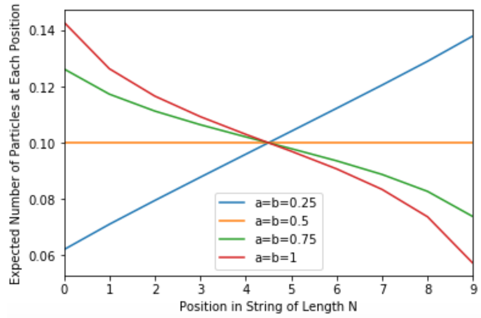

FIGURE 8. Suppose $N = 10, q = 0$, and a and b take on the values of .25, .5, .75, and 1. The inflow and the outflow of the particles remain equal, but it is varied over the range $[0, 1]$ for the fixed value of $q = 0$. All other parameters remain equal throughout this scenario.

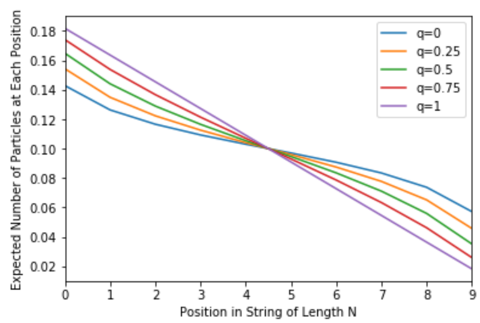

FIGURE 9. Suppose $N = 10, a = 1, b = 1$, and q takes on the values of 0, .25, .5, .75, and 1. The inflow and outflow of the particles is equal, N remains constant, and the parameter q is varied.

can create a uniform distribution just as in the periodic case. However, most values of a and b render non-linear behavior and non-uniform steady states.

In Figure 9, we can see the expected number of particles at each position in a string of length $N = 10$ with varying q in the steady state. Unlike the periodic boundary, there is not necessarily a uniform distribution for the expected number of particles at every position in the steady state. This is due to the fact that particles can only enter from the left side of the system and particles can only exit from the right side of the system. This asymmetry causes a non-uniform distribution even when the inflow and outflow of the particles is equal and $q = 0$. As we increase q, there seems to be greater insensitivity to the boundary in the steady state. We suspect that because particles can only move left, the particles cannot even themselves out as efficiently once they are clumped up. It makes sense that as we increase q to 1, we recover the standard linear diffusion pattern that emerges in certain symmetric diffusion processes, such as heat flow.

These simulations allowed us to understand how the asymmetry of the boundary parameters interacted with other parameters in the system, and how that ultimately affected the steady state behavior and shape.

4. Discussion

In this paper, we have completely characterized the equilibrium distribution of particles and the equilibrium distribution of flow in the periodic system. Specifically, we showed that the system always reaches a uniform equilibrium. Furthermore, we demonstrated that the distributions of flow converge to $(1-q)\alpha(1-\alpha)$ where $\alpha = k/N$ is held fixed but N is taken to infinity. This proves that we expect the periodic system to have well-defined, precise, macroscopic properties as we take the system to the continuum limit.

The results presented in this paper provide an understanding of how particles move within the periodic boundary and open boundary. The conclusions that we attained in the periodic boundary provided some key insights as a way to understand and simplify the open boundary. In both the periodic and open boundary, the transition matrices are Markov matrices and have an eigenvalue of one. Therefore, the eigenvector corresponding to the eigenvalue of one is the steady state eigenvector. In the periodic case, the steady state eigenvector is of size $\binom{N}{k}$ x 1 and has a uniform probability for each entry. However, in the open boundary, the steady state eigenvector y_{eq} is of size 2^N x 1 and is not necessarily uniform.

We sought to find a general closed form for the steady state eigenvector y_{eq} for all values of N in terms of q, a, and b. Finding the symbolic steady state eigenvector corresponding to values larger than $N = 6$ is incredibly difficult on the computer, because the corresponding transition matrix of size 2^N x 2^N is so large. However, we attempted to derive a method using pattern recognition to generate the symbolic steady state eigenvector for any value of N. This construction would have been incredibly valuable, because it would have allowed us to build the steady state eigenvector for larger values of N without solving for the eigenvalues of the transition matrix. This would have been useful for plotting and understanding the steady state behavior in the continuum limit.

Our preliminary inspection revealed that when $q = 0$, the components of steady state eigenvector can be written as polynomials in a and b, and that there exists a recursive relationship between these polynomials corresponding to the states that they represent. Ultimately, formalizing this relationship is left as an open question, but the research is promising.

With a complete analytical description of the equilibrium, we would be able to analyze the open boundary system in the continuum limit. This would allow us to completely characterize the effect of asymmetry in the system.