Return Probabilities of Skip-Free Random Walks

Andrew Lin

Mentors: Professor Leonard Schulman and Jenish Mehta

August 24, 2021

Abstract

We consider a discrete random walk on the integers in which the range of allowed jumps is limited to an interval $\{-1, 0, \cdots, (k-1)\}$ for some integer k, and we study its return probabilities after various numbers of steps. Specifically, we wish to show that if the first k return probabilities are small, so is the $(k+1)$st one. Our problem can be interpreted as an infinite-site version of a trace conjecture by Mehta and Schulman, which would verify the validity of a matrix construction demonstrating the breakdown of Cheeger's inequality for nonsymmetric matrices. We examine the random walk problem in several contexts, and we present combinatorial results demonstrating conditions under which our conjecture holds. Finally, we describe ideas for connecting the infinite-site random walk problem to the finite-matrix trace conjecture.

Introduction

The study of convergence of Markov chains in probability theory brings together mathematical tools from various fields, like theoretical computer science, numerical linear algebra, and combinatorics, which measure the properties of Markov chains in different ways. Recently [2], our mentors showed that the **edge expansion** and **eigenvalue gap**, two ways of measuring the "separation" of a Markov chain, can be wildly different, even though the classical Cheeger's inequality shows that they are much more closely related for symmetric Markov chains.

Examining the breakdown of Cheeger's inequality was the initial direction of our project, and in doing so, an initial direction of research was to resolve a matrix trace conjecture. Essentially, this trace conjecture involves adding entries to a matrix one diagonal at a time and analyzing the **traces** of various powers of that matrix. However, it was soon discovered that progress could be made more easily by studying a related problem: by interpreting matrices as Markov chains on line segments, an "infinite-line" version of the trace conjecture could be formulated. This infinite-line conjecture involves a random walk on the integer line which can move up to $(k-1)$ steps to the right or 1 step to the left, and it aims to study properties of the **return probabilities** to the initial starting point after a fixed number of steps.

In this report, we describe our main findings for those return probabilities and their application to a special case of the original trace conjecture. In particular, we show that if the first k return probabilities are small enough – less than $\frac{2}{k(k+1)}$ – then so is the $(k+1)$st. We also use that reasoning to show that the trace conjecture holds for Toeplitz matrices (identical entries on each diagonal) up until the $\frac{\sqrt{n}}{2}$th diagonal. Proofs of key results have been written out in detail in subsequent sections to document our methods and strategies for future work.

This research is joint work with Lin Lin Lee, a fellow SURF student.

Motivation: edge expansion and eigenvalue gap

Our project is initially based off the differences between the following two (paraphrased) theorems, which relate the **edge expansion** of a doubly stochastic matrix (a graph theoretic measurement of disconnectness) to its **eigenvalue gap** (which essentially determines the speed of mixing for the corresponding Markov chain):

> **Theorem 1** (Cheeger's inequality)
> Let M be any symmetric doubly stochastic matrix with edge expansion ϕ and eigenvalue gap g. Then $\frac{g}{2} \leq \phi \leq \sqrt{2g}$.

> **Theorem 2** (Mehta–Schulman, 2019 for symmetric doubly stochastic matrices)
> Let M_n be any $n \times n$ (not necessarily symmetric) doubly stochastic matrix with edge expansion ϕ and eigenvalue gap g. Then $\frac{g}{\phi} \leq 35n$. However, there are specific $n \times n$ doubly stochastic matrices A_n satisfying $\frac{g}{\phi} \geq \sqrt{n}$.

In other words, Cheeger's inequality breaks down when the matrices (and thus the corresponding Markov chains) are allowed to be nonsymmetric. (The proof of Cheeger's inequality, along with additional intuition, can be found at [5].) To determine the lower bound of Theorem 2, a family of matrices A_n was constructed with $\phi \sim \frac{1}{\sqrt{n}}$ but $g = 1$. This was done by imposing the following conditions:

- Require all eigenvalues of A_n except one to be 0. Specifically, require that the Jordan form of A_n looks like $A_n = U_n T_n U_n^*$, where U is a unitary matrix and T_n is upper triangular of the form

$$T_n = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & r_n & 0 & 0 \\ 0 & 0 & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & r_n \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

  Though doing this is not strictly necessary, it simplifies the calculations for a few reasons. For example, there are fewer free variables to adjust, and we can write down the singular values of the matrix more easily.

- Choose (based on numerical calculations for small cases) the unitary matrix U to also take on a simple form. In particular, if we choose the first row and column to be constant, we can easily characterize the first column of the resulting matrix A_n.

- Optimize (minimize) the edge expansion to make the first vertex as disconnected from the rest as possible.

The resulting construction is an $n \times n$ matrix with only six distinct nonzero entries, making it relatively easy to describe. However, this type of construction provably only achieves $\phi \sim \frac{1}{\sqrt{n}}$, and to get a tight bound in Theorem 2, we want to instead construct a family of matrices that achieves $\phi \sim \frac{1}{n}$ instead.

To do this, the following alternative strategy was proposed by Mehta and Schulman instead to construct a family of matrices B_n satisfying $\frac{\phi}{g} \sim n$:

- Pick the matrix to be Hessenberg, meaning that the entry $B_n(i,j)$ is only nonzero if $i - j \geq -1$. (In other words, the matrix is almost lower triangular.) This automatically makes $\phi \leq \frac{2}{n}$, since the first half and second half of the vertices in the Markov chain will be only connected by a single edge (in one direction).

- Choose the matrix to be almost Toeplitz. Specifically, other than the first column and last row (which will be chosen to make the matrix doubly stochastic), make each diagonal constant. This makes the number of distinct entries of the matrix $O(n)$ instead of $O(n^2)$.
- Enforce an eigenvalue gap of 1 by choosing the traces of B_n and all of its powers to be 1. This is the key point: we can control the traces more directly than with the $\sqrt{n}$ construction, because the diagonal with $i-j=k$ only starts affecting traces B_n^{k+1} and later.

This means that the matrix is constructed one diagonal at a time. First, the entries $B_n(i, i+1)$ are all chosen to be the same value r. Then the following process is repeated n times (using the counter $1 \leq c \leq n$): the "edge" entries $B_n(c, 1)$ and $B_n(n, n+1-c)$ are chosen so that the cth row and $n-c$th column sum to 1, and the remaining entries on the diagonal satisfying $i-j=c-1$ are chosen so that the trace of B_n^c is 1. Finally, after completing that recursive process, the value of r is chosen so that B_n^n also has trace 1 – from numerical calculations and small cases, this value is $r = n^{-1/(n-1)}$.

What remains to be done in this construction is twofold: first, it must be proved that $r = n^{-1/(n-1)}$ is the correct value to ensure all traces are 1, and second, it must be shown that the entries of the matrix are all nonnegative in this recursive process, so that we have a valid doubly stochastic matrix.

Numerical calculations have showed that the entries are indeed nonnegative even for values like $n = 2000$. In fact, asymptotics can be conjectured – the smallest positive entry of B_n seems to be $B_n(n-1, 2)$, and it looks to be approximately $\frac{0.48}{n^2}$. But we have found no direct ways of proving this fact, so our analysis has taken a turn. As mentioned, entries far from the diagonal only start affecting the traces later on in the process. Since adding positive entries to our matrix only increases traces, if we want to make all the traces 1 eventually, we just need to make sure traces don't get too large too quickly in the recursive process. This motivates the following conjecture:

Conjecture 3
Let M be a matrix with nonzero entries only within the diagonals $-1 \leq i-j \leq k-1$. If M is substochastic, and the traces of $M, M^2, \cdots, M^k$ are all at most 1, then the trace of M^{k+1} is at most 1.

Proving this trace conjecture would almost be enough to verify the validity of the B_n construction – it would additionally also need to be proven that taking the first step in the repeated process above (choosing $B_n(c, 1)$ and $B_n(n, n+1-c)$) would still keep the trace below 1.

From the finite segment to the infinite line

While the trace conjecture can be stated without many extra conditions, and the problem statement itself is natural, there are many degrees of freedom ($O(nk)$) in its current stated form, and this makes studying the eigenvalues or traces of potential matrices difficult. Thus, the first simplification we may make is to consider only Toeplitz matrices (constant along each diagonal), which would give us only $O(k)$ degrees of freedom. Still, though, the formulas for the traces would then depend not only on the entries of the matrix but also the size n of the matrix because of boundary conditions.

This motivates us to eliminate the boundary conditions and the dependence on n completely. To do this, we think of a Toeplitz matrix as a random walk on a finite line segment with absorbing boundary conditions, in which the probabilities of making jumps of each length are the same regardless of the current position. (For example, if the main diagonal entries are 0.3, then at each step, there is a 30 percent chance that the walk stays in place.) We can then imagine taking the length of this line segment to infinity, giving us a translation-invariant random walk. Now instead of looking at the traces of the (now-infinite) transition matrices, we can just look at a single diagonal entry (which is essentially a rescaled trace). We can now state our modified "infinite-line" problem:

Problem 4

Suppose that we fix some integer k, just like in the trace conjecture, and we have jump probabilities of $a_{-1}, a_0, \cdots, a_{k-1}$ in our random walk at each step. Define p_i to be the probability that, if the walk on the integer line starts at the origin, we are back at the origin after i steps. (Note that this definition is suppressing the dependence on k and the jump probabilities $a_{-1}, a_0, \cdots, a_{k-1}$). Prove that if $p_1, \cdots, p_k \le \frac{1}{k+1}$, then $p_{k+1} \le \max(p_1, \cdots, p_k)$.

A random walk on the integer line where we can only move at most one step to the left is known as a **skip-free** random walk, because the walk cannot skip over vertices as it moves to the left. The reason for the choice $\frac{1}{k+1}$ is partially because our matrix must be at least size $(k+1)$ to have a meaningful kth diagonal to "add entries" as in the B_n construction, but further justification will be shown later.

Intuitively, here is one way to think about our infinite-line conjecture: in this skip-free random walk, we can only move to the left with small steps, so any step of length L to the right needs to be balanced out by L steps of one to the left, creating a total "cycle" of length $L+1$. Since the longest such cycle we are allowed finishes in k steps, it makes sense to believe that the chance of being back will not suddenly be large in the long run.

Brute-force calculations

We begin by describing the most direct approach for computing these return probabilities. As described in the section above, we can write the return probabilities p_i as polynomials in our jump probabilities $a_{-1}, a_0, \cdots, a_{k-1}$.

Example 5

For $k = 3$ (meaning that we're allowed to have jumps of size $-1, 0, 1$, or 2), we have $p_1 = a_0, p_2 = a_0^2 + 2a_1a_{-1}, p_3 = a_0^3 + 6a_1a_0a_{-1} + 3a_2a_{-1}^2$, and $p_4 = a_0^4 + 12a_1a_0^2a_{-1} + 12a_2a_{-1}^2 + 6a_1^2a_{-1}^2$.

Once we have these algebraic expressions, relatively direct approaches can be taken to arrive at results in special cases. First of all, it is natural to try to show that not only is the $(k+1)$th return probability smaller than the first k, but also all of the future return probabilities are also smaller. This is because of the following special case:

Lemma 6

If the only allowed jumps in our skip-free random walk are of size -1 or $\ell - 1$ (for some $\ell \le k$), then the return probabilities satisfy $p_\ell > p_{2\ell} > p_{3\ell} > \cdots$, and all other return probabilities are zero.

Proof. In this situation, all jumps add -1 to our position mod ℓ, meaning that we only have a nonzero return probability when the number of steps is a multiple of ℓ. The closed-form expression for the return probabilities is

$$p_{c\ell} = \binom{c\ell}{c} p^{(\ell-1)c}(1-p)^c,$$

since we return after $c\ell$ steps by making c right jumps and $(\ell-1)c$ left jumps, which can occur in any order. It suffices to prove that $\frac{p_{(c+1)\ell}}{p_{c\ell}} < 1$ for all c. Indeed, notice that

$$\frac{p_{(c+1)\ell}}{p_{c\ell}} = \frac{\binom{(c+1)\ell}{c+1}}{\binom{c\ell}{c}} \frac{p^{(\ell-1)(c+1)}(1-p)^{c+1}}{p^{(\ell-1)c}(1-p)^c} = \frac{(c\ell+\ell)!c!(c\ell-c)!}{(c\ell)!(c+1)!(c\ell+\ell-c-1)!} p^{(\ell-1)}(1-p)$$

(expanding out the binomial coefficients). The only dependence on p in this expression is the last part, $p^{(\ell-1)}(1-p) = p^{\ell-1} - p^{\ell}$, and this expression is maximized when the derivative satisfies $(\ell-1)p^{\ell-2} - \ell p^{\ell-1} = 0 \implies p = \frac{\ell-1}{\ell}$. In other words, the **ratio between return probabilities is maximized in the case where our walk has zero drift**. Plugging this in and simplifying the factorials slightly yields

$$\frac{p_{(c+1)\ell}}{p_{c\ell}} \le \frac{(c\ell+\ell)_\ell}{(c\ell+\ell-c-1)_{\ell-1}} \cdot \frac{1}{c+1} \cdot \left(\frac{\ell-1}{\ell}\right)^{\ell-1} \cdot \frac{1}{\ell},$$

where $n_k = n(n-1)\cdots(n-k+1)$ is the falling factorial. This expression further simplifies by pulling the $(c\ell+\ell)$ term out of the numerator and cancelling with the $c+1$ and ℓ in the denominator:

$$\frac{p_{(c+1)\ell}}{p_{c\ell}} \le \frac{(c\ell+\ell-1)_{\ell-1}}{(c\ell+\ell-c-1)_{\ell-1}} \cdot \left(\frac{\ell-1}{\ell}\right)^{\ell-1}.$$

Finally, pair up the corresponding terms in the falling factorials, which gives us fractions of the form $\frac{c\ell+j}{c\ell-c+j}$ for $j \in \{1, 2, \cdots, \ell-1\}$. Since

$$\frac{c\ell+j}{c\ell-c+j} < \frac{c\ell}{c\ell-c} = \frac{\ell}{\ell-1}$$

for each j, we finally find that (applying this inequality to all $\ell-1$ terms of the falling factorial)

$$\frac{p_{(c+1)\ell}}{p_{c\ell}} < \left(\frac{\ell}{\ell-1}\right)^{\ell-1}\left(\frac{\ell-1}{\ell}\right)^{\ell-1} = 1,$$

as desired. Thus the return probabilities (for multiples of ℓ) are indeed monotonic. □

In particular, since p_ℓ will always be the largest return probability and $\ell \le k$, this proves that Problem 4 holds in this very simple case. The idea then is that each "raw cycle type" is decreasing in probability (since any positive jump always needs to be canceled out with -1s due to the allowed set of jumps), so even when we combine different types of cycles together, the probabilities should still generally decrease. However, this does **not** hold in general:

Example 7

Let $k = 10$, and suppose we have a situation where the only nonzero jump probabilities are $a_0 = 0.1$, $a_9 = 0.09$, and $a_{-1} = 0.81$. Then $p_{11} > p_{10}$, so it is not true that one of the first k return probabilities is the largest (and in fact, $p_{k+1} > \max(p_1, \cdots, p_{10})$ here).

Intuitively, the only way to be back at the origin after a path of $(k+1)$ steps is to attach an a_0 self-loop to a path of length k steps. So every path has its probability multiplied by a factor of a_0, but we gain up to a factor of $(k+1)$ more paths of each type because of the "entropy" coming from choosing where that a_0 is added (less if the paths already contained a_0s). And in this special example above, the factor of a_0 is large enough (in particular, bigger than $\frac{1}{k+1}$) that the entropic term dominates. This explains why we require $p_1, \cdots, p_k \le \frac{1}{k+1}$ in our problem, and in fact this argument can be turned around to show a special case of Problem 4:

Lemma 8

Suppose that in our skip-free random walk, the only allowed positive jumps are of size at least $\frac{k}{2}$ (but jumps of 0 or -1 are allowed). If $p_1, \cdots, p_k \le \frac{1}{k+1}$, then $p_{k+1} \le \max(p_1, \cdots p_k)$.

We omit the details, but the main step in the proof is that when we only have large jumps, we can only take at most one of them to get back to the origin in k or $k+1$ steps.

It is difficult to say more in generality with these polynomials in a_i, because their closed form expressions can be thought of as a weighted sum over partitions. In particular, notice that we can write the (not very illuminating, but basically another way to write out the casework in Example 5) expression

$$p_i = \sum_{\substack{(c_1,\cdots,c_i)\\ c_1+\cdots+c_i=i\\ 0\le c_1,\cdots,c_i\le k+1}} \binom{i}{c_1,\cdots,c_i} \prod_{j=1}^{i} a_{c_j}.$$

This essentially means we have a sum over partitions of i (except missing one term for p_{k+1}), but the theory of partitions rarely seems to weight partitions in this particular way. And even though there are only k free variables in our jump probabilities, leading us to always be able to write p_{k+1} as a linear combination of $p_1, \cdots, p_k$, this does not immediately help us solve our problem. In particular, nonlinear constraints such as $p_2 > p_1^2$ still need to be imposed for our infinite-line conjecture to hold, and those are difficult to impose if we ignore the jump probabilities and try to reason algebraically.

For this reason, we turn to more computational efforts again. The intuition behind this next result comes partially from the proof of Lemma 6 and also partially in the following statement: a walk which tends to move to the left or to the right will be even less likely to return to the origin than one which generally stays in place.

Proposition 9

It suffices to prove the infinite-line conjecture when the **drift** $\mu = -a_{-1} + a_1 + 2a_2 + \cdots + (k-1)a_{k-1}$ of the walk is zero.

Proof. The return probabilities after a fixed number of steps are polynomials in our variables $a_{-1}, a_0, \cdots, a_{k-1}$, and if we make the substitution (call this a **re-bias**)

$$(a_{-1}, a_0, a_1, \cdots, a_{k-1}) \mapsto (\tfrac{a_{-1}}{x}, a_0, a_1 x, \cdots, a_{k-1}x^{k-1}),$$

for any positive x, that keeps the polynomials invariant. Define a function $f : \mathbb{R}^+ \to \mathbb{R}^+$ via $f(x) = \frac{a_{-1}}{x} + a_0 + a_1 x + \cdots + a_{k-1}x^{k-1}$ (this is the sum of the jump rates). Notice that our re-biased walk will have zero drift when

$$-1 \cdot \frac{a_{-1}}{x} + 1 \cdot a_1 x + \cdots + (k-1) \cdot a_{k-1}x^{k-1} = 0,$$

and it will do so for some positive value of x (because for $x \to 0$ the walk is extremely left-biased, and for $x \to \infty$ it is extremely right-biased). However, also notice that the function f attains a minimum[1] when

$$f'(x) = -1 \cdot \frac{a_{-1}}{x^2} + 1 \cdot a_1 + \cdots + (k-1)a_{k-1}x^{k-2} = 0.$$

Since $x = 0$ is not a solution, multiplying this equation by x yields this equation, which means that the **minimum value for $f(x)$ occurs exactly when the walk has zero drift**. Furthermore, because this minimum occurs when (multiplying by x again) $a_{-1} = a_1 x^2 + \cdots + (k-1)a_{k-1}x^k$, and the right-hand side is strictly increasing for positive x, there is a **unique** solution for x. In other words, the zero-drift case does indeed minimize the function $f(x)$.

So if we have an arbitrary set of jump probabilities $\{a_{-1}, a_0, \cdots, a_{k-1}\}$ satisfying $a_{-1} + a_0 + \cdots + a_{k-1} = 1$, and we re-bias it with some value x to make it drift-free, the re-biased, drift-free "jump rates" will have sum $f(x) \le f(1) = 1$. (Remember that $f(1) = a_{-1} + a_0 + \cdots + a_{k-1}$ is the original sum of probabilities.)

But now the re-biased (drift-free) $\{\frac{a_{-1}}{x}, a_0, a_1 x, \cdots, a_{k-1}x^{k-1}\}$ is a "scaled-down" version of some drift-free prob-

[1] One important detail (for why this is a minimum and not a maximum): the following paragraph shows that there is **only one solution** x to this equation. Since f is very large for $x \to 0$ and $x \to \infty$, the critical point at x must be a minimum.

ability distribution. Notice that if we have already proved the drift-free case for probability distributions, reducing all jump rates by some constant factor c will still make our return probability conjecture hold (since return probabilities are multiplied by a factor of c^j, making later ones smaller). So the re-biased (drift-free) $\{\frac{a_{-1}}{x}, a_0, a_1 x, \cdots, a_{k-1}x^{k-1}\}$ will satisfy our conjecture, which means that $\{a_{-1}, a_0, \cdots, a_{k-1}\}$ will also satisfy our conjecture, because it has the same return probabilities as $\{\frac{a_{-1}}{x}, a_0, a_1 x, \cdots, a_{k-1}x^{k-1}\}$. □

This result means that we can eliminate yet another free variable in our problem, and for instance, it allows us to prove the case $k = 3$ for Problem 4:

Example 10

For $k = 3$, the drift-free case can be characterized by setting $a_1 = x, a_2 = y, a_{-1} = x + 2y, a_0 = 1 - 2x - 3y$. Then p_1, p_2, p_3, p_4 can be written as polynomials in x and y, and $\max(p_1, p_2, p_3) - p_4$ can be shown to always be nonnegative in the allowed range of values for x and y. Thus $p_4 \leq \max(p_1, p_2, p_3)$ and the $k = 3$ case is proved.

These are the results we have obtained through direct manipulation of our jump probabilities, but these methods are relatively crude. Other results have been found through more theoretical techniques, which we will now describe.

Linear algebraic results

Recall that we switched our attention from a finite matrix and its traces to the infinite-line random walk and its return probabilities, because we wanted to increase the amount of symmetry in the problem and allow for more direct calculations. However, we can actually make progress on the infinite-line conjecture by switching back to matrices in the following way:

- Instead of using doubly stochastic matrices like in the original motivation, we construct **Toeplitz or circulant matrices** with the jump probabilities (that is, matrices with equal entries M_{ij} on each diagonal of constant $(j-i)$, potentially with wraparound).
- Calculate the eigenvalues of these new matrices, and use them to find the traces of matrix powers (sums of powers of eigenvalues).
- Notice that, up to some "edge effects," the traces that we compute are basically n times the return probabilities in the infinite-line walk. Taking $n \to \infty$, if we can prove that the edge effects are lower order than the traces that are computed, then the return probabilities will need to satisfy the same inequalities as the traces.

We can see this in action in the following application:

Proposition 11

The infinite-line conjecture holds for $k = 2$. In fact, p_2 will be larger than **all** future return probabilities, so $p_a \leq \max(p_1, p_2)$ for all $a > 2$.

Proof. Suppose we have jump probabilities a_{-1}, a_0, a_1. Construct a Toeplitz matrix T_n with a_0 on the main diagonal, a_1 on the diagonal immediately above it, a_{-1} on the diagonal immediately below it, and 0s everywhere else. By [1], this matrix has eigenvalues

$$\lambda_k = a_0 - 2\sqrt{a_1 a_{-1}} \cos\left(\frac{k\pi}{n+1}\right).$$

In particular, notice that the magnitude of the second term is at most $2\sqrt{a_1 a_{-1}} \le a_1 + a_{-1} = 1 - a_0$ by AM-GM, so **all eigenvalues are real and of magnitude at most** 1. Therefore, $\mathrm{Tr}(T_n^2)$ will be at least as large as $\mathrm{Tr}(T_n^a)$ for any $a > 2$, because $\lambda_k^a \le \lambda_k^2$ for each eigenvalue λ_k.

But now if we think about the diagonal entries of our matrix T_n^a, any diagonal entry that is not in the first or last a rows will be exactly p_a, the infinite-line return probability after a steps. (This is because we can only move by one row per step, so the edge effects will not be able to make an impact.) Since each diagonal entry is between 0 and 1 inclusive (as this Toeplitz matrix is strictly dominated by a stochastic one), this means that for each positive integer a, we have

$$(n-2a)p_a \le \mathrm{Tr}(T_n^a) \le (n-2a)p_a + 2a.$$

Finally, using the fact that $\mathrm{Tr}(T_n^a) \le \mathrm{Tr}(T_n^2)$, we find that

$$(n-2a)p_a \le \mathrm{Tr}(T_n^a) \le \mathrm{Tr}(T_n^2) \le (n-4)p_2 + 4.$$

(Remember that this inequality must hold for all n, for any fixed a.) Assuming for the sake of contradiction that $p_a = p_2 + \varepsilon$ for some positive ε, we find that

$$(n-2a)(p_2+\varepsilon) \le (n-4)p_2 + 4 \implies (n-2a)\varepsilon \le (2a-4)p_2 + 4,$$

and this yields a contradiction by picking any $n \ge 2a + \frac{(2a-4)p_2+4}{\varepsilon}$. Thus $p_a \le p_2$, as desired. □

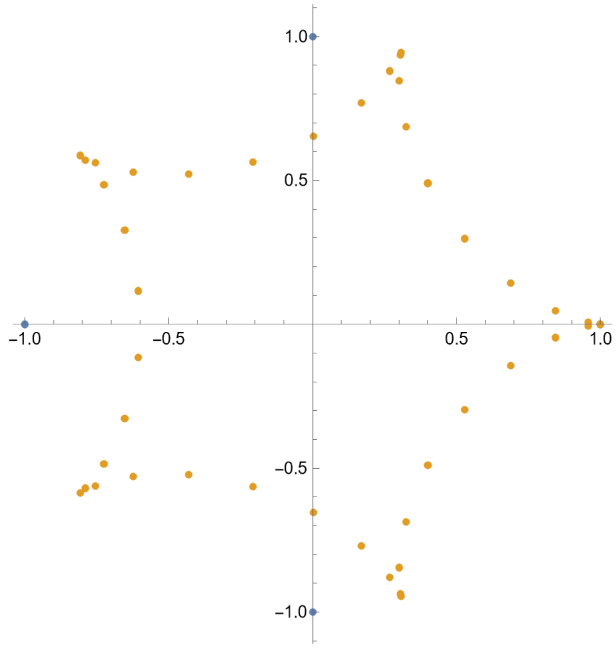
Unfortunately, this strategy does not extend to larger k. In [4], a method for computing the eigenvalues of more general banded Toeplitz matrices is described: essentially, we treat the eigenvalue equation as a linear recurrence with fixed boundary conditions, which requires us to solve a polynomial equation of degree k and then plug in those roots into a Vandermonde-style determinant. This indirect strategy means that finding closed-form expressions for the eigenvalues is much more difficult, and we have not extended the $k = 2$ case with Toeplitz matrices.

However, circulant matrices have even more structure than Toeplitz matrices, and that helps us when calculating eigenvalues and eigenvectors. Instead of requiring M_{ij} to be constant across each value of $j - i$, we now require M_{ij} to be constant across each value of $(i - j) \bmod n$, meaning that the columns of the matrix are just cyclic shifts of each other. The following fact can be directly checked by matrix computation:

Fact 12

In any $n \times n$ circulant matrix with diagonal entries $M_{ij} = a_\ell$ for $j - i = \ell \bmod n$, the vector $v = (1, \omega, \cdots, \omega^{n-1})^T$ is always a right eigenvector for any nth root of unity ω, with corresponding eigenvalue $a_0 + \omega a_1 + \omega^2 a_2 + \cdots + \omega^{n-1} a_{n-1}$.

To analyze those eigenvalues more explicitly, we follow the strategy of Lemma 8 and consider the case where only two of the a_is are nonnegative, specifically $a_{-1} = a_{n-1}$ and $a_{\ell-1}$ for some positive integer ℓ. For example, below is a plot of eigenvalues (in yellow) corresponding to the "drift-free" case $a_{-1} = 0.8$, $a_4 = 0.2$ for a 43×43 circulant matrix. The eigenvalues appear to trace out a star-like shape in the complex plane:



We can explain this shape and generalize it with the argument below.

Lemma 13

Let $M_{n,k}$ be the circulant matrix with $a_{-1} = a_{n-1} = \frac{k-1}{k}$ and $a_{k-1} = \frac{1}{k}$. Then all eigenvalues of $M_{n,k}$ will lie on the **hypocycloid** created by rotating a circle of radius $\frac{1}{k}$ inside the unit circle.

Proof. Since $\omega^{n-1} = \frac{1}{\omega}$ for any nth root of unity ω, all eigenvalues of this circulant matrix will be of the form

$$\lambda = \frac{1}{k}\omega^{k-1} + \frac{k-1}{k}\frac{1}{\omega}$$

for some nth root of unity ω. Letting $\omega = \cos(x) + i\sin(x)$, we find that

$$\mathrm{Re}(\lambda) = \frac{1}{k}\cos((k-1)x) + \frac{k-1}{k}\cos x, \quad \mathrm{Im}(\lambda) = \frac{1}{k}\sin((k-1)x) - \frac{k-1}{k}\sin(x).$$

This is exactly the parameterization of the hypocycloid described above. □

In addition, because the matrices $M_{n,k}$ all have the same eigenvectors for fixed n, and circulant matrices are diagonalizable, we can actually take linear combinations of these $M_{n,k}$ matrices. For example, we can now consider

$$M(s,t) = sM_{n,1} + tM_{n,2} + (1-s-t)M_{n,3},$$

which is a family of matrices allowing for "arbitrary jump probabilities a_1, a_2, a_3, a_{-1} that are drift-free." In such a situation, our eigenvalues will lie in the convex hull of the hypocycloids traced out by $M_{n,1}$, $M_{n,2}$, and $M_{n,3}$, and this is potentially useful because we can now bound the magnitude of the eigenvalues along various complex arguments. (For example, if we are trying to make progress on the infinite-line or trace conjecture for $k = 5$, then we need to bound $\mathrm{Tr}(A^6)$, and thus it can be helpful to know that eigenvalues cannot have magnitudes too close to 1 when their arguments are close to $\frac{2\pi a}{6}$ for some integer a.) However, turning this intuition into a more precise statement has not yielded any inequalities similar to the trace or infinite-line conjectures yet.

Combinatorial results

We now turn our attention to methods of studying the infinite-line random walk which do not rely on such involved algebraic calculations as above. In particular, we can study the structure of the path that the random walk traces out to obtain a more qualitative result:

Lemma 14

Consider a path of length $(k+1)$ on the integer line in which every step taken is one of $\{-1, \cdots, (k-1)\}$ and the walk returns to the origin. Then the path must revisit some integer.

Proof. First of all, the set of sites visited by this path forms a (connected) line segment of the integers which contains 0. This is because any negative site x visited must have been visited by taking a jump of -1 from $x+1$ (induct up to 0), and any positive site y visited must return to the origin by taking a jump of -1 to $y-1$ (induct down to 0). In particular, this also means that if the line segment is of length ℓ, there must be at least $\ell - 1$ jumps of size -1.

However, there can only be at most $k-1$ jumps of size -1, because if there are k jumps of size -1, the total displacement of the path is at most $k \cdot -1 + 1 \cdot (k-1) < 0$. So the number of jumps is at most $k-1$ and at least $\ell - 1$, meaning that $\ell \leq k$. So by the Pigeonhole principle, since there are more opportunities to visit sites than distinct sites that can be visited, at least one site must be visited twice. □

This observation can be restated as "some of the cyclic shifts of any path of length $(k+1)$ returning to the origin must actually hit the origin again in the middle" (if we cyclically shift to the first return point of the repeated vertex). This is enough to prove a weaker version of the infinite-line conjecture for general k:

Theorem 15

Suppose that we have a skip-free random walk where the only allowed jumps are $\{-1, \cdots, (k-1)\}$, and $p_1, \cdots, p_k \leq p$ for some $p \leq \frac{2}{k(k+1)}$. Then $p_{k+1} \leq p$, and therefore $p_{k+1} \leq \max(p_1, \cdots, p_k)$.

(The original problem wished to show this for $p \leq \frac{1}{k+1}$.)

Proof. Let x be the total probability that after $(k+1)$ steps, the random walk has returned to the origin, and it also revisited the origin somewhere in the middle of the path. Notice that $x \leq p_1 p_k + p_2 p_{k-1} + \cdots + p_k p_1$ by the definition of the p_is (since the right-hand side is a union bound on paths that return to the origin after $(k+1)$ steps, doing casework on the first return step).

On the other hand, we have $\frac{2}{k+1} p_{k+1} \leq x$. This is because of the following logic: group the set of all paths that return after $(k+1)$ steps into equivalence classes of cyclic shifts. By our previous lemma, there are two points in the walk that visit the same vertex, so if we cyclically shift our path within one of these equivalence classes, using either of those two points as starting vertices will give us a walk that returns to the origin in the middle (at the other point). If the two cyclic shifts are distinct, then because there are at most $(k+1)$ cyclic shifts possible, at least $\frac{2}{k+1}$ of the paths in this equivalence class do indeed return. If they are not, that means the walk is c copies of the same sequence appended together for some $c \geq 2$, and thus there are $\frac{k+1}{c} \leq \frac{k+1}{2}$ cyclic shifts in total (and $\frac{1}{(k+1)/2} = \frac{2}{k+1}$). Since the probability of all sequences of a given equivalence class are equal (they are all the product of the corresponding a_is), at least $\frac{2}{k+1}$ of the total return probability after $(k+1)$ steps must return to the origin, as desired.

So putting these inequalities together, we have

$$\frac{2}{k+1} p_{k+1} \leq p_1 p_k + \cdots + p_k p_1 \implies p_{k+1} \leq \frac{k+1}{2}(p_1 p_k + \cdots + p_k p_1) \leq \frac{k(k+1)}{2} p^2,$$

which is at most p by our problem statement. □

One way to summarize this proof in words is that "if it is unlikely for the path to return once to the origin, it is even more unlikely for it to return multiple times, which must happen a substantial fraction of the time for longer paths." Though some of the inequalities here are crude (for example, uniformly bounding $p_1, \cdots, p_k$ by p), there are special cases under which each of them are nearly tight, so trying to obtain stronger constant factors on the inequalities is difficult to do directly. However, Theorem 15 can indeed be generalized by reasoning about walks that are longer than $(k+1)$ steps as well. We write out the proofs below, but the next two results are basically more algebraically involved versions of the proof above.

Lemma 16

Consider a walk in which every step taken is one of $\{-1, \cdots, (k-1)\}$ and the walk returns to the origin. Then if the path is of length ℓ, and $ak < \ell \leq (a+1)k$, then at least a of the cyclic shifts of this walk will return to the origin in the middle of the walk.

Proof. Again, the set of sites visited by this path forms a (connected) line segment of the integers which contains 0. (This part of the proof is the same as in the above Pigeonhole argument.)

However, there can be at most $\ell - a - 1$ jumps of size -1 in this path. Suppose otherwise, so that there are at least $\ell - a$ jumps of size -1. Then the largest final position that is possible for the path is

$$(\ell - a) \cdot (-1) + a \cdot (k-1) = -\ell + ak < 0,$$

a contradiction with the walk returning to the origin. Therefore, the line segment is of length at most $\ell - a$. Since there are at least a more steps than there are spots the path could have landed on, at least a of the steps must be returning to a previously visited vertex, meaning that starting the walk from that point (cyclically) would result in an early return to the origin, as desired. □

Proposition 17

Consider a walk in which every step taken is one of $\{-1, \cdots, (k-1)\}$. Assume that the first k return probabilities $(1 \leq m \leq k)$ satisfy $p_m \leq \frac{c}{km^x}$ for some $x > 1$ and $c \leq \frac{1}{\zeta(x)4^{x+1}}$. Then $p_m \leq \frac{c}{km^x}$ for all m.

Proof. We proceed by induction. The base case is the first k return probabilities given to us, and for the induction step, assume that $p_m \leq \frac{c}{km^x}$ for $1 \leq m \leq n$. Write out all sequences of jumps of length $(n+1)$ that return to the origin, and group them into groups of cyclic shifts. Applying Lemma 16 for $m = n+1$, we can (crudely) say that at least $\frac{1}{2k}$ of the cyclic shifts will return to the origin (since regardless of the value of m, the fraction of returns is always at least $\frac{a}{(a+1)k} \geq \frac{1}{2k}$). Let the set of paths that return to the origin have probability x; we are saying here that

$$\frac{p_{n+1}}{2k} \leq x.$$

However, because all blue paths must return to the origin at some point, we have

$$x \leq p_1 p_n + p_2 p_{n-1} + \cdots + p_n p_1.$$

Putting these together, we have the inequality

$$p_{n+1} \leq 2k(p_1 p_n + \cdots + p_n p_1).$$

We now apply the induction hypothesis and find that

$$p_{n+1} \leq 2k \cdot \frac{c^2}{k^2}\left(\frac{1}{1^x n^x} + \frac{1}{2^x n^x} + \cdot + \frac{1}{n^x 1^x}\right).$$

Now pull out a factor of 2 from the sum (only keeping the first half) and then lower bound the second term in the denominator by $\left(\frac{n}{2}\right)^x$ to get

$$\begin{aligned} p_{n+1} &\leq \frac{4c^2}{k}\left(\frac{1}{(n/2)^x}\right)\left(\frac{1}{1^x} + \frac{1}{2^x} + \cdots + \frac{1}{\lceil n/2\rceil^x}\right) \\ &\leq \frac{4^{x+1}c^2}{k(n+1)^x}\left(\frac{1}{1^x} + \frac{1}{2^x} + \cdots\right) \\ &\leq \frac{c}{k(n+1)^x}\zeta(x)4^{x+1}c \\ &\leq \frac{c}{k(n+1)^x}, \end{aligned}$$

where we've made the crude bound $\frac{n}{2} \geq \frac{n+1}{4}$ in the second line. This completes the induction and proof. ☐

Corollary 18

If we have the same walk as above, and the first k return probabilities are all at most $\frac{c}{k^{x+1}}$ for some $c \leq \frac{1}{\zeta(x)4^{x+1}}$, then one of the first k return probabilities will be the largest (by picking c that achieves equality).

We conclude this section by mentioning one thought which has been pursued regarding the fraction of paths that do return to the origin but has not been fruitful so far. From numerical testing, it seems plausible that if the allowed jumps in a given path have magnitude at most $\frac{k}{2}$, then at least a constant fraction c of its permutations will return to the origin. In particular, this would turn the inequality $p_{k+1} \leq \frac{k+1}{2}(p_1 p_k + \cdots + p_k p_1)$ into one that looks like $p_{2k+1} \leq \frac{1}{c}(p_1 p_{2k} + \cdots + p_{2k} p_1)$, which would allow us to reason about probabilities $p \sim \frac{1}{k}$ instead of $\frac{1}{k^2}$. We also have the following special case to support this hypothesis:

Lemma 19

Consider a walk with only jumps of $+r$ or -1 that returns to the origin, so that we have n $+r$ jumps and nr -1 jumps for some $n \geq 2$. Then at least half of its permutations will return to the origin in the middle.

Proof. We can restate the problem as follows: consider the set of all up-right paths from $(0,0)$ to (rn, n). We can place this set in bijection with the permutations of the jumps in the lemma statement by representing $+r$s as up steps and -1s as down steps, and returning to the origin is equivalent to hitting the line $x = ry$ at some lattice point other than $(0,0)$ and (rn, n). The paper [3] provides a closed-form solution for the number of paths that do not hit the line $x = ry$ (this corresponds to $k = 1$ in the paper), and it is

$$\frac{r(r+1)}{(r+1)n-1}\binom{(r+1)n-1}{rn}.$$

Dividing this by the total number of paths, $\binom{(r+1)n}{rn}$, yields $\frac{r}{(r+1)n-1}$. Since this denominator is at least $2r$ when $n \geq 2$, the fraction of paths that do not return to the origin in our original formulation is at most $\frac{1}{2}$, and indeed at least half of the permutations do return to the origin in the midde of the path. ☐

To extend this to the case where there are different positive jumps, an even more generalized ballot theorem or other technique will be necessary than what is provided in [3], but we have not been able to provide such a result.

Returning to the original problem

Recall that the trace conjecture is as follows: suppose we have an $n \times n$ matrix in which the only nonzero diagonals are $a_{-1}, \cdots, a_{k-1}$. If we further assume that the matrix is substochastic and that $\text{Tr}(A), \cdots, \text{Tr}(A^k) \le 1$, then $\text{Tr}(A^{k+1}) \le 1$. In full generality, we have no results, but the infinite-line conjecture does help us prove a particular case of this trace conjecture:

Proposition 20

The trace conjecture holds if we have a Toeplitz matrix, and $k \le \frac{\sqrt{n}}{2}$ (or equivalently $n \ge 4k^2$).

Proof. First, we return to the original blue-paths inequality in the form (this is slightly weaker than the original)

$$p_{k+1} \le k(p_1 p_k + \cdots + p_k p_1).$$

Notice that if $p_1, \cdots, p_k \le \frac{1}{k^2}$, then $p_{k+1} \le \frac{1}{k^2}$ as well. This time, we will use a stronger result, essentially that if $p_1, \cdots, p_k \le \frac{1}{2k^2}$, then $p_{k+1} \le \frac{1}{4k^2}$. Specifically, **if** $\max(p_1, \cdots, p_k) \le \frac{1}{2k^2}$, then $p_{k+1} \le \frac{1}{2}\max(p_1, \cdots, p_k)$.

Now notice that because our matrix is of size $n \ge 4k^2$, and our walk cannot stray more than k^2 steps away from the origin within $(k+1)$ steps, the middle diagonal entries of A^m for $1 \le m \le k+1$ (all except the k^2 at the top and bottom) will just be p_m, like in the infinite-line case. (This value of k^2 is extremely loose, but it's needed for another assumption later anyway.) Therefore, we have

$$\text{Tr}(A^m) \ge (n - 2k^2)p_m \ge \frac{n}{2} p_m$$

for all $1 \le m \le k$. On the other hand, because restricting to a finite matrix only decreases the diagonal entries from the infinite case, we also have

$$\text{Tr}(A^{k+1}) \le n p_{k+1}.$$

Putting everything together, this means that under our assumption $\max(p_1, \cdots, p_k) \le \frac{1}{2k^2}$, we have

$$\text{Tr}(A^{k+1}) \le n p_{k+1} \le \frac{n}{2}\max(p_1, \cdots, p_k) \le \max(\text{Tr}(A), \text{Tr}(A^2), \cdots, \text{Tr}(A^k)).$$

Finally, we remove the need for our additional assumption as follows: if we further assume $\text{Tr}(A), \cdots, \text{Tr}(A^k) \le 1$, that means that we also assume $\frac{n}{2} p_m \le \text{Tr}(A^m) \le 1$ for all $1 \le m \le k$. But that means we are assuming $p_m \le \frac{2}{n} \le \frac{1}{2k^2}$ already, so that assumption is satisfied and we get the desired statement that we want. □

In summary, what's going on here is that the factor-of-2-smaller p_{k+1} allows us to lose a factor of 2 when going from the infinite line to the finite matrix, keeping $\text{Tr}(A^{k+1})$ smaller than the previous traces. And as long as the matrix is large enough, we no longer need to make an artificial assumption about the probabilities being small, since that's already encoded in the traces being less than 1. It may be surprising that we need to lose a factor of 2 here, but the fraction of "probability" lost when going from the infinite-line case to the finite Toeplitz matrix is not always monotone, so this is the cleanest bound that we are able to produce at the moment.

Few of the other strategies employed for the infinite-line problem (such as rebiasing) work for general **non**-Toeplitz matrices, but there is one current direction of research that may yield some insight. There is noticeable structure in the eigenvalues of matrices which have nonzero entries in only two bands (for example -1 and $(k-1)$ for some positive integer k), and we have the following conjecture:

Conjecture 21
Let M be a matrix which has nonnegative entries M_{ij} on the diagonals $j - i = -1$ and $j - i = (k - 1)$ for some $k \geq 2$, and zeros everywhere else. Then all eigenvalues lie on the rays connecting 0 to one of the kth roots of unity in the complex plane, and the set of eigenvalues is k-fold symmetric about the origin.

The k-fold symmetry can be seen by noticing that the only nonzero coefficients of the characteristic polynomial are $\lambda^n, \lambda^{n-k}, \lambda^{n-2k}, \cdots$, so up to a few factors of λ, we have a polynomial in λ^k which we must show has real roots. However, this has not been done. Furthermore, because the eigenvectors of different matrices of this form can look wildly different, there is not currently a known way to generalize this to the banded matrices in the trace conjecture like we have for the circulant case.

Acknowledgements

The author would like to thank Professor Leonard Schulman and Jenish Mehta for their guidance in obtaining background, attacking the problem, providing new insights, and working through calculations throughout the SURF, as well as the SFP SURF office and the NSF for providing funding and managing the logistics of the program.

References

[1] D. Kulkarni, D. Schmidt, and S.-K. Tsui. Eigenvalues of tridiagonal pseudo-toeplitz matrices. *Linear Algebra and its Applications*, 297(1):63–80, 1999.

[2] J. C. Mehta and L. J. Schulman. Edge Expansion and Spectral Gap of Nonnegative Matrices. https://epubs.siam.org/doi/pdf/10.1137/1.9781611975994.73, 2019.

[3] M. Spivey. Enumerating lattice paths touching or crossing the diagonal at a given number of lattice points. *The Electronic Journal of Combinatorics [electronic only]*, 19, 08 2012.

[4] W. F. Trench. On the eigenvalue problem for toeplitz band matrices. *Linear Algebra and its Applications*, 64:199–214, 1985.

[5] L. Trevisan. Lecture Notes on Expansion, Sparsest Cut, and Spectral Graph Theory. http://theory.stanford.edu/~trevisan/cs359g, 2015.