Electronic Journal of Probability

Electron. J. Probab. **22** (2017), no. 9, 1–29.
ISSN: 1083-6489 DOI: 10.1214/16-EJP11

Mixing time of the fifteen puzzle

Ben Morris * Anastasia Raymer †

Abstract

We show that the mixing time for the fifteen puzzle in an $n \times n$ torus is on the order of $n^4 \log n$.

Keywords: fifteen puzzle; mixing time; interchange process.
AMS MSC 2010: 60J10.
Submitted to EJP on September 1, 2016, final version accepted on October 17, 2016.

1 Introduction

The fifteen puzzle, often credited to Sam Loyd, was a craze in 1880. The game consists of a 4×4 grid with fifteen tiles, labeled 1,2,...,15, and an empty space (the "hole"). In a move, the player pushes a tile into the hole. The tiles start in "mixed up" order and the goal is to sort the tiles and move the hole to the lower right corner, as shown in Figure 1. There are also 3×3 and 2×4 versions of the game. In this paper we study the problem, posed by Diaconis [1], of finding the *mixing time* of the fifteen puzzle: starting from a solved game, how many steps are required to "mix up" the tiles again, if at each step we choose a move uniformly at random? (See Section 2 for a precise definition of the mixing time).

We can define the fifteen puzzle on any finite graph G as follows. In a configuration, the tiles and hole occupy the vertices of G. In a move, the hole is interchanged with a tile in an adjacent vertex. If G is bipartite, then there are some configurations that are not reachable from a given starting state. To see this, note that if G is bipartite then we can define a parity for each vertex in G. Thus, if we view configurations as permutations π on the vertex set of G, and define

$$\Omega = \{\pi : \mathrm{parity}(\pi) = \mathrm{parity}(\text{hole})\}; \tag{1.1}$$

$$\Omega^c = \{\pi : \mathrm{parity}(\pi) \neq \mathrm{parity}(\text{hole})\}; \tag{1.2}$$

then it is impossible to transition between Ω and Ω^c, using a legal move. Suppose that the game is started in a configuration in Ω. We say the game is *solvable* if every

*University of California, Davis. Research supported by NSF grant CNS-1228828. E-mail: morris@math.ucdavis.edu
†Cornell University. E-mail: araymer@math.cornell.edu

Fifteen puzzle

4	9	8	3
11	10	1	
14	2	7	6
15	13	12	5

Start

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	

End

Figure 1:

configuration in Ω is reachable by legal moves. If G is not bipartite, we say the game is solvable if every configuration is reachable by legal moves. The fifteen puzzle is known to be solvable on most graphs (see [15]); in particular, it is solvable on an $m \times n$ grid provided that m and n are both at least 2 (see [7]).

In the present paper, we analyze the fifteen puzzle in the $n \times n$ torus $G_n := \mathbf{Z}_n^2$. We consider the Markov chain, which we call the *Loyd process,* in which at each step a uniform random move is made. (We actually consider the *lazy* version of the Loyd process, where we add a holding probability of $\frac{1}{2}$ to each state to avoid periodicity.) The Loyd process is related to the *interchange process* on G_n, which is defined as follows. In a configuration, each vertex in G_n is occupied by a particle. At each step, choose a pair of neighboring particles uniformly at random and then interchange them. Yau [16] famously showed that the log Sobolev constant for the interchange process is on the order of n^{-4}, which implies that the mixing time is $\mathrm{O}(n^4 \log n)$, and there is a matching lower bound [11]. The Loyd process can be viewed as a variant of the interchange process, where there is a special particle (the hole) that is conditioned to be involved in each step.

Our main result is to determine the mixing time of the Loyd process to within constant factors. Let T_V be the mixing time in total variation for the Loyd process (see Section 2 for a precise definition). We show that there are universal constants $c > 0$ and $C > 0$ such that

$$cn^4 \log n \leq T_V \leq Cn^4 \log n.$$

For the upper bound, we use the comparison techniques for random walks on groups developed in [2], which allow us to bound the mixing time for the Loyd process using known bounds for shuffling by random transpositions. A difficulty that arises here is that G_n is bipartite when n is even, which implies that there is a restricted state space. To handle this, we use a method to compare mixing times across different state spaces. To compare our chain with shuffling by random transpositions, we introduce three intermediate chains and then make a total of four comparisons.

For the lower bound, we use a variation on Wilson's method [14]. (For a good introduction to Wilson's method, see section 13.2 of [9].) Wilson's method is useful when the Markov chain can be described as a system with a large number of particles where the motion of each individual particle is itself a Markov chain. (In the Loyd process the movement of a single tile is *not* a Markov chain; however, we can get around this by considering the process only at times when the hole is to its immediate right.) In Wilson's method, one often analyzes a distinguishing statistic of the form

$$\sum_p f(\text{position of particle } p),$$

where the sum is over a certain set of particles, and f is an eigenfunction for the motion of a single particle. In a typical application of Wilson's method, only a bounded number of particles are involved in each move, and hence the distinguishing statistic is slowly decaying. However, in the Loyd process, each move of the hole affects the distribution of the final position of each tile, which makes the "Wilson statistic" hard to analyze. Fortunately, by making use of some surprising cancellations we are able to prove a lower bound of the correct order $cn^4 \log n$. Note that our proof, although based on Wilson's method, is self-contained and does not rely on any of the results in [14].

2 Mixing time, harmonic extension, and continuous-time chain

Let $(X_0, X_1, \dots)$ be an irreducible, aperiodic Markov chain on a finite state space S with transition probabilities $p(x,y)$, and suppose that the stationary distribution π is uniform over S. For probability measures μ and ν on S, define the *total variation distance* $||\mu - \nu|| = \frac{1}{2}\sum_{x\in S} |\mu(x) - \nu(x)|$, and define the *$\epsilon$-mixing time in total variation*

$$T_V(\epsilon) = \min\{t : ||p^t(x,\,\cdot) - \pi|| \le \epsilon \text{ for all } x \in S\}\,. \tag{2.1}$$

The *mixing time in total variation* is $T_V = T_V(\frac{1}{4})$.

For a measure μ on S, write

$$\chi^2(\mu,\pi) := \sum_{x\in S} \pi(x)\left(\frac{\mu(x)}{\pi(x)} - 1\right)^2 = \left(\sum_{x\in S}\frac{\mu(x)^2}{\pi(x)}\right) - 1.$$

By Cauchy-Schwarz,

$$2\|\mu - \pi\| = \left\|\frac{\mu(\,\cdot\,)}{\pi(\,\cdot\,)} - 1\right\|_{L^1(\pi)} \le \left\|\frac{\mu(\,\cdot\,)}{\pi(\,\cdot\,)} - 1\right\|_{L^2(\pi)} = \chi(\mu,\pi)\,. \tag{2.2}$$

Define the *ϵ-mixing time* in L^2 by

$$T_{\text{mix}}(\epsilon) = \min\{t : \chi(p^t(x,\,\cdot),\pi) \le \epsilon \text{ for all } x \in S\}\,. \tag{2.3}$$

The mixing time in L^2 is $T_{\text{mix}} = T_{\text{mix}}(1/4)$. By (2.2), for any $\epsilon > 0$ we have

$$T_V(\epsilon) \le T_{\text{mix}}(2\epsilon). \tag{2.4}$$

For $S' \subset S$, let $\tau_1 < \tau_2 < \cdots$ be the times when the chain is in S'. The *restriction* of the Markov chain to S' is the new Markov chain $(X_{\tau_1}, X_{\tau_2}, \dots)$. For $f : S' \to \mathbf{R}$, the *harmonic extension of f to S* is the function $\tilde{f}$ that agrees with f on S' and is harmonic on $S \setminus S'$, which can be defined by

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \in S'; \\ \mathbb{E}_x(f(X_{T_{S'}})) & \text{otherwise,} \end{cases}$$

where $\mathbb{E}_x\Big(\,\cdot\,\Big) := \mathbb{E}\Big(\,\cdot\,\Big|\, X_0 = x\Big)$ and $T_{S'} = \min\{t \ge 0 : X_t \in S'\}$ is the hitting time of S'.

If P is the transition matrix of the chain, we consider the continuous-time Markov process that moves at rate 1 according to P. Thus, for $i \ne j$ the continuous-time chain moves from i to j at rate $p(i,j)$. Let H_t be the transition probabilities for the continuous-time chain, defined by

$$H_t(x,y) = \sum_{i=0}^{\infty} \frac{e^{-t}t^i}{i!} p^i(x,y).$$

For the continuous-time chain, the ϵ-*mixing time* in L^2 is defined by

$$\tau_{\mathrm{mix}}(\epsilon) = \min\{t \geq 0 : \chi(H_t(x,\cdot),\pi) \leq \epsilon \text{ for all } x \in S\}\,, \tag{2.5}$$

and the mixing time in L^2 is $\tau_{\mathrm{mix}} = \tau_{\mathrm{mix}}(1/4)$.

3 Random walks on groups and comparison techniques

Let G be a finite group and let p be a probability measure supported on a set of generators of G. The random walk on G driven by p is the Markov chain with the following transition rule. If the current state is x, choose y at random according to p, and then move to xy.

In the present paper we shall use a slightly more general definition of a random walk on a group. For a finite group G, we write G^* for the set of *strings* over G, that is, finite sequences of elements of G. If $g_1 g_2 \cdots g_k \in G^*$, we define its *evaluation* as the group element $g_1 \cdot g_2 \cdots g_k$ (where $\cdot$ is the group operation). As an abuse of notation, we use the string itself as notation for its evaluation. (Thus there exist strings y and y' such that $y \neq y'$ in G^*, but $y = y'$ in G.) If two strings evaluate to the same group element, we say that one is a *representation* of the other.

Let H be a subgroup of G, let p be a probability measure on G^*, and suppose that

$$\{g \in G : g \text{ is the evaluation of a string in the support of } p\}$$

is a generating set for H. The random walk on H driven by p is the Markov chain with the following transition rule. If the current state is $x \in H$:

1. choose the string y at random according to p;
2. move to xy.

We call strings in the support of p *moves*. For x and y in G^* we write xy for the concatenation of x and y. Note that since the random walk on H driven by p is doubly stochastic the stationary distribution is uniform over H.

3.1 Comparison techniques

We say that p is *symmetric* if $p(g_1 \cdots g_k) = p(g_k^{-1} \cdots g_1^{-1})$ for every $g_1 \cdots g_k \in G^*$. Let p and $\tilde{p}$ be symmetric probability measures on G^* that drive random walks on a subgroup H of G. Think of $\tilde{p}$ as driving a known chain and p as driving an unknown chain. Let E be the support of p. For each y in the support of $\tilde{p}$, we give a *random* representation of y of the form $Z_1 Z_2 \cdots Z_K$, where K is possibly random, and each of the Z_i are random elements of E. Given such a representation, we write $|y|$ for the value of K. For $z \in E$, let

$$N(z,y) \quad = \quad \text{number of times } z \in E \text{ occurs in the representation of } y.$$

Theorem 3.1. *([4]) If*

$$A = \max_{z \in E} \frac{1}{p(z)} \mathbb{E}\Big(\sum_{y \in G^*} |y| N(z,y) \tilde{p}(y)\Big)\,,$$

then for any $\epsilon > 0$ the ϵ-mixing times in L^2 for the continuous-time random walks satisfy

$$\tau_{\mathrm{mix}}(\epsilon) \leq A \tilde{\tau}_{\mathrm{mix}}(\epsilon).$$

Remark 3.2. Note that the quantity A can be written as

$$A = \max_{z \in E} \frac{1}{p(z)} \mathbb{E}\left(|Y| N(z, Y)\right),$$

where Y is chosen at random according to $\tilde{p}$.

Remark 3.3. In the original formulation in [4] the representations are deterministic. In this case the quantity A can be written as

$$A = \max_{z \in E} \frac{1}{p(z)} \sum_{y \in G^*} |y| N(z, y) \tilde{p}(y).$$

4 Mixing time upper bound: main theorem

Before stating the mixing time upper bound, we give a more formal description of the Loyd chain, and we also describe some other chains that are used in comparisons. Suppose $n \geq 2$ and let V_n be the vertex set of the $n \times n$ torus G_n. Note that if we give each tile and the hole a unique label in V_n, then we can view configurations as permutations on V_n. For reasons that will become clear later, we give the hole the label $h := (0, 0)$. For $y = (y_1, y_2) \in V_n$, call y *even* if $y_1 + y_2$ is even, and define Ω and Ω^c as in equations (1.1) and (1.2). Since the fifteen puzzle is solvable in a grid of size 2×2 or larger, any pair of states in Ω (respectively, Ω^c) communicate. Furthermore, there are transitions between Ω and Ω^c if and only if n is odd. It follows that the state space is restricted to half the permutations exactly when n is even. If we start from a configuration in Ω, then the state space is

$$\begin{cases} \Omega & \text{if } n \text{ is even;} \\ \text{all permutations on } V_n & \text{if } n \text{ is odd.} \end{cases}$$

As stated in the Introduction, we prove the upper bound by comparing the Loyd chain with shuffling by random transpositions, using a number of intermediate chains. For easy reference we give a short description of each of these chains below. We shall describe each chain in discrete time, but in fact our comparisons will involve the mixing times of the continuous-time versions of each chain.

1. Loyd chain: interchange the hole with one of four adjacent tiles, chosen uniformly at random.
2. Hole-conditioned chain (HC): Interchange the hole with a tile chosen uniformly at random.
3. Shuffling by random transpositions (RT): Choose two particles uniformly at random and then swap them. (Here *particle* refers to both the tiles and the hole.)

The following two chains are defined when n is even.

4. Parity-conditioned chain (PC): Choose a tile whose position has opposite parity to that of the hole, uniformly at random, and then interchange it with the hole.
5. Ω-restricted chain (OR): The hole-conditioned chain, restricted to Ω. That is, if $T_1 < T_2 < \cdots$ are the times when the hole-conditioned chain X_t is in Ω, then the Ω-restricted chain is $\{X_{T_j} : j \geq 1\}$.

Our main theorem concerns the *lazy version* of the Loyd process, which is obtained from the Loyd process by adding a holding probability of $\frac{1}{2}$. That is, at each step we do nothing with probability $\frac{1}{2}$; else move like in the Loyd process.

Theorem 4.1. *The mixing time in total variation for the lazy version of the Loyd process satisfies* $T_{\text{mix}} = \mathrm{O}(n^4 \log n)$.

Proof. By (2.4) and Theorem 20.3 of [9], it is sufficient to show that the bound holds for the L^2 mixing time of the continuous-time Loyd chain. Hereafter, we will simply write *mixing time* for the L^2 mixing time of a continuous-time chain.

The mixing time $\tau_{\mathrm{mix\,RT}}$ for shuffling n^2 cards by random transpositions in continuous time satisfies

$$\tau_{\mathrm{mix\,RT}} \leq an^2 \log n,$$

for a universal constant $a > 0$; see [3, 8, 13].

For the case when n is even, the theorem follows from the following relations between mixing times: For any $\epsilon > 0$, we have

$$\tau_{\mathrm{mix\,HC}}(\epsilon) \leq 12\tau_{\mathrm{mix\,RT}}(\epsilon); \quad \tau_{\mathrm{mix\,OR}}(2\epsilon) \leq 2\tau_{\mathrm{mix\,HC}}(\epsilon) + C_\epsilon n^2 \log n; \quad \tau_{\mathrm{mix\,PC}}(\epsilon) \leq 822\tau_{\mathrm{mix\,OR}}(\epsilon);$$

$$\tau_{\mathrm{mix\,Loyd}}(\epsilon) \leq \alpha n^2 \tau_{\mathrm{mix\,PC}}(\epsilon);$$

which we prove below as Lemmas 5.1, 6.2, 6.3 and 6.4, respectively. Here, $\tau_{\mathrm{mix\,PC}}$ denotes the mixing time of the parity-conditioned chain, with similar notation for the other chains, α is a universal constant, and C_ϵ is a constant that depends only on ϵ.

For the case when n is odd, Theorem 4.1 follows from the following relations between mixing times:

$$\tau_{\mathrm{mix\,HC}} \leq 12\tau_{\mathrm{mix\,RT}}; \qquad \tau_{\mathrm{mix\,Loyd}} \leq \alpha' n^2 \tau_{\mathrm{mix\,HC}};$$

which we prove below as Lemmas 5.1 and 6.5, respectively. Here, α' is a universal constant.

The proof of Lemma 5.1 can be found in Section 5. The proofs of Lemmas 6.2, 6.3, 6.4 and 6.5 can be found in Section 6. □

5 Comparison of hole-conditioned chain with random transpositions

Lemma 5.1. *The mixing times $\tau_{\mathrm{mix\,RT}}$ and $\tau_{\mathrm{mix\,HC}}$ satisfy*

$$\tau_{\mathrm{mix\,HC}} \leq 12\tau_{\mathrm{mix\,RT}}$$

Proof. Let G be the symmetric group on V_n with the group operation defined by

$$\pi\mu = \mu \circ \pi.$$

For permutations π on V_n, if we think of $\pi(j)$ as representing the label of the particle in position j, then we can view shuffling by random transpositions (respectively, the hole-conditioned chain) as the random walk on G driven by $\tilde{p}$ (respectively, p), where

$$\begin{aligned} \tilde{p} &= \text{uniform distribution on permutations of the form } (i,j) \text{ with } i \neq j \text{ and } i,j \in V_n; \\ p &= \text{uniform distribution on permutations of the form } (h,i) \text{ with } i \neq h \text{ and } i \in V_n. \end{aligned}$$

We compare the hole-conditioned chain with shuffling by random transpositions using Theorem 3.1. If $i < j$ we represent the permutation (i,j) by $(h,i)(h,j)(h,i)$. Consider the move (h,i) in the support of p. Note that (h,i) is in the representation of $n^2 - 1$ elements, each of the form (i,j). Since $p((h,i)) = 1/(n^2-1)$ and $\tilde{p}((i,j)) = 1/\binom{n^2}{2}$, applying Theorem 3.1 and using the bounds $N(z,y) \leq 2$ and $|y| \leq 3$ gives

$$A \leq 6(n^2-1)^2 / \binom{n^2}{2} < 12.$$

□

6 Comparisons involving the remaining chains

The subsequent chains that we analyze are random walks on a different group. Note that the hole-conditioned chain, Loyd chain, and parity-conditioned chain all can be described as follows. At each step:

1. choose y according to some distribution on V_n;
2. if the hole is in position x, interchange it with the tile in position $x + y$.

To see that these are random walks on a group, let $\widehat{V}_n = V_n \setminus (0,0)$ and note that a configuration can be specified by an ordered pair (x, f), where $x \in V_n$ is the position of the hole, and $f : \widehat{V}_n \to \widehat{V}_n$, is the permutation defined by

$$f(z) = (\text{position of tile } z) - x.$$

(Thus f gives the positions of the tiles relative to the hole; note that f maps tiles to positions, whereas for the permutations in Section 5 it was the other way around.)

Let G be the group whose elements are $\{(x, f) : x \in V_n, f \text{ is a permutation on } \widehat{V}_n\}$ and with the group operation

$$(x, f) \cdot (y, g) = (x + y, g \circ f).$$

Thus G is the direct product of V_n and the symmetric group on $\widehat{V}_n$. For $y \in V_n$, the transition that translates the hole by y is right multiplication by the group element (y, π_y), where π_y is the permutation defined by

$$\pi_y(z) = \begin{cases} z - y & \text{if } z \neq y; \\ -y & \text{if } z = y. \end{cases} \tag{6.1}$$

As an abuse of notation, we write y for the move (y, π_y). We write $\uparrow, \downarrow, \rightarrow$, and $\leftarrow$ for the moves $(0, 1), (0, -1), (1, 0)$, and $(-1, 0)$, respectively.

The Ω-restricted chain. We write 0 for the identity element $(0, 0)$, of G. If n is even, and we define

$$\Omega = \{(x, f) : \text{parity}(x) = \text{parity}(f)\},$$

then Ω is the set of states reachable from 0 in the Loyd chain. Note that Ω is closed under products and inverses and hence is a subgroup of G.

It is not hard to show that the permutation π_y defined in (6.1) is odd unless $y = 0$. This implies that the move y is in Ω if and only if y is odd or 0. We will call such moves *good* and the other moves *bad.* Note that the product of moves $y_1 y_2 \cdots y_m$ is in Ω if and only if an even number of the y_i are bad.

The Ω-restricted chain is a random walk on Ω where each move is generated as follows:

1. Let $y_1, y_2, \ldots$ be i.i.d. moves of the hole-conditioned chain, and let

$$T = \min\{m \geq 1 : \text{an even number of the moves } y_1, \ldots, y_m \text{ are bad}\};$$

2. Let the move be $y_1 y_2 \cdots y_T$.

6.1 Comparison of hole-conditioned chain with Ω-restricted chain

Note that the Ω-restricted chain is a "sped up" version of the hole-conditioned chain; this suggests that its mixing time is at most of the same order. In this section we show that this is indeed the case. The key step in the proof is to show that the time the hole-conditioned chain spends in Ω over the interval $[0,t]$ is tightly concentrated around $t/2$.

More precisely, let $\widetilde{X}_t$ be the continuous-time version of the hole-conditioned chain and define the occupation measure $\mathcal{L}_t(\Omega) = \int_0^t \mathbf{1}(\widetilde{X}_s \in \Omega)\, ds$. For $t \geq 0$, define $T_t = \inf\{s : \mathcal{L}_s(\Omega) > t\}$ and define $X_t = \widetilde{X}_{T_t}$. Note that X_t is the continuous-time version of the Ω-restricted chain.

When t is large the random variable T_t is typically about $2t$. In the following lemma we give some large-deviation bounds. Note that there are $\frac{n^2}{2} - 1$ bad vertices, and $\frac{n^2}{2}$ good vertices, respectively, in $\widehat{V}_n$.

Lemma 6.1. *Let $\lambda = \frac{n^2-2}{2(n^2-1)}$ be the proportion of vertices in $\widehat{V}_n$ that are bad. For any $a > 0$ there exist positive constants α_a and β_a such that*

$$\mathbb{E}\Big((T_t - (2+a)t)^+\Big) \leq \frac{\alpha_a}{\lambda} e^{-\beta_a \lambda t}, \tag{6.2}$$

and

$$\mathbb{E}\Big((T_t - (2-a)t)^-\Big) \leq \frac{\alpha_a}{\lambda} e^{-\beta_a \lambda t}. \tag{6.3}$$

Proof. We shall prove that the Lemma holds with

$$\alpha_a = \frac{6\sqrt{2}(2+a)}{a} \qquad \beta_a = \frac{a^2}{6(2+a)}.$$

Note that $\mathbf{1}(\widetilde{X}_t \in \Omega)$ is a continuous-time Markov chain on $\{0,1\}$ with generator

$$G = \begin{bmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{bmatrix},$$

The second eigenvalue of G is -2λ. Thus, Theorem 3.4 of [10] implies that for any $h > t/2$ we have

$$\mathbb{P}(\mathcal{L}_t(\Omega) > h) \quad \leq \quad \sqrt{2}\exp(-\lambda(2h-t)^2/6t),$$

and for any $h < t/2$ we have

$$\mathbb{P}(\mathcal{L}_t(\Omega) < h) \quad \leq \quad \sqrt{2}\exp(-\lambda(2h-t)^2/6t),$$

Hence, for any $s > 0$ we have

$$\begin{aligned} \mathbb{P}(T_t > (2+s)t) &= \mathbb{P}(\mathcal{L}_{(2+s)t} < t) \\ &\leq \sqrt{2}\exp\Big(-s^2\lambda t/6(2+s)\Big), \end{aligned} \tag{6.4}$$

and

$$\begin{aligned} \mathbb{P}(T_t < (2-s)t) &= \mathbb{P}(\mathcal{L}_{(2-s)t} > t) \\ &\leq \sqrt{2}\exp\Big(-s^2\lambda t/6(2-s)\Big). \end{aligned} \tag{6.5}$$

First we show that (6.2) follows from (6.4). Using the identity $\mathbb{E}(X) = t\int_0^\infty \mathbb{P}(X > tu)\, du$, valid for any $t > 0$ and non-negative random variable X, gives

$$\begin{aligned} \mathbb{E}\Big((T_t - (2+a)t)^+\Big) &= t\int_0^\infty \mathbb{P}(T_t - (2+a)t > tu)\, du \\ &= t\int_a^\infty \mathbb{P}(T_t > (2+s)t)\, ds \qquad (6.6) \\ &\leq \sqrt{2}t\int_a^\infty \exp\Big(-s^2\lambda t/6(2+s)\Big)\, ds, \qquad (6.7) \end{aligned}$$

where in the second line we have made the change of variable $s = u + a$, and the inequality holds by (6.4). Since $\frac{s}{2+s} \geq \frac{a}{2+a}$ whenever $s \geq a$, the quantity (6.7) is at most

$$\sqrt{2}t \int_a^\infty \exp\Big(-as\lambda t/6(2+a)\Big)\,ds \quad = \quad \frac{6\sqrt{2}(2+a)}{a\lambda} \exp\Big(-a^2\lambda t/6(2+a)\Big), \tag{6.8}$$

verifying (6.2). One can similarly show that (6.3) follows from (6.5). (Note that the bound in (6.5) is stronger than the bound in (6.4), but the right-hand sides of (6.2) and (6.3) are the same.) This completes the proof. □

Lemma 6.2. *Suppose that n is even. For any $\epsilon > 0$ there is a $C_\epsilon > 0$ such that the mixing times of the Ω-restricted and hole-conditioned chain satisfy*

$$\tau_{\mathrm{mixOR}}(2\epsilon) \leq \max\Big(2\tau_{\mathrm{mixHC}}(\epsilon), C_\epsilon n^2 \log n\Big);$$

Proof. Let p and $\tilde{p}$ be transition probabilities for the Ω-restricted and hole-conditioned chains, respectively, and let π and $\tilde{\pi}$ be the corresponding stationary distributions. For $t \geq 0$, define $d_2(t) = \chi(p^t(x,\,\cdot), \pi)$, with a similar definition for $\tilde{d}_2(t)$. Let x be a state in Ω. Since p is symmetric, conditioning on X_t gives that $p^{2t}(x,x) = \sum_{y\in\Omega} p(x,y)^2$, and hence

$$d_2^2(t) = Np^{2t}(x,x) - 1, \tag{6.9}$$

where $N = |\Omega|$. Similary,

$$\tilde{d}_2^2(t) = (2N)\tilde{p}^{2t}(x,x) - 1. \tag{6.10}$$

(The factor 2 is present because the state space for the hole-conditioned chain is twice as large as that of the Ω-restricted chain.) Note that replacing t by $t/2$ in (6.10) gives

$$\tilde{p}^t(x,x) = \frac{1}{2N}(1 + \tilde{d}_2^2(t/2)). \tag{6.11}$$

Since $p^t(x,x)$ is non-increasing in t (see [4]), we have

$$p^{2t}(x,x) \leq \frac{1}{t}\int_t^{2t} p^u(x,x)\,du. \tag{6.12}$$

Fix $\epsilon > 0$ and suppose that a satisfies $0 \leq a \leq \min(\frac{1}{8}, \frac{\epsilon}{8})$, so that

$$2a + 2a\epsilon \leq \frac{\epsilon}{2}. \tag{6.13}$$

The integral in (6.12) can be interpreted as the expected amount of time that the Ω-restricted chain spends at x, between time t and $2t$. This is equal to the expected amount of time that the Ω-restricted chain spends at x between T_t and T_{2t}, which is at most

$$\int_{2t(1-a)}^{4t(1+a)} \tilde{p}^u(x,x)\,du + \mathbb{E}\Big((T_t - 2t(1-a))^-\Big) + \mathbb{E}\Big((T_{2t} - 4t(1+a))^+\Big) \tag{6.14}$$

$$\leq \quad 2t(1+2a)\tilde{p}^t(x,x) + B_\epsilon e^{-D_\epsilon t}, \tag{6.15}$$

where B_ϵ and D_ϵ are constants that depend only on ϵ. The first term in (6.15) bounds the integral in (6.14) because the integrand is bounded above by $p^t(x,x)$ (since $a \leq \frac{1}{2}$), and the second term in (6.15) bounds the expectations in (6.14) by Lemma 6.3. (Note that the proportion λ of vertices in $\widehat{V}_n$ that are bad can be bounded below by $\frac{1}{3}$ for $n \geq 2$.)

Now, note that whenever $t \geq 2\tau_{\mathrm{mixHC}}(\epsilon)$, we have $\tilde{d}_2^2(t/2)) \leq \epsilon$, and hence

$$\tilde{p}^t(x,x) \leq \frac{1}{2N}(1+\epsilon)$$

by (6.11). It follows that the quantity (6.15) is at most

$$t(1+2a)\frac{1}{N}(1+\epsilon)) + B_\epsilon e^{-D_\epsilon t}.$$

Since this is an upper bound on the integral in (6.12), we have

$$\begin{aligned} p^{2t}(x,x) &\leq (1+2a)\frac{1}{N}(1+\epsilon) + \frac{1}{t}B_\epsilon e^{-D_\epsilon t} \\ &= \left[1+2a+2a\epsilon+\epsilon+\frac{N}{t}B_\epsilon e^{-D_\epsilon t}\right]\frac{1}{N}, \end{aligned}$$

and hence

$$d(t) \leq 2a + 2a\epsilon + \epsilon + \frac{N}{t}B_\epsilon e^{-D_\epsilon t}$$

by (6.9). Combining this with (6.13) gives

$$d(t) \leq \tfrac{3}{2}\epsilon + \frac{N}{t}B_\epsilon e^{-D_\epsilon t}.$$

Finally, since $N = \frac{1}{2}(n^2)! < \exp(2n^2 \log n)$, there is a constant $C_\epsilon > 0$ such that

$$\frac{N}{t}B_\epsilon e^{-D_\epsilon t} \leq \frac{\epsilon}{2}, \tag{6.16}$$

whenever $t \geq C_\epsilon n^2 \log n$. It follows that

$$d(t) \leq 2\epsilon$$

whenever $t \geq 2\tau_{\mathrm{mixHC}}(\epsilon)$ and $t \geq C_\epsilon n^2 \log n$, proving the Lemma. □

6.2 Comparison of parity-conditioned chain to Ω-restricted chain

Lemma 6.3. *Suppose that n is even. Then the mixing times τ_{mixPC} and τ_{mixOR} satisfy*

$$\tau_{\mathrm{mixPC}} \leq (882)\tau_{\mathrm{mixOR}}.$$

Proof. In order to compare the Ω-restricted chain with the parity-conditioned chain we intoduce an intermediate chain, which we denote BGB . A move of the BGB chain is a concatenation consisting of between 1 and 3 moves of the HC chain, generated as follows. Let b_1 and b_2 be uniform random bad moves, and let g be a uniform random good move, respectively, in $\widehat{V}_n$. The BGB move is

$$x = \begin{cases} g & \text{with probability } 1/3; \\ b_1 b_2 & \text{with probability } 1/3; \\ b_1 g b_2 & \text{with probability } 1/3. \end{cases}$$

We shall use Theorem 3.1 twice, first to compare BGB with the Ω-restricted chain, then to compare PC with BGB.

Comparison of BGB with the Ω-restricted chain chain. We need to show how to represent moves of the Ω-restricted chain using BGB moves. Consider a move y of the Ω-restricted chain. Then y is of the form g, $b_1 b_2$ or $b_1 g_1 g_2 \cdots g_k b_2$, where we write b's for bad moves and g's for good moves. If $y = g$ (respectively, $y = b_1 b_2$) then we can represent it as g (respectively, $b_1 b_2$), since this is also a BGB move. Suppose now that

$$y = b_1 g_1 g_2 \cdots g_k b_2.$$

In this case we represent it as $z_1 \cdots z_k$, where the z_i are defined by

$$\underbrace{(b_1 g_1 B_1)}_{z_1}\underbrace{(B_1^{-1} g_2 B_2)}_{z_2}\underbrace{(B_2^{-1} g_3 B_3)}_{z_3}\cdots\underbrace{(B_{k-1}^{-1} g_k b_2)}_{z_k},$$

for uniform random bad moves $B_1, \ldots, B_{k-1}$.

We apply Theorem 3.1, letting

$$\begin{aligned}\tilde{p} &= \text{measure corresponding to the } \Omega\text{-restricted chain;}\\ p &= \text{measure corresponding to the BGB chain.}\end{aligned}$$

We need to bound the quantity $A = \max_z A(z)$, where

$$A(z) = \frac{1}{p(z)}\mathbb{E}\Big(\sum_y \tilde{p}(y) N(z,y)|y|\Big).$$

Recall that there are $\frac{n^2}{2} - 1$ bad vertices, and $\frac{n^2}{2}$ good vertices, respectively, in $\widehat{V}_n$. If $z = g$ then z is used only in the representation of g itself, and since

$$\tilde{p}(g) = \frac{1}{n^2 - 1}, \qquad p(g) = \frac{2}{3n^2};$$

we have $A(z) = \frac{\tilde{p}(g)}{p(g)} = \frac{3n^2}{2(n^2-1)}$. Similarly, if $z = b_1 b_2$ then since

$$\tilde{p}(b_1 b_2) = \frac{1}{(n^2-1)^2}; \qquad p(b_1 b_2) = \frac{1}{3(\frac{n^2}{2} - 1)^2},$$

then $A(z) = \frac{\tilde{p}(b_1 b_2)}{p(b_1 b_2)} = \frac{3(n^2-2)^2}{4(n^2-1)^2}$.

It remains to check the case when z is of the form $b_1 g b_2$. Note that $A(z)$ can be written as $\frac{1}{p(z)}\mathbb{E}\left(N(z,Y)|Y|\right)$, where Y is a random move chosen from $\tilde{p}$. Define the random variable K by

$$K = \begin{cases} \text{k} & \text{if } Y = b_1 g_1 g_2 \cdots g_k b_2;\\ 0 & \text{if } Y \text{ is of the form } g \text{ or } b_1 b_2.\end{cases}$$

Note that $\mathbb{P}(K = k) = \left(\frac{1}{2}\right)^{k+2}$ for $k \geq 1$. Furthermore, conditional on $K = k$, the distributions of $Z_1, \ldots, Z_k$ are uniform over moves of the form $b_1 g b_2$. It follows that

$$\begin{aligned}\mathbb{E}\Big(N(z,Y)|Y|\,\Big|\, K = k\Big) &= k\mathbb{E}\Big(N(z,Y)\,\Big|\, K = k\Big)\\ &= k\,\frac{k}{|S|},\end{aligned}$$

where S is the set of moves of the form $b_1 g b_2$. It follows that

$$\begin{aligned}\mathbb{E}\left(N(z,Y)|Y|\right) &= \sum_{k\geq 1}\left(\tfrac{1}{2}\right)^{k+2}\frac{k^2}{|S|}\\ &= \frac{3}{2|S|}.\end{aligned}$$

Since $p(z) = \frac{1}{3|S|}$, we have $A(z) = \frac{1}{p(z)}\mathbb{E}\left(N(z,Y)|Y|\right) = 9/2$. Hence $A = 9/2$ as well, and hence

$$\tau_{\mathrm{mixBGB}}(\epsilon) \leq 822\tau_{\mathrm{mixOR}}(\epsilon). \tag{6.17}$$

Comparison of PC chain with BGB chain. We need to show how to represent a BGB move with PC moves. Consider a move y of the BGB chain. If $y = g$ then we represent it as g itself. To handle moves of the form $b_1 b_2$ and $b_1 g b_2$, we first note that if $e_1, e_2 \in \widehat{V}_n$ are even and $o \in \widehat{V}_n$ is odd, then we can represent the BGB move $e_1 o e_2$ as

$$(e_1 + o)(-o)(o + e_2)(-e_1 - o - e_2)(e_1 + o)(-o)(o + e_2). \tag{6.18}$$

Note that the moves in (6.18) are moves of the PC chain, since the corresponding elements of V_n are odd. If y is of the form $b_1 g b_2$, we can represent it with PC moves using (6.18). If y is of the form $b_1 b_2$, we first give it the intermediate representation $(b_1 G B)(B G b_2)$, where B and G are uniform random bad and good moves, respectively, and then represent both the $b_1 G B$ and $B G b_2$ using (6.18). Note that the maximum length of the representation of any y is 14

We apply Theorem 3.1 again, this time letting $\tilde{p}$ (respectively, p) be the measure corresponding to the BGB chain (respectively, PC chain). We need to bound the quantity

$$A = \max_z \frac{1}{p(z)} \mathbb{E}\left(N(z, Y)|Y|\right),$$

where Y is chosen according to $\tilde{p}$. Let $Y = Z_1 \cdots Z_K$ be the representation of Y. Note that for all $k \in \{1, 7, 14\}$ the conditional distribution of $Z_1, \ldots, Z_k$, given $|Y| = k$ is uniform over the set of PC moves. It follows that for every PC move z we have

$$\begin{aligned} \mathbb{E}\Big(N(z, Y)|Y| \,\Big|\, |Y| = k\Big) &= k\mathbb{E}\Big(N(z, Y) \,\Big|\, K = k\Big) \\ &= k\,\frac{k}{|PC|} \end{aligned}$$

where we write $|PC|$ for the number of PC moves. It follows that, for any PC move z, we have

$$\begin{aligned} \mathbb{E}\left(N(z, Y)|Y|\right) &= \frac{1}{|PC|}\mathbb{E}\left(|Y|^2\right) \\ &\leq \frac{196}{|PC|}, \end{aligned}$$

where the last line holds because $|Y| \leq 14$. Since p is the uniform distribution over PC moves, we have $p(z) = \frac{1}{|PC|}$, and hence $\frac{1}{p(z)}\mathbb{E}\left(N(z, Y)|Y|\right) \leq 196$. Hence $A \leq 196$, which implies that

$$\tau_{\mathrm{mix\,PC}}(\epsilon) \leq 822\tau_{\mathrm{mix\,BGB}}(\epsilon). \tag{6.19}$$

Combining this with (6.17) yields the lemma. □

6.3 Comparisons of parity-conditioned and hole-conditioned chains with Loyd chain

Lemma 6.4. *Suppose that n is even. Then the mixing times $\tau_{\mathrm{mix\,Loyd}}(\epsilon)$ and $\tau_{\mathrm{mix\,PC}}(\epsilon)$ satisfy*

$$\tau_{\mathrm{mix\,Loyd}}(\epsilon) \leq \alpha n^2 \tau_{\mathrm{mix\,PC}}(\epsilon),$$

for a universal constant α.

Proof. In order to apply Theorem 3.1, we need to show how to represent any move of the PC chain using moves of the Loyd chain. We will actually show how to represent PC moves using a different Markov chain, which we call *near Loyd* (NL). In the NL chain, each move is a move y of the PC chain with y conditioned to satisfy $|y_1| + |y_2| \in \{1, 3\}$, where for $u \in \mathbf{Z}_n$ we define $|u| = \min(u, n - u)$. That is, each step of the NL chain swaps the hole with tile at L^1-distance 1 or 3 away from it. A representation using NL moves is sufficient because any NL move can be represented using a bounded number of Loyd moves: if the L^1-distance between the hole and tile T is at most 3, then there is a 3×3 square grid that contains both the hole and tile T, and the fifteen puzzle is solvable in a 3×3 grid.

We now show how to represent a PC move with NL moves. There are three cases to consider.

Case 1: swapping the hole with a tile one row higher. We first consider the case where the move $y = (y_1, y_2)$ is such that $y_2 = 1$. That is, the move swaps the hole with a tile one row higher.

Suppose that tile T is located in the row immediately above the hole. To swap the hole with tile T, leaving everything else the same, perform the following algorithm:

1. repeat: $\uparrow, \rightarrow, \downarrow, \rightarrow$, until the hole is swapped with T.
2. do $\leftarrow, \downarrow, \rightarrow$ once.
3. repeat: $\uparrow, \leftarrow, \leftarrow, \downarrow, \rightarrow$, until T is in the position that the hole initially occupied.
4. repeat: alternate $\rightarrow, \uparrow$ and $\rightarrow, \downarrow$ until the hole is in the position initially occupied by T.

Note that each move here is actually a move of the Loyd chain.

Figures 2–6 show an application of the algorithm. In this example, the hole is swapped with the tile of label 9.

1	2	5	6	9
h	3	4	7	8

Figure 2: Initial configuration

2	3	6	7	h
1	4	5	8	9

Figure 3: After step 1

2	3	6	8	7
1	4	5	9	h

Figure 4: After step 2

1	2	4	5	8
9	h	3	6	7

Figure 5: After step 3

1	2	5	6	h
9	3	4	7	8

Figure 6: After step 4 (final position)

To see why these moves swap the hole and tile T, leaving everything else unchanged, suppose that we label the tiles so that the numbers increase in the pattern $\cdots$ between the hole and tile T. If k is the label of tile T, then for i with $1 \leq i \leq k$, we write *position* i for the initial position of the tile of label i, and *position* 0 for the initial position of the hole. We can represent each intermediate configuration of tiles as a $k+1$ dimensional vector, writing $(u_0, u_1, \dots, u_k)$ for the configuration where tile u_i is located in position i for $0 \leq i \leq k$. Note that under this convention the initial configuration is $(H, 1, 2, \dots, k)$.

Analysis of the Algorithm: Step 1 moves tile i to position $i-1$ for $1 \leq i \leq k$, and moves the hole to position k. Hence after step 1 the configuration is $(1, 2, \dots, k, H)$. (This is shown in Figure 3 in the case where $k = 9$.) Step 2 moves only the rightmost four tiles, moving tiles $k-1, k-2, k-3$ and the hole to positions $k-2, k-3, k$ and $k-1$, respectively. The resulting configuration is $(1, 2, \dots, k-1, k, H, k-2)$ (see Figure 4). In step 3, tiles $k-1$ and $k-2$ (initially the two rightmost tiles on the top row) each move only once and go to positions k and $k-1$, respectively. Tile 2 (initially the leftmost tile on the top row) moves only once and goes to position 2. Every other tile on the top row is moved twice, first to the right then down, increasing its position by 2 units. Tile 1 (initially the leftmost tile of the bottom row) moves only once and goes to position 1. Every other tile on the bottom row is moved twice, first up and then to the right, increasing its position by 2 units. The resulting configuration is $(k, 1, 2, H, 3, 4, \dots, k-2, k-1)$ (see Figure 5). Finally, step 4 moves the hole to position k, and tiles $3, 4, \dots, k-1$ down by one position each, resulting in configuration $(k, 1, 2, \dots, k-1, H)$ (see Figure 6).

Case 2: swapping the hole with a tile on the same row. Let $\mathcal{C}$ be a configuration in which the hole and tile T are on the same row. To swap the hole with tile T: Choose a tile T' on the row one step higher such that T and T' share one edge. Let $\mathcal{C}'$ be the configuration obtained from $\mathcal{C}$ by interchanging T and T'. Let f be the permutation on V_n that transposes the positions of tiles T and T' in configuration $\mathcal{C}$. Since in $\mathcal{C}'$ tile T is one row higher than the hole, we can use the algorithm for Case 1 to swap the hole and T starting from configuration $\mathcal{C}'$. Let l_k be the label of the tile swapped with the hole in the kth step when performing this algorithm. To swap the hole with tile T starting from configuration $\mathcal{C}$, we use the sequence of moves defined by the same label sequence $(l_1, l_2, \dots)$. Note that if a tile is in position x after k steps of the algorithm starting from $\mathcal{C}'$, then it is in position $f(x)$ after k steps of the algorithm starting from $\mathcal{C}$. Since the algoithm for Case 1 performs only Loyd moves, the resulting algorithm for $\mathcal{C}$ swaps the hole with tiles at a distance either 1 or 3 from it, that is, it performs only NL moves.

Case 3: swapping the hole with a tile not on the same row or next row up. Now we consider the situation not covered in Case 1 or Case 2. The cases where tile T is in the column to the immediate right of the hole or in the same column as the hole are similar to above, so assume neither of these situations hold, as in Figure 7. Let $\mathcal{C}$ be the configuration shown in Figure 7 and let $\mathcal{C}'$ be the configuration shown in Figure 8. Let f be the bijection from locations in $\mathcal{C}$ to locations in $\mathcal{C}'$ that leaves the horizontal part unchanged and rotates and inverts the vertical part (which consists of locations in the column of T and in the column one unit to the left of T) so that the location of tile T is sent to the row second from the bottom. Since in $\mathcal{C}'$ tile T is in the row second from the bottom, we can use the algorithm for Case 1 to swap the hole with tile T, using only Loyd moves, starting from configuration $\mathcal{C}'$. As before, we can use the labels of the tiles moved at each step to define an algorithm starting from configuration $\mathcal{C}$. Note that if positions x and y are adjacent in $\mathcal{C}'$ then $f^{-1}(x)$ and $f^{-1}(y)$ are at distance 1 or 3 from each other in $\mathcal{C}$. It follows that the algorithm for configuration $\mathcal{C}$ swaps the hole with tiles at a distance 1 or 3 from it, that is, performs only NL moves.

Note that the maximum length of the representation of a PC move using NL moves is at most Bn, for a universal constant B. This also applies to the resulting representation using Loyd moves.

We apply Theorem 3.1 again, this time letting $\tilde{p}$ (respectively, p) be the measure corresponding to the PC chain (respectively, Loyd chain). We need to bound the quantity

$$A = \max_z \frac{1}{p(z)} \mathbb{E}\left(N(z,Y)|Y|\right),$$

where Y is chosen according to $\tilde{p}$. Recall that $N(z,Y)$ denotes the number of times z occurs in the representation of Y. This is at most $|Y|$, since $|Y|$ is the total number of moves in the representation of Y. Hence $N(z,Y) \le |Y| \le Bn$, and since $p(z) = 1/4$ for $z \in \{\leftarrow, \rightarrow, \uparrow, \downarrow\}$, we have

$$A \le \alpha n^2$$

for a universal constant α, and hence

$$\tau_{\mathrm{mix}\,\mathrm{Loyd}} \le \alpha n^2 \tau_{\mathrm{mix}\,\mathrm{PC}}. \tag{6.20}$$

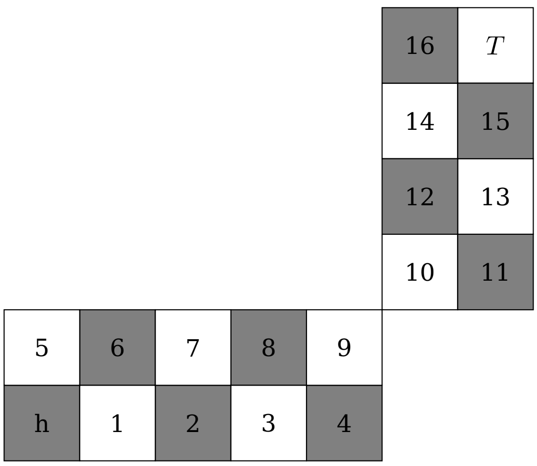

Figure 7:

5	6	7	8	9	11	13	15	T
h	1	2	3	4	10	12	14	16

Figure 8:

□

Lemma 6.5. *Suppose that n is odd. Then the mixing times $\tau_{\text{mixLoyd}}(\epsilon)$ and $\tau_{\text{mixHC}}(\epsilon)$ satisfy*

$$\tau_{\text{mixLoyd}}(\epsilon) \le \alpha' n^2 \tau_{\text{mixHC}}(\epsilon),$$

for a universal constant α'.

Proof. The proof follows the proof of Lemma 6.4 closely. We will show how to represent any HC move using Loyd moves. Consider a move $y = (y_1, y_2)$ of the HC chain. If y is odd then it is also a PC move and hence we can represent it using Loyd moves using the algorithm from the proof of Lemma 6.4. If y is even, then $(-y_1, y_2)$ is odd, and we can represent y using Loyd moves as follows: we perform the algorithm from the proof of Lemma 6.4 to swap the hole with the tile in position $(-y_1, y_2)$, but we interchange the roles of $\leftarrow$ and $\rightarrow$ moves. The resulting algorithm will swap the hole with the tile in position $(y_1, y_2) = y$.

We have shown that any HC move can be represented by $\mathrm{O}(n)$ Loyd moves, so the theorem follows by calculations similar to those leading up to equation (6.20). □

7 Lower bound

In this section we prove a lower bound on the order of $n^4 \log n$ for the mixing time of the Loyd chain. For the lower bound, a key fact is that if we look at a tile at times when the hole is immediately to its right, the x-coordinate is doing a random walk on $\mathbf{Z}_n$. More precisely, let $\{L_t : t = 0, 1, 2, \ldots\}$ be a (discrete time) lazy Loyd process. We write $L_t(u)$ for the position of tile u at time t. For a configuration L and tile u let $X(L, u)$ denote the x-coordinate of tile u in configuration L, and define $X_t(u) := X(L_t, u)$. Define $\tau_1(u), \tau_2(u), \ldots$ inductively as follows. Let $\tau_1(u)$ be the first time t such that the hole is immediately to the right of tile u at time t, and for $k > 1$, let $\tau_k(u)$ be the first time $t > \tau_{k-1}(u)$ such that the hole is immediately to the right of tile u at time t. The process $\{X_{\tau_k(u)}(u) : k \ge 0\}$ is a symmetric random walk on $\mathbf{Z}_n$, which we shall call the u *random walk*. To see this, note that if $m_1 m_2 \cdots m_l$ is a sequence of moves between times $\tau_1(u)$ and $\tau_2(u)$ that changes $X_t(u)$ from x to $x + 1 \pmod n$, then the sequence of moves $m_l^{-1}, m_{l-1}^{-1}, \ldots, m_1^{-1}$, which occurs with the same probability, would change x to $x - 1 \pmod n$ over the same time interval. Note that each step of the u random walk has a positive holding probability, which is the probability that between times $\tau_k(u)$ and $\tau_{k+1}(u)$ the value of $X_t(u)$ does not change.

Recall that for simple symmetric random walk on a cycle of length n, $f(x) = \cos \frac{2\pi x}{n}$ is an eigenfunction with corresponding eigenvalue $\cos \frac{2\pi}{n}$. Thus f is an eigenfunction for the u random walk as well. Since the u random walk has a holding probability the corresponding eigenvalue $\lambda > \cos \frac{2\pi}{n}$.

The rough idea behind the lower bound will be to show that the tiles that start with an x-coordinate close to 0 will tend to stay that way if the number of random walk steps is too low. Let $\mathcal{S}$ be the set of tiles u such that $f(X_0(u)) > 1/2$. and suppose that the hole is not initially adjacent to any tile in $\mathcal{S}$. Let μ be large enough so that

$$\left(\cos\frac{2\pi}{n}\right)^{n^2} \geq e^{-\mu}, \tag{7.1}$$

for all $n \geq 2$. (Such a μ exists because $\cos x$ has the power series expansion $1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots$.) Next, define

$$\epsilon = \tfrac{1}{8}\mu^{-1}; \qquad \widehat{T} = \lfloor 1 + \epsilon n^2 \log n \rfloor; \qquad T = (n^2-1)\widehat{T}. \tag{7.2}$$

Since there are $n^2 - 1$ tiles, we can think of the quantity $\widehat{T}$ as the typical number of times that the hole has been to the immediate right of any given tile if the Loyd process has made T steps.

We shall bound the mixing time from below by T, which is on the order of $n^4 \log n$. We accomplish this using as a *distinguishing statistic* the random variable W_{dist} defined by

$$W_{\text{dist}} = \sum_{u \in \mathcal{S}} f(X_T(u)).$$

Let $k = |\mathcal{S}|$ and let W be the sum of k samples without replacement from a population consisting of values of $\cos\frac{2\pi x}{n}$ for vertices $(x, y) \in V_n$. The lower bound follows from Lemmas A and B below, which together imply that $\|W_{\text{dist}} - W\|_{TV} \to 1$ as $n \to \infty$. In the statements of Lemmas A and B, the random variables depend implicitly on the parameter n of the Loyd process.

Lemma A. There is a universal constant $c > 0$ such that

$$\mathbb{P}(W_{\text{dist}} > cn^{15/8}) \to 1,$$

as $n \to \infty$.

Lemma B. For any $c > 0$ we have

$$\mathbb{P}(W > cn^{15/8}) \to 0,$$

as $n \to \infty$.

Theorem 7.1. *Let L_t be the Loyd process on G_n, and let π be the stationary distribution. There is a universal constant $c > 0$ such that for any $\epsilon > 0$, when n is sufficiently large, we have*

$$T_{\text{mix}}(\epsilon) > cn^4 \log n.$$

Proof. Lemmas A and B together imply that $\|W_{\text{dist}} - W\|_{TV} \to 1$ as $n \to \infty$. This implies the Theorem since W_{dist} is measurable with respect to L_T and $T \geq cn^4 \log n$ for a universal constant $c > 0$. $\square$

We prove Lemma A in subsection 7.1. Lemma B is a straightforward consequence of Hoeffding's bounds for sampling without replacement in [6], which we recall now.

Theorem 7.2. *Let $X_1, \ldots, X_k$ be samples, without replacement, from a population whose values are in the interval $[a, b]$, and suppose that the population mean $\mathbb{E}(X_1) = 0$.*

Then for $\alpha > 0$,

$$\mathbb{P}\left(\sum_{i=1}^{k} X_i \geq \alpha\right) \leq e^{-2\alpha^2/k(b-a)^2}. \tag{7.3}$$

Proof of Lemma B. Let $k = |\mathcal{S}|$. Applying Theorem 7.2 to k samples from a population consisting of values of $\cos\frac{2\pi x}{n}$ for vertices $(x, y) \in V_n$ gives

$$\mathbb{P}\left(\sum_{i=1}^{k} X_i \geq n^{15/8}\right) \leq \exp\left(-n^{15/4}/2k\right). \tag{7.4}$$

Since $k \leq n^2$, the quantity (7.4) converges to 0 as $n \to \infty$. $\square$

7.1 Proof of Lemma A

For $u \in \mathcal{S}$, let $N_t(u)$ be the number of times that the hole has been to the immediate right of tile u, up to time t. Note that for all t, if $N_t(u) > 0$ then

$$\tau_{N_t(u)}(u) \le t < \tau_{N_t(u)+1}(u).$$

Recall that $f(x) = \cos \frac{2\pi x}{n}$. It follows that $f'(x) = -\frac{2\pi}{n} \sin \frac{2\pi x}{n}$ and hence $|f'(x)| \le \frac{2\pi}{n}$ for all x. Thus the mean value theorem implies that for every x and k we have

$$|f(x+k) - f(x)| \le \frac{2\pi|k|}{n}. \tag{7.5}$$

We will prove Lemma A by approximating W_{dist} by the random variable $Z := \sum_{u \in \mathcal{S}} X_{\tau_{\hat{T}}(u)}(u)$. The random variable Z is easier to analyze than W_{dist} (but couldn't be used as a distinguishing statistic itself because it is not measurable with respect to L_t for any t). For the proof of Lemma A we will need the following propositions.

Proposition 7.3. *For any $b > 0$ we have*

$$\mathbb{P}\left(\left|\sum_{u \in \mathcal{S}} f(X_{\tau_{N_T(u)}}(u)) - Z\right| > bn^{7/4}\right) \to 0$$

as $n \to \infty$.

Proposition 7.4. *For any $b > 0$ we have*

$$\mathbb{P}\left(|Z - \mathbb{E}(Z)| > bn^{7/4}\right) \to 0$$

as $n \to \infty$.

We defer the proofs of Propositions 7.3 and 7.4 to subsection 7.2. We now give a proof of Lemma A, assuming Propositions 7.3 and 7.4.

Proof of Lemma A. Recall that $W_{\text{dist}} = \sum_{u \in \mathcal{S}} f(X_T(u))$. Since for any tile $u \in \mathcal{S}$ we have $|X_{\tau_{N_T(u)}}(u) - X_T(u)| \le 1$, it follows that $|f(X_{\tau_{N_T(u)}}(u)) - f(X_T(u))| \le \frac{2\pi}{n}$, by (7.5). Thus

$$\begin{aligned} \left|W_{\text{dist}} - \sum_{u \in \mathcal{S}} f(X_{\tau_{N_T(u)}}(u))\right| &\le \sum_{u \in \mathcal{S}} \left|f(X_{\tau_{N_T(u)}}(u)) - f(X_T(u))\right| && (7.6) \\ &\le 2\pi n, && (7.7) \end{aligned}$$

where the last line holds because $|\mathcal{S}| \le n^2$. The main remaining step of the proof is to compute $\mathbb{E}(Z)$. We claim that $\mathbb{E}(Z) \ge cn^{15/8}$, for a universal constant c. Combining this with Propositions 7.3 and 7.4 and (7.7) implies that there exist positive constants b and c such that

$$\mathbb{P}(W_{\text{dist}} \ge cn^{15/8} - 2bn^{7/4} - 2\pi n) \to 1$$

as $n \to \infty$. For sufficiently large n the quantity $cn^{15/8} - 2cn^{7/4} - 2\pi n$ is larger than $\frac{c}{2} n^{15/8}$. Incorporating an extra factor of $\frac{1}{2}$ into the constant c yields Lemma A.

So it remains only to verify that $\mathbb{E}(Z) \ge cn^{15/8}$, for a universal constant c. Recall that $\tau_k(u)$ denotes the kth time that the hole is to the right of tile u, and $(X_{\tau_1(u)}(u), X_{\tau_2(u)}(u), \dots)$ is a simple symmetric random walk on $\mathbf{Z}_n$ with a holding probability. Since the second eigenvalue for this walk λ satisfies $\lambda > \cos \frac{2\pi}{n}$, it follows that for all t we have $\mathbb{E}\left(f(X_{\tau_t(u)}(u) \,|\, X_{\tau_1(u)}(u)\right) \ge f(X_{\tau_1(u)}(u))\lambda^{t-1}$, and since $f(X_{\tau_1(u)}(u)) \ge f(X_0(u)) - \frac{2\pi}{n}$ it follows that

$$\mathbb{E}\left(f(X_{\tau_t(u)}(u))\right) \ge \left(f(X_0(u)) - \frac{2\pi}{n}\right)\lambda^{t-1}.$$

Substituting $t = \widehat{T}$ and summing over $u \in \mathcal{S}$ gives

$$\mathbb{E}Z \geq \left[\sum_{u \in \mathcal{S}} f(X_0(u)) - \frac{2\pi|\mathcal{S}|}{n}\right] \lambda^{\widehat{T}-1}.$$

The expression in square brackets can be bounded below by cn^2 for a universal constant c, since for every $u \in \mathcal{S}$ we have $f(X_0(u)) \geq \frac{1}{2}$. Furthermore, since $\widehat{T} - 1 \leq \epsilon n^2 \log n$ by (7.2) and $\lambda^{n^2} \geq e^{-\mu}$ by (7.1), it follows that

$$\begin{aligned} \mathbb{E}Z &\geq cn^2 \exp(-\mu\epsilon \log n) \\ &= cn^{15/8}. \end{aligned}$$

(Recall that $\mu\epsilon = 1/8$.) This verifies the claim and hence proves the lemma. $\square$

7.2 Proofs of Propositions 7.3 and 7.4

It remains to prove propositions 7.3 and 7.4, which were used in the proof of Lemma A. This is done is subsections 7.2.1 and 7.2.2, respectively.

7.2.1 Proof of Proposition 7.3

Recall that $N_t(u)$ denotes the number of times the hole has been to the immediate right of tile u, up to time t. The main step in the proof of Proposition 7.3 is to show that N_t is well approximated by $t(n^2-1)^{-1}$. We accomplish this using the second moment method.

In order to bound the mean and variance of $\mathbb{E}(N_t(u))$, we use the fact that the position of the hole relative to tile u (that is, the position of the hole minus the position of tile u) behaves like a random walk on a certain graph. Let $\widetilde{G}_n$ be the graph obtained from G_n by deleting the origin and adding an edge from $(-1,0)$ to $(1,0)$ and an edge from $(0,1)$ to $(0,-1)$. (Figure 9 shows $\widetilde{G}_n$ when $n=5$.) Note that if H_t denotes the the position of the hole at time t in the Loyd chain, then $H_t - L_t(u)$ is the same random process as a random walk on $\widetilde{G}_n$. The times $\tau_k(u)$ coincide with the times when the random walk on $\widetilde{G}_n$ is at the vertex $(1,0)$. In Lemmas 7.5 and 7.6 below, we use the connection to the random walk on $\widetilde{G}_n$ to bound the mean and variance of $N_t(u)$.

Lemma 7.5. *There is a universal constant A such that for any tile u and time t we have*

$$\left|\mathbb{E}(N_t(u)) - t(n^2-1)^{-1}\right| \leq A \log t. \tag{7.8}$$

Proof. Let $\{p(x,y)\}$ be transition probabilities for the random walk on $\widetilde{G}_n$. Lemma 8.2 in Appendix A states that there is a universal constant $A > 0$ such that

$$\left|p^t(x,y) - \pi(y)\right| \leq \frac{A}{t}, \tag{7.9}$$

for all $t \geq 1$, where $\pi(y)$ is the stationary probability $(n^2-1)^{-1}$. Since the hole is not initially to the right of tile u, using (7.9) with $x = H_t - L_0(u)$ and $y = (1,0)$ gives

$$\begin{aligned} \left|\mathbb{E}(N_t(u)) - t\pi(y)\right| &\leq \sum_{k=1}^{t} \frac{A}{k} && (7.10) \\ &\leq A \log t. && (7.11) \end{aligned}$$

$\square$

Next we bound the variance of $N_t(u)$.

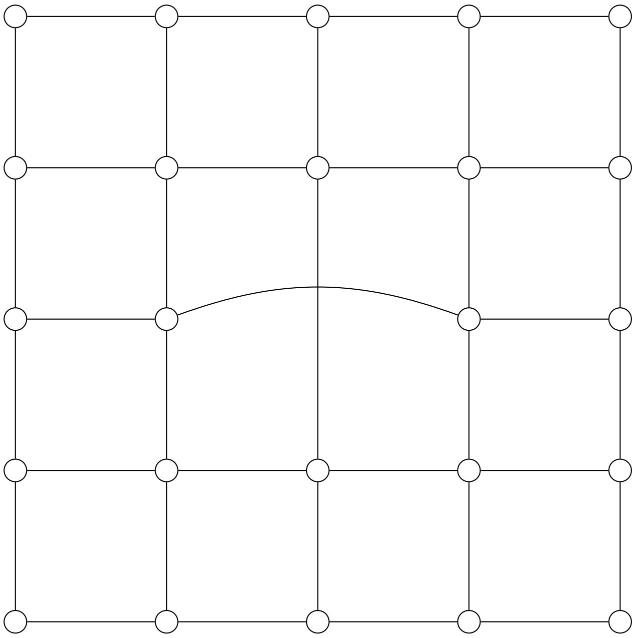

Figure 9: Graph $\widetilde{G}_n$. (Edges connecting top row to bottom row and edges connecting leftmost row to rightmost row are not shown.)

Lemma 7.6. *There is a universal constant C such that for any tile u we have*

$$\mathrm{var}(N_t(u)) \leq Cn^{-2} t \log t,$$

whenever $n^2 \log n \leq t \leq n^5$.

Proof. Fix a tile u and for i with $1 \leq i \leq t$, let I_i be the indicator of the event that the hole is to the right of tile u at time i. Then $N_t(u) = \sum_{i=1}^t I_i(u)$, and hence

$$\mathrm{var}(N_t(u)) \quad = \quad \sum_{i=1}^{t} \mathrm{var}(I_i) + 2 \sum_{1 \leq i < j \leq t} \mathrm{cov}(I_i I_j). \tag{7.12}$$

The first term is at most $\mathbb{E}(N_t(u))$ (since for each i we have $\mathrm{var}(I_i) \leq \mathbb{E}(I_i^2) \leq \mathbb{E}(I_i)$) and recall that Lemma 7.5 implies that $\mathbb{E}(N_t(u))$ is at most $t(n^2-1)^{-1} + A \log t$. To bound the second term in (7.12), note that for each i and j with $i < j$ we have

$$\begin{aligned} \mathrm{cov}(I_i, I_j) \quad &= \quad \mathbb{E}(I_i I_j) - \mathbb{E}(I_i)\mathbb{E}(I_j) \\ &\leq \quad \Big(\pi(y) + \frac{A}{i}\Big)\Big(\pi(y) + \frac{A}{j-i}\Big) - \Big(\pi(y) - \frac{A}{i}\Big)\Big(\pi(y) - \frac{A}{j}\Big), \end{aligned}$$

where in the last line we used Lemma 8.2 to bound each expectation. Expanding each product and then collecting terms gives

$$\Big[\frac{2A}{i} + \frac{A}{j} + \frac{A}{j-i}\Big]\pi(y) + A^2\Big[\frac{1}{i(j-i)} - \frac{1}{ij}\Big].$$

If we sum this over j with $i < j \leq t$, then the result is at most

$$\Big[\frac{2At}{i} + 2A \log t\Big]\pi(y) + \frac{A^2}{i} \log t.$$

If we sum this over i with $1 \le i \le t$, then the result is at most

$$\Big[2At\log t + 2At\log t\Big]\pi(y) + A^2\log^2 t, \tag{7.13}$$

which is of the form $\mathrm{O}(n^{-2}t\log t) + \mathrm{O}(\log^2 n)$. (Note that since $t \le n^5$, we have $\log^2 t = \mathrm{O}(\log^2 n)$.) The result follows if we note that $\log n = n^{-2}(n^2\log n)$, which is at most $n^{-2}t$ whenever $n^2\log n \le t$. □

We will need one more lemma before proving Proposition 7.3, but first we recall Hoeffding's bounds for sums of independent random variables.

Theorem 7.7. *([6]) Let $Y_1, Y_2, \dots$ be i.i.d. random variables and suppose that $\mathbb{E}(Y_1) = 0$ and $|Y_1| \le 1$. Define $S_m = \sum_{i=1}^m Y_i$. Then for all positive integers s and t we have*

$$\mathbb{P}(|S_t - S_s| \ge \alpha) \le 2e^{-\alpha^2/2|t-s|}.$$

Lemma 7.8. *Let $Y_1, Y_2, \dots$ be i.i.d. random variables and suppose that $\mathbb{E}(Y_1) = 0$ and $|Y_1| \le 1$. Define $S_m = \sum_{i=1}^m Y_i$. Fix constants $C > 0$ and β with $\frac{1}{2} < \beta < \frac{3}{4}$. For positive integers n define $M_n = \max \frac{|S_t - S_s|}{|t-s|^\beta}$, where the maximum is over s and t such that*

$$0 \le s \le Cn^4\log n; \qquad 0 \le t \le Cn^4\log n; \qquad |s-t| \ge \sqrt{n}. \tag{7.14}$$

Then for every $p > 1$ there is a constant C_p, which depends only on p, such that

$$\mathbb{E}(M_n^p) \le C_p.$$

Proof. Since each M_n is bounded it is enough to show that $\limsup_{n\to\infty}\mathbb{E}(M_n^p) < \infty$. If $|s-t| > \sqrt{n}$ then applying Hoeffding's bounds with $\alpha = c|t-s|^\beta$ gives

$$\mathbb{P}\Big(\frac{|S_t - S_s|}{|t-s|^\beta} > c\Big) \le 2\exp\left(-\frac{c^2}{2}n^{\beta-\frac{1}{2}}\right). \tag{7.15}$$

Define $p_n(c) := \mathbb{P}(M_n > c)$. There are at most C^2n^{10} pairs (s,t) that satisfy the conditions in (7.14). Thus if n is large enough so that for all $c \ge 1$ we have

$$2C^2n^{10}\exp\left(-\frac{c^2}{2}n^{\beta-\frac{1}{2}}\right) \le e^{-c^2},$$

a union bound implies that for all $c \ge 1$ we have $p_n(c) \le e^{-c^2}$ and hence

$$\begin{aligned}\mathbb{E}(M_n^p) &= \int_0^\infty \mathbb{P}(M_n^p > t)\,dt \\ &\le \int_0^\infty p_n(t^{1/p})\,dt \\ &< \infty.\end{aligned}$$

□

Now that we have Lemmas 7.5, 7.6 and 7.8, we are ready to prove Proposition 7.3

Proof of Proposition 7.3. Since T is $\mathrm{O}(n^4\log n)$, applying Lemma 7.6 with $t = T$ implies that when n is sufficiently large, we have $\mathrm{var}(N_T(u)) \le Cn^2\log^2 n$. It follows that

$$\begin{aligned}\mathbb{E}\Big|N_T(u) - \widehat{T}\Big| &\le \mathbb{E}\Big(\Big|N_T(u) - \mathbb{E}N_T(u)\Big|\Big) + \Big|\mathbb{E}N_T(u) - \widehat{T}\Big| \\ &\le \sqrt{C}n\log n + A\log n,\end{aligned}$$

where in the second line we have used the inequality $\mathbb{E}|X - \mathbb{E}X| \leq \mathrm{sd}(X)$, valid for all random variables X, to bound the first term and Lemma 7.5 to bound the second term. It follows that

$$\mathbb{E}\left|N_T(u) - \widehat{T}\right| \leq Bn \log n, \tag{7.16}$$

for a universal constant B.

Let $S_k = X_{\tau_k(u)}(u) - X_{\tau_1(u)}(u)$, that is, the change of the u random walk after $k-1$ steps. Note that we can write S_k as $Y_1 + Y_2 + \cdots Y_k$, where the Y_i are i.i.d. ± 1 random variables. Fix $\beta \in (\frac{1}{2}, \frac{3}{4})$. Since $|S_{N_T(u)} - S_{\widehat{T}}| \leq |N_T(u) - \widehat{T}|$, and since $N_T(u)$ and $\widehat{T}$ can both be bounded above by $Cn^4 \log n$ for a universal constant C, it follows that if M_n is defined as in the statement of Lemma 7.8, then

$$|S_{N_T(u)} - S_{\widehat{T}}| \leq M_n |N_T - \widehat{T}|^\beta + \sqrt{n}\mathbf{1}\Big(|N_T(u) - \widehat{T}| \leq \sqrt{n}\Big). \tag{7.17}$$

Let C_p be the constant from Lemma 7.8. Applying Hölder's inequality with $p = \frac{1}{1-\beta}$ and $q = \frac{1}{\beta}$ gives

$$\begin{aligned}\mathbb{E}\left(M_n |N_T(u) - \widehat{T}|^\beta\right) &\leq \mathbb{E}(M_n^p)^{1/p}\left(\mathbb{E}|N_T(u) - \widehat{T}|\right)^\beta \\ &\leq C_p^{1/p}\left[Bn\log n\right]^\beta,\end{aligned}$$

where in the last line we have used Lemma 7.8 to bound $\mathbb{E}(M_n^p)^{1/p}$ and (7.16) to bound $\mathbb{E}|N_T(u) - \widehat{T}|$.

Taking expectations in (7.17) gives

$$\mathbb{E}|S_{N_T(u)} - S_{\widehat{T}}| \quad \leq \quad C_p^{1/p}\left[Bn\log n\right]^\beta + \sqrt{n}.$$

Now, let $\gamma \in (\beta, \frac{3}{4})$. Then there is a constant B' such that

$$\mathbb{E}|S_{N_T(u)} - S_{\widehat{T}}| \quad \leq \quad B' n^\gamma. \tag{7.18}$$

Since $S_{N_T(u)}(u) - S_{\widehat{T}}(u) = X_{\tau_{N_T(u)}}(u) - X_{\tau_{\widehat{T}}(u)}(u)$ from the definition of S_k, combining (7.18) with (7.5) gives

$$\begin{aligned}\mathbb{E}\Big(\Big|f(X_{\tau_{N_T(u)}}(u)) - f(X_{\tau_{\widehat{T}}(u)}(u))\Big|\Big) &\leq \frac{2\pi}{n}B'n^\gamma &(7.19)\\ &= \frac{2\pi}{n}B'n^{\gamma-1} &(7.20)\end{aligned}$$

for a constant B'. Define $X_{\mathrm{final}}(u) = X_{\tau_{\widehat{T}}(u)}(u)$. Summing (7.20) over $u \in \mathcal{S}$ gives

$$\mathbb{E}\Big(\Big|\sum_{u\in\mathcal{S}} f(X_{\tau_{N_T(u)}}(u)) - \sum_{u\in\mathcal{S}} f(X_{\mathrm{final}}(u))\Big|\Big) \leq \frac{2\pi}{n}B'n^{\gamma+1}.$$

Combining this with Markov's inequality yields the proposition, since $\gamma < \frac{3}{4}$. □

7.2.2 Proof of Proposition 7.4

We prove Proposition 7.4 using the method of bounded differences. The main step is to show that each step of the Loyd process has a small effect on the conditional expectation of Z, which we prove via Lemma 7.9 below.

Define $f_{\mathrm{final}}(u) := f(X_{\mathrm{final}}(u))$, so that we can write Z as

$$Z = \sum_{u\in\mathcal{S}} f_{\mathrm{final}}(u).$$

Let $\mathcal{H}_t = (L_0, L_1, \dots, L_t)$ be the history of the Loyd process up to time t. We call the Markov chain $(\mathcal{H}_t : t \geq 0)$ the *history process*. If $H = (L_0, \dots, L_k)$ is a state of the history process, we write $L(H)$ for the Loyd configuration L_k.

Let $\mathcal{H} \to \widehat{\mathcal{H}}$ be a possible transition of the history process. We aim to compare the distribution of Z when the history process starts at $\mathcal{H}$ versus when it starts from $\widehat{\mathcal{H}}$. We shall refer to the history process started from $\mathcal{H}$ (respectively, $\widehat{\mathcal{H}}$) as the *primary* (respectively, *secondary*) history process.

Convention. If a random variable W is defined in terms of the primary process, we write $\widehat{W}$ for the corresponding random variable defined in terms of the secondary process, and similarly for events.

Lemma 7.9. *We have*

$$|\mathbb{E}(Z) - \mathbb{E}(\widehat{Z})| \leq \frac{D \log n}{n},$$

for a universal constant D.

Proof. Our main tool is coupling. Note that to demonstrate a coupling of the primary and secondary history processes, it is sufficient to demonstrate a coupling of the Loyd process started from $L := L(\mathcal{H})$ and the Loyd process started from $\widehat{L} := L(\widehat{\mathcal{H}})$. We call these processes the primary and secondary Loyd processes, respectively.

We start by bounding $|\mathbb{E}(f_{\mathrm{final}}(u)) - \mathbb{E}(\widehat{f}_{\mathrm{final}}(u))|$ for the case when u is the tile swapped with the hole in the transition from L to $\widehat{L}$. We can couple the secondary Loyd process with the primary Loyd process so that the way that the hole moves after the first time it is to the right of tile u is the same in both processes. Since with this coupling we have $|X_{\mathrm{final}}(u) - \widehat{X}_{\mathrm{final}}(u)| \leq 1$, equation (7.5) implies that

$$|\mathbb{E}(f_{\mathrm{final}}(u)) - \mathbb{E}(\widehat{f}_{\mathrm{final}}(u))| \leq \frac{2\pi}{n}. \tag{7.21}$$

Let $\mathcal{S}'$ be the set of tiles in $\mathcal{S}$ that are not swapped with the hole in the transition from L to $\widehat{L}$. We now consider the tiles in $\mathcal{S}'$. It will be convenient to group the tiles in columns (i.e., group them according to their x-coordinates) and then consider the columns one at a time.

Let H_t be the location of the hole at time t in the primary Loyd process, and suppose that $H_0 = (h_x, h_y)$. Let C be a column in V_n, that is, a set of the form $\{(j,k) : k \in \mathbf{Z}_n\}$ for some $j \in \mathbf{Z}_n$, and suppose that $|h_x - j| = d$ (that is, the hole is initially a distance d from C), where $d \in \{0, 1, 2, \dots\}$. We claim that there is a universal constant D such that

$$\Big| \sum_{u \in \mathcal{S}' \cap C} \mathbb{E}(f_{\mathrm{final}}(u) - \widehat{f}_{\mathrm{final}}(u)) \Big| \leq \frac{D}{n(d+1)}. \tag{7.22}$$

Summing this over columns C and combining this with (7.21) proves the Lemma.

We now prove the claim. We verify (7.22) by constructing a coupling of the primary Loyd process and the secondary Loyd process. The coupling is designed so that if the hole is initially far away from column C, then H_t is likely to couple with $\widehat{H}_t$ before it gets close to column C.

Let C_L and C_R be the columns to the immediate left and right, respectively, of C. We now give a rough description of the coupling. The nature of the coupling will depend on whether the hole moves horizontally or vertically in the transition from L to $\widehat{L}$. If the hole moves horizontally (respectively, vertically), then the trajectory of H_t is the reflection of the trajectory of $\widehat{H}_t$ about a vertical (respectively, horizontal) axis, up until the time when either the holes have coupled or one of them has reached column C, C_R or C_L. We now give a more formal description in the case where $\widehat{H}_0 = (h_x + 1, h_y)$. (The other cases are similar. In the case where the hole moves vertically in the transition from L to $\widehat{L}$, the coupling is the same, except that the roles of vertical and horizontal moves are reversed.)

The coupling in the case where $\widehat{H}_0 = (h_x + 1, h_y)$

1. If $H_t = \widehat{H}_t$, then we couple so that $H_{t+1} = \widehat{H}_{t+1}$;
2. else, if either H_t or $\widehat{H}_t$ is in column C, C_R or C_L, then the holes move independently;
3. else, if H_t is to the immediate left of $\widehat{H}_t$, we use the following rule.

primary	secondary	probability
$\leftarrow$	$\rightarrow$	$1/8$
$\rightarrow$	do nothing	$1/8$
$\uparrow$	$\uparrow$	$1/8$
$\downarrow$	$\downarrow$	$1/8$
do nothing	$\leftarrow$	$1/8$
do nothing	do nothing	$3/8$

4. else, we use the following rule.

primary	secondary	probability
$\leftarrow$	$\rightarrow$	$1/8$
$\rightarrow$	$\leftarrow$	$1/8$
$\uparrow$	$\uparrow$	$1/8$
$\downarrow$	$\downarrow$	$1/8$
do nothing	do nothing	$1/2$

Note that if the x-coordinate of H_t takes the value $h_x + 1$ before either H_t or $\widehat{H}_t$ hits C, C_R or C_L then the holes couple before either of them affects tile s.

Let $\widetilde{T}$ be the first time either H_t or $\widehat{H}_t$ hits columns C, C_L or C_R. Let E be the event that the holes have not coupled before time $\widetilde{T}$. We claim that there is a universal constant $\mathcal{K} > 0$ such that

$$\mathbb{P}(E) \leq \frac{\mathcal{K}}{d}. \tag{7.23}$$

Recall that d is the initial distance between the hole and column C. We may assume that $d > 0$ since otherwise the bound is trivial provided that $\mathcal{K} \geq 1$.

It is enough to verify (7.23) in the folowing two cases, since we can always reduce to one of these cases by interchanging the roles of H_t and $\widehat{H}_t$ if necessary:

1. $\widehat{H}_0$ is to the immediate right of H_0.
2. $\widehat{H}_0$ is immediately below H_0.

In the first case, (7.23) follows from part (i) of Lemma 9.1 in Appendix B, since the event E occurs only if time $\widetilde{T}$ occurs before the x-coordinate of H_t takes the value $h_x + 1$. In the second case, (7.23) follows from part (ii) of Lemma 9.1, since in this case the event E occurs only if time $\widetilde{T}$ occurs before the y-coordinate of H_t takes the value $h_y - 1$.

Let T_C be the first time that the hole is in column C. For tiles $u \in \mathcal{S}$ that are initially in column C, let $T_R(u)$ (respectively, $T_L(u)$) be the first time that the hole is to the immediate right (respectively, left) of tile u. Let R_u be the event that $T_R(u) = \min(T_R(u), T_L(u), T_C)$ and let L_u be the event that $T_L(u) = \min(T_R(u), T_L(u), T_C)$. Define

$$z_R = \mathbb{E}\left(f_{\text{final}}(u) \,|\, T_R < T_L\right), \qquad z_L = \mathbb{E}\left(f_{\text{final}}(u) \,|\, T_L < T_R\right).$$

Note that (7.5) implies that

$$|z_R - z_L| \leq \frac{2\pi}{n}. \tag{7.24}$$

We say that the hole is *beside* a tile if it is to its immediate right or immediate left. Note that if the hole starts in the same column as tile s, then the next time the hole is beside tile s it is equally likely to be to its right as to its left. It follows that

$$\mathbb{E}(f_{\text{final}}(u)) = \mathbb{P}(R_u)z_R + \mathbb{P}(L_u)z_L + [1 - \mathbb{P}(R_u) - \mathbb{P}(L_u)](\tfrac{1}{2}z_R + \tfrac{1}{2}z_L).$$

Rearranging terms gives

$$\mathbb{E}(f_{\text{final}}(u)) = \tfrac{1}{2}\Big[(z_R + z_L) + \mathbb{P}(R_u)(z_R - z_L) + \mathbb{P}(L_u)(z_L - z_R)\Big]. \tag{7.25}$$

Similarly, we also have

$$\mathbb{E}(\widehat{f}_{\text{final}}(u)) = \tfrac{1}{2}\Big[(z_R + z_L) + \mathbb{P}(\widehat{R}_s)(z_R - z_L) + \mathbb{P}(\widehat{L}_s)(z_L - z_R)\Big]. \tag{7.26}$$

Replacing each probability in (7.25) and (7.26) with the expectation of an appropriate indicator random variable, and then subtracting (7.26) from (7.25), gives

$$\mathbb{E}(f_{\text{final}}(u)) - \mathbb{E}(\widehat{f}_{\text{final}}(u)) = \tfrac{1}{2}\mathbb{E}(\mathbf{1}_{R_u} - \mathbf{1}_{\widehat{R}_s})\Delta - \tfrac{1}{2}\mathbb{E}(\mathbf{1}_{L_u} - \mathbf{1}_{\widehat{L}_s})\Delta, \tag{7.27}$$

where $\Delta := z_R - z_L$. Hence

$$\Big|\mathbb{E}(f_{\text{final}}(u)) - \mathbb{E}(\widehat{f}_{\text{final}}(u))\Big| \le |\Delta| \max(\mathbb{E}(\mathbf{1}_{R_u} - \mathbf{1}_{\widehat{R}_s}), \mathbb{E}(\mathbf{1}_{L_u} - \mathbf{1}_{\widehat{L}_s})).$$

Note that $\mathbf{1}_{R_u} - \mathbf{1}_{\widehat{R}_s}$ and $\mathbf{1}_{L_u} - \mathbf{1}_{\widehat{L}_s}$ are both 0 on the event that the holes couple before either one hits C_R or C_L. It follows that

$$\Big|\mathbb{E}(f_{\text{final}}(u)) - \mathbb{E}(\widehat{f}_{\text{final}}(u))\Big| \le |\Delta| \cdot \mathbb{E}(Y(u) + \widehat{Y}(u)), \tag{7.28}$$

where $Y(u)$ is the indicator of the event that the hole is beside tile u before time T_C. Let $Y = \sum_{u\in C} Y(u)$ be the total number of positions in column C_L and C_R visited before time T_C. Summing over $u \in C$ gives

$$\sum_{u\in C}\Big|\mathbb{E}(f_{\text{final}}(u)) - \mathbb{E}(\widehat{f}_{\text{final}}(u))\Big| \le |\Delta| \cdot \mathbb{E}(Y + \widehat{Y}) \tag{7.29}$$

Note that Y and $\widehat{Y}$ are both 0 unless the event E occurs and recall that (7.23) gives $\mathbb{P}(E) \le \frac{\mathcal{K}}{d+1}$. Furthermore, the condional distribution of both Y and $\widehat{Y}$ given E is geometric($\frac{1}{4}$), since each time the hole is in column C_R or C_L, it moves to column C in the next step with probability $\frac{1}{4}$. It follows that

$$\begin{aligned}\mathbb{E}(Y + \widehat{Y}) &\le \frac{\mathcal{K}}{d+1}\mathbb{E}(Y + \widehat{Y} \mid E) &(7.30)\\ &= \frac{8\mathcal{K}}{d+1}. &(7.31)\end{aligned}$$

Finally, recall that $\Delta = z_R - z_L$ and hence $|\Delta| \le \frac{2\pi}{n}$ by (7.24). Combining this with (7.29) and (7.31) vertifies (7.22), which proves the lemma. □

Now that we know there are bounded differences, we are ready to prove Proposition 7.4:

Proof of Proposition 7.4. We need to show that for any $b > 0$ we have

$$\mathbb{P}(|Z - \mathbb{E}(Z)| > bn^{7/4}) \to 0$$

as $n \to \infty$, where $Z = \sum_{u\in\mathcal{S}} f(X_{\text{final}}(u))$.

Recall that $\tau_k(u)$ is the kth time at which the hole is to the immediate right of tile u. Define $\tau = \max_{u\in\mathcal{S}} \tau_{\widehat{T}}(u)$. Let $\mathcal{F}_t = \sigma(L_1, \dots, L_t)$ and consider the Doob martingale

$$M_t := \mathbb{E}(Z \,|\, \mathcal{F}_t).$$

The idea of the proof will be to evaluate the martingale at a suitably chosen time K. The value of K will be chosen to be large enough so that $\tau \le K$ with high probability, but small enough so that the Azuma-Hoeffding inequality will give a good large deviation bound for M_K. To these ends, we choose $K = n^5$. Note that Z is determined by time τ. Hence $M_K = Z$ unless $\tau > K$. Furthermore, we have $\mathbb{E}(M_K) = \mathbb{E}(Z)$. It follows that

$$\mathbb{P}(|Z - \mathbb{E}(Z)| > bn^{7/4}) \le \mathbb{P}(|M_K - \mathbb{E}(M_K)| > bn^{7/4}) + \mathbb{P}(\tau > K). \tag{7.32}$$

We now bound each term on the righthand side of (7.32). We start with the first term. Lemma 7.9 implies that

$$|M_t - M_{t-1}| \le \frac{D \log n}{n},$$

for t with $1 \le t \le K$. Thus the Azuma-Hoeffding bound gives

$$\mathbb{P}(|M_K - \mathbb{E}(M_K)| \ge x) \quad\le\quad 2\exp\left(\frac{-x^2}{2\sum_{i=1}^K \alpha^2}\right), \tag{7.33}$$

where $\alpha = \frac{D\log n}{n}$. Substituting $x = bn^{7/4}$ and $K = n^5$ into (7.33) gives

$$\begin{aligned}\mathbb{P}(|M_K - \mathbb{E}(M_K)| \ge bn^{7/4}) \quad&\le\quad 2\exp\left(\frac{-b^2 n^{7/2}}{2n^3 D^2 \log^2 n}\right) &(7.34)\\ &=\quad 2\exp\left(\frac{-b^2 n^{1/2}}{2D^2\log^2 n}\right), &(7.35)\end{aligned}$$

which converges to 0 as $n \to \infty$.

Next, we bound $\mathbb{P}(\tau > K)$. Note that $\tau_{\widehat{T}}(u) \le K$ whenever $N_K(u) \ge \widehat{T}$. Furthermore, since $K = n^5$, Lemmas 7.5 and 7.6 imply that for sufficiently large n we have

$$\mathbb{E}(N_K(u)) \ge n^3 - \mathrm{O}(\log n); \qquad \mathrm{var}(N_K(u)) = \mathrm{O}(n^3 \log n).$$

Note also that $\widehat{T}$ is $\mathrm{o}(n^3)$. Thus Chebyshev's inequality implies that $\mathbb{P}(N_K(u) < \widehat{T})$ is $\mathrm{O}\left(\frac{\log n}{n^3}\right)$, and hence $\mathbb{P}(\tau_{\widehat{T}} > K)$ is $\mathrm{O}\left(\frac{\log n}{n^3}\right)$. Thus a union bound implies that $\mathbb{P}(\tau > K)$ is $\mathrm{O}\left(\frac{\log n}{n}\right)$, and hence converges to 0 as $n \to \infty$. This completes the proof. □

8 Appendix A: Probability bounds for random walk on $\widetilde{G}_n$

In this section we derive bounds on transition probabilities for random walk on $\widetilde{G}_n$. First, we give some definitions and extract some necessary results from [12].

Let $\{q(x,y)\}$ be transition probabilities for a Markov chain on a finite state space V with stationary distribution π. For $S \subset V$, define the "boundary size" $|dS| = \sum_{x\in S, y\in S^c} \pi(x)q(x,y)$. Following [5], we call $\Phi_S := \frac{|\partial S|}{\pi(S)}$ the *conductance* of S. Write $\pi_* := \min_{x\in V} \pi(x)$ and define $\Phi(r)$ for $r \in [\pi_*, 1/2]$ by

$$\Phi(r) = \inf\left\{\Phi_S : \pi(S) \le r\right\}. \tag{8.1}$$

For $r > 1/2$, let $\Phi(r) = \Phi(1/2)$. We call Φ the *isoperimetric profile*. We recall the following theorem from [12].

Theorem 8.1. *Suppose that $q(x,x) \ge \frac{1}{2}$ for all $x \in V$. If*

$$t \ge 1 + \int_{\pi_*}^{4/\epsilon} \frac{4du}{u\Phi^2(u)}\,, \tag{8.2}$$

then

$$\left|\frac{q^t(x,y) - \pi(y)}{\pi(y)}\right| \le \epsilon. \tag{8.3}$$

Lemma 8.2. *Let $\{p(x,y\}$ be transition probabilities for the lazy random walk on $\widetilde{G}_n$ and let π be the stationary distribution. There is a universal constant $A > 0$ such that*

$$\left|p^t(x,y) - \pi(y)\right| \le \frac{A}{t}, \tag{8.4}$$

for all $t \ge 1$.

Proof. Recall that G_n denotes the $n \times n$ torus $\mathbf{Z}_n^2$. We write Φ (respectively, $\widetilde{\Phi}$) for the conductance profile for the lazy random walk on G_n (respectively, $\widetilde{G}_n$). It is well known that Φ satisfies

$$\Phi(u) \ge \frac{C}{n\sqrt{u}}, \tag{8.5}$$

for a universal constant $C > 0$.

Let $\widetilde{V}_n$ be the vertex set of $\widetilde{G}_n$. Since for $S \subset \widetilde{V}_n$, the boundary size and stationary probability of S, with respect to random walk on $\widetilde{G}_n$, are within constant factors of the corresponding quantities with respect to random walk on G_n, it follows that the conductance profile $\widetilde{\Phi}$ for random walk on $\widetilde{G}_n$ satisfies the similar inequality

$$\widetilde{\Phi}(u) \ge \frac{\widetilde{C}}{n\sqrt{u}}, \tag{8.6}$$

for a universal constant $\widetilde{C} > 0$.

Fix $0 < \alpha < 1$. Using Theorem 8.1 with $\epsilon = \alpha/\pi(y)$ gives

$$\left|p^t(x,y) - \pi(y)\right| \le \alpha \tag{8.7}$$

whenever

$$t \ge 1 + \int_{\pi_*}^{4\pi(y)/\alpha} \frac{4du}{u\widetilde{\Phi}^2(u)}. \tag{8.8}$$

Equation (8.6) implies that the righthand side of (8.8) is at most

$$\begin{aligned} 1 + \int_{\pi_*}^{4\pi(y)/\alpha} 4\widetilde{C}^{-2} n^2\, du \quad &\le \quad 1 + \frac{16\pi(y)n^2}{\widetilde{C}^2\alpha} \\ &\le \quad \frac{A}{\alpha}, \end{aligned}$$

for a universal constant $A > 0$, where the last line follows from the fact that $\pi(y)$ is $\mathrm{O}(n^{-2})$. It follows that

$$\left|p^t(x,y) - \pi(y)\right| \le \frac{A}{t}, \tag{8.9}$$

for all $t \ge 1$, and the proof is complete. □

9 Appendix B

Lemma 9.1. *Let $W_t = (X_t, Y_t)$ be a simple random walk on $\mathbf{Z}^2$, started at $(0,1)$. Fix a positive integer k and let L_1, L_2 and L_3 be the lines $y = 0$, $y = k$ and $|x| = k$, respectively. Let T_2 and T_3 be the hitting times of $L_1 \cup L_2$ and $L_1 \cup L_3$, respectively.*

(i)

$$\mathbb{P}(W_{T_2} \in L_2) = \frac{1}{k}.$$

(ii)

$$\mathbb{P}(W_{T_3} \in L_3) \le \frac{2}{k}.$$

Proof. **(i)** This is immediate by the optional stopping theorem because Y_t is a bounded martingale and T_2 is a stopping time.
(ii) Let $T = \min(T_2, T_3)$. Note that T_3 and T are stopping times. A routine calculation shows that $Y_t^2 - X_t^2$ is a martingale. It follows that $Y_{t\wedge T}^2 - X_{t\wedge T}^2$ is a bounded submartingale. Thus the optional stopping theorem implies that

$$\mathbb{E}\left(Y_T^2 - X_T^2\right) = \mathbb{E}\left(Y_0^2 - X_0^2\right) = 1,$$

and hence

$$\mathbb{E}(X_T^2) < \mathbb{E}(Y_T^2). \tag{9.1}$$

But since $Y_{t\wedge T}^2$ is a bounded submartingale and $T \le T_2$, we have

$$\begin{aligned} \mathbb{E}\left(Y_T^2\right) &\le \mathbb{E}\left(Y_{T_2}^2\right) \\ &= k^2 \mathbb{P}(Y_{T_2} = k) \\ &= k, \end{aligned}$$

where the last line holds because $\mathbb{P}(Y_{T_2} = k) = \frac{1}{k}$ by part (i) of the lemma. Combining this with (9.1) gives

$$\mathbb{E}\left(X_T^2 + Y_T^2\right) < 2k. \tag{9.2}$$

It follows that

$$\begin{aligned} \mathbb{P}(W_T \in L_2 \cup L_3) &= \mathbb{P}(X_T^2 + Y_T^2 \ge k^2) \\ &\le \frac{1}{k^2}\mathbb{E}(X_T^2 + Y_T^2) \\ &\le \frac{2}{k}, \end{aligned}$$

where first inequality is Markov's and the second follows from (9.2). This verifies (ii) because $W_T \in L_2 \cup L_3$ whenever $W_T \in L_3$. □

References

[1] Diaconis, P. Group representations in probability and statistics. Institute of Mathematical Statistics, 1988. MR-0964069

[2] Diaconis, P. and Saloff-Coste, L. Comparison techniques for random walk on finite groups. *Annals of Probability* **21** (1993), pp.2131–2156. MR-1245303

[3] Diaconis, P. and Saloff-Coste, L. Logarithmic Sobolev inequalities for finite Markov chains. *Annals of Applied Probability* **6**(3) (1996), pp.695–750. MR-1410112

[4] Diaconis, P. and Saloff-Coste, L. Random walks on finite groups: a survey of analytic techniques. In Probability Measures on Groups and Related Structures 11 (Z.H. Heyer, ed.) 44–75.

[5] Jerrum, M. R. and Sinclair, A. J. (1989). Approximating the permanent. *SIAM Journal on Computing* **18**, 1149–1178. MR-1025467

[6] Hoeffding, W. (1963), Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association* **58** (301), pp. 13–30. MR-0144363

[7] Johnson, W. and Story, W. (1879) Notes on the "15" puzzle. *American Journal of Mathematics* **2** (4), pp.397–404.

[8] Lee, T.Y. and Yau, H.T. (1998). Logarithmic Sobolev inequality for some models of random walks. *Annals of Probability* **26**, pp.1855–1873.

[9] Levin, D., Peres, Y., and Wilmer, E. Markov chains and mixing times. American Mathematical Society, Providence, RI, 2009. MR-2466937

[10] Lezaud, P. Chernoff-type bound for finite Markov chains. *Annals of Probability* **8**, pp.849–867.

[11] Morris, B. (2006). The mixing time for simple exclusion. *Annals of Applied Probability* **16**, pp.615–635.

[12] Morris, B. and Peres, Y. (2005). Evolving sets, mixing and heat kernel bounds. *Probability Theory and Related Fields* **133**, pp.245–266.

[13] Saloff-Coste, L. and Zuniga, J. Refined estimates for some basic random walks on the symmetric and alternating groups, Latin American Journal of Probability and Mathematical Statistics 4, 359-392, 2008.

[14] Wilson, D. (2004) Mixing times of lozenge tiling and card shuffling Markov chains. *Ann. Appl. Prob.* **14**, pp. 274–325.

[15] Wilson, M. (1974). Graph puzzles, homotopy, and the alternating group. *Journal of Combinatorial Theory* Series B. **16**, pp.86–96.

[16] Yau, Horng-Tzer (1997). Logarithmic Sobolev inequality for generalized simple exclusion processes. *Probability Theory and Related Fields* **109**, pp.507–538.

Acknowledgments. We are grateful to I. Benjamini for bringing the problem of the mixing time of the fifteen puzzle to our attention.