




MIXING TIME BOUNDS FOR EDGE FLIPPING ON REGULAR GRAPHS

YUNUS EMRE DEMİRÇİ,* *Queen's University*

ÜMİT IŞLAK,** *Boğaziçi University*

ALPEREN ÖZDEMİR ,*** *Georgia Institute of Technology*

Abstract

An edge flipping is a non-reversible Markov chain on a given connected graph, as defined in Chung and Graham (2012). In the same paper, edge flipping eigenvalues and stationary distributions for some classes of graphs were identified. We further study edge flipping spectral properties to show a lower bound for the rate of convergence in the case of regular graphs. Moreover, we show by a coupling argument that a cutoff occurs at $\frac{1}{4}n \log n$ for the edge flipping on the complete graph.

Keywords: Coupling; edge flipping; hyperplane arrangements; Markov chains; regular graphs

2020 Mathematics Subject Classification: Primary 60C05

Secondary 60F05; 60G42

1. Introduction

Edge flipping is a random process on graphs where an edge of a graph is randomly chosen at each step and both its endpoints are colored either blue with probability p or red with probability $q = 1 - p$. We are interested in the long-term vertex color configurations of the graph and a derived statistic, the frequency of blue vertices. In most cases, the initial configuration of vertex colors is taken to be monochromatic (all blue or all red). Stationary distributions for paths and cycles [7] and for complete graphs [6] have been studied, and various asymptotic results on stationary color configurations obtained for those cases. Our emphasis will be on the mixing time of the chain.

The set of color configurations of a graph with n vertices is in fact the n -dimensional hypercube, $\{0, 1\}^n$, so it is expected that edge flipping has similarities with other Markov chains on the space. Take, for example, the Ehrenfest urn model. In this model there are two urns containing n balls in total. At every stage a ball is randomly chosen and is moved to the opposite urn. In our case, the urns can be associated with the two different colors, and the dynamics is picking two balls at random and placing them together in one of the two urns. The crucial difference is that edge flipping is non-reversible, unlike the former model. In order to see that, consider two configurations, one with all the vertices blue and the other with all but one vertex

Received 5 February 2022; revision received 13 December 2022.

* Postal address: Queen's University, Department of Mathematics and Statistics, Kingston, Ontario, Canada. Email: 21yed@queensu.ca

** Postal address: Boğaziçi University, Department of Mathematics, Istanbul, Turkey. Email: umit.islak1@boun.edu.tr

*** Postal address: Georgia Institute of Technology, School of Mathematics, Atlanta, GA, USA. Email: [aозdemir6@gatech.edu](mailto:aozdemir6@gatech.edu)

© The Author(s), 2023. Published by Cambridge University Press on behalf of Applied Probability Trust.

blue. Then the chain moves from the latter to the former with positive probability, but the other direction has zero probability.

Edge flipping can also be considered a random walk on the set of chambers of Boolean arrangements, which are particular cases of hyperplane arrangements. The eigenvalues of hyperplane chamber walks were first obtained in [3]. Markov chains such as the Tsetlin library and various riffle shuffle models were shown to be examples of these walks in the same paper. The transition matrix of the random walk was shown to be diagonalizable in [5] using algebraic topology, and a non-compelling condition for the uniqueness of the stationary distribution was provided. In a broader context, [4] gave a ring-theoretical proof for the diagonalizability of random walks on left-regular bands, which include hyperplane arrangements. Finally, a combinatorial derivation of the spectrum of the random walk was shown in [1]. The eigenvectors of hyperplane arrangements were studied in [14, 16].

We use the spectral analysis of the Markov chain, as developed in the aforementioned papers, and combine it with various probabilistic techniques to address the rate of convergence to its stationary distribution with respect to total variation distance. See [13] for an example of separation distance bounds on hyperplane random walks. The spectral techniques are limited since edge flipping is non-reversible. In particular, a frequently used ℓ^2 bound on the total variation distance by the sum of squares of eigenvalues is not available (see [11, Lemma 12.16]). Rather, an upper bound involving the sum of the eigenvalues was obtained by combining coupling techniques with spectral analysis in [3] and [5].

We are ready to present our results on the mixing time of edge flipping. The mixing time of a Markov chain is the time required for the distance between the law of the chain and the stationary distribution to be small. The formal definitions for the distance notion and the mixing time are given at the beginning of Section 3. Our first result is for the most general case. In Theorem 4, we show an upper and a lower bound on the mixing time for edge flipping on connected graphs with n vertices, which differ by a logarithmic factor of n . Then we specialize to regular graphs and obtain bounds of order $n \log n$ with different constants in Theorem 5. This is known as *pre-cutoff*, which will not be defined here; we refer to [11, Section 18.1] for its definition. A finer result is known as *cutoff*, which refers to a sharp transition in the distance of the Markov chain from its stationary distribution. We refer the reader to Section 3.3 for its definition.

A complete characterization of cutoff is still an open problem. There are conditions that guarantee its existence in broad classes of Markov chains. For example, the *product condition* implies the existence of cutoff for some classes of reversible chains [2], yet there are counterexamples too; see [11, Section 18, Notes]. Recently, an entropic criterion was proposed in [15] which requires *symmetry of probabilities*, a less restrictive property than reversibility. However, that condition does not apply to edge flipping either for the aforementioned reason of it not being reversible. So, we are not able to tell immediately whether edge flipping shows a cutoff or not. The main result of this paper is to find matching constants for the upper and lower bounds on the mixing time in the case of complete graphs.

Theorem 1. *Let K_n be the complete graph with n vertices. Edge flipping on K_n exhibits a cutoff at time $\frac{1}{4}n \log n$ with a window of order n .*

In order to develop an intuition for this exact rate, let us reconsider random walks on hypercubes. In particular, let us take the lazy random walk on $\{0, 1\}^n$. In comparison to the lazy random walk on the hypercube, the number of vertices that change color at each step is 1 on average instead of $\frac{1}{2}$, so we expect to have twice as fast a chain as the former, which happens to be the case; see [11, Theorem 18.3].

The outline of the paper is as follows. The next section provides background on hyperplane arrangements and the Markov chains defined on them. In Section 3 we formally present our results. Then we discuss how to bound the rate of convergence to the stationary distribution, and supply the proofs of our theorems.

2. Hyperplane random walks

Let $\mathcal{A} = \{H_i\}_{i=1}^n$ be a collection of hyperplanes in a real vector space V . The collection is called *central* if $\bigcap_{i=1}^n H_i \neq \emptyset$. Each H_i divides V into two open half-spaces. Let us denote them by H_i^+ and H_i^- , and let H_i^0 stand for the hyperplane itself. A *face* F is a non-empty subset of V determined by hyperplanes as $F = \bigcap_{i=1}^n H_i^{\sigma_i(F)}$, where $\sigma_i(F) \in \{+, -, 0\}$. So, we can represent any face by a sign sequence $\{\sigma_i(F)\}_{i=1}^n$. The faces with $\sigma_i(F) \neq 0$ for all i are called *chambers*. The set of faces and the set of chambers are denoted by \mathcal{F} and \mathcal{C} respectively. Then we define a left-multiplication on \mathcal{F} by

$$\sigma_i(FF') = \begin{cases} \sigma_i(F) & \text{if } \sigma_i(F) \neq 0, \\ \sigma_i(F') & \text{if } \sigma_i(F) = 0. \end{cases} \tag{1}$$

The operation defined above generates a random walk on \mathcal{C} with transition probabilities

$$P(C, C') = \sum_{\substack{F \in \mathcal{F} \\ FC=C'}} w(F), \tag{2}$$

where $C, C' \in \mathcal{C}$, and w is any probability distribution over faces.

To state the results on the eigenvalues of the matrix associated with the transition probabilities, we first consider a partial order on \mathcal{A} defined via (1) as $F \leq F' \Leftrightarrow FF' = F'$. This implies that $F \leq F'$ if and only if either $\sigma_i(F) = \sigma_i(F')$ or $\sigma_i(F) = 0$ for each $i = 1, \dots, n$. We can view left-multiplication by F as a projection from \mathcal{A} onto $\mathcal{A}_{\geq F} := \{F' \in \mathcal{A} : F \leq F'\}$.

Next, we define another partial order by

$$F \preceq F' \Leftrightarrow F'F = F'. \tag{3}$$

In other words, $F \preceq F'$ if and only if $\sigma_i(F') = 0$ implies $\sigma_i(F) = 0$. The equivalence classes of (3) give a semilattice L , and L is a lattice if and only if \mathcal{A} is central [17]. The equivalence class of F with respect to (3) is called the *support* of F , and is denoted by $\text{supp } F$. In lattice terminology, we have $\text{supp } FF' = \text{supp } F \vee \text{supp } F'$. The elements of L are called *flats*. Furthermore, if $\text{supp } F = \text{supp } F'$, then the projection space $\mathcal{A}_{\geq F}$ is isomorphic to $\mathcal{A}_{\geq F'}$. Thus the following is well-defined: $c_X = |\mathcal{A}_{\geq F}|$ for any $F \in \mathcal{F}$ with $X = \text{supp } F$.

Observe that the support of the chambers is maximal in L , and the set of chambers is closed under left-multiplication. Therefore, the Markov chain is well-defined over the set of chambers. Its spectral profile is as follows.

Theorem 2. ([5, Theorem 1.]) *Let $\{w(F)\}_{F \in \mathcal{F}}$ be a probability distribution over the faces of a central hyperplane \mathcal{A} . The eigenvalues of the random walk given by the transition probabilities (2) are indexed by the lattice L . For each flat $X \in L$, the associated eigenvalue is $\lambda_X = \sum_{\text{supp } F \subseteq X} w(F)$, with multiplicity m_X satisfying $\sum_{X \preceq Y} m_Y = c_X$.*

We note that the multiplicities have the explicit expression $m_X = \sum_{X \preceq Y} \mu(X, Y)c_Y$ by Möbius inversion, where μ is the Möbius function of the lattice L . For a Boolean arrangement,

the partial order (3) is set inclusion, and its Möbius function is simply $\mu(X, Y) = (-1)^{|Y|-|X|}$ by the inclusion–exclusion principle. In this case, the eigenvalues are indexed by the Boolean lattice.

Finally, we look at the edge flipping case. Let G be a connected graph with n vertices and m edges. Denote the vertex set of G by $V(G)$ and the edge set by $E(G)$. Each vertex i is associated with a hyperplane H_i in \mathbb{R}^n and the arrangements are central. All color configurations where the vertex i is blue reside in the half-space H_i^+ , and the color configurations where the vertex i is red fall into H_i^- .

The faces that correspond to color configurations on subsets of vertices are the chambers of the hyperplane where the sign of $\sigma_k(C)$ gives the color of the vertex k in the subset. In agreement with the half-spaces defined above, if $\sigma_k(F) = +(-)$, the vertex of the subset labeled by k is blue (red). The faces that generate the random walk are colored edges. A blue edge connecting the vertices i and j is given by the face which has the sign sequence

$$\sigma_k(F_b) = \begin{cases} + & \text{if } k \in \{i, j\}, \\ 0 & \text{otherwise.} \end{cases}$$

In the same way, a red edge connecting i and j is given by

$$\sigma_k(F_r) = \begin{cases} - & \text{if } k \in \{i, j\}, \\ 0 & \text{otherwise.} \end{cases}$$

The probabilities assigned to these faces are uniform for the same color, and are given by $w(F_b) = p/m$ and $w(F_r) = q/m$.

3. Edge flipping

We are ready to study the stationary distribution and the rate of convergence to it for edge flipping. We first define the distance notion to be used. The *total variation distance* between μ and ν on the state space Ω is defined as

$$\|\mu - \nu\|_{TV} = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)| = \max_{S \subseteq \Omega} |\mu(S) - \nu(S)|. \tag{4}$$

In the context of Markov chains, we will use the following notation. For a Markov chain defined on Ω , let the probability assigned to $x \in \Omega$ be $P_z^t(x)$ after running the chain for t steps with the initial state z , and let it be $\pi(x)$ at the stationary distribution. A shorthand notation will be used for the total variation distance between P^t and π : $d_{TV}(t) = \max_{z \in \Omega} \|P_z^t - \pi\|_{TV}$. In case that it does not make difference from which state we initiated the Markov chain, we will denote its law after t steps by P^t . In fact, it can be shown that the maximum can only be achieved by point masses on color configurations. For the rate of convergence, we define the mixing time, $t_{\text{mix}}(\varepsilon) = \min\{t: d_{TV}(t) \leq \varepsilon\}$.

3.1. Distance to stationary distribution

The following theorem gives an upper bound for the distance to stationary distributions in terms of the eigenvalues given in the previous section.

Theorem 3. ([3, Theorem 5.1].) *Consider a central hyperplane arrangement with set of flats L . Let L^* be the set of all flats in L except the unique maximal one, and μ be the Möbius function*

of the lattice L . Then $d_{TV}(t) \leq - \sum_{X \in L^*} \mu(X, V(G)) \lambda_X^t$, where λ_X , the eigenvalue associated with flat X , is as defined in Theorem 2.

As mentioned in the introduction, this upper bound is not obtained by eigenvalue techniques, symmetric matrices, etc. This is due to a coupling argument.

We provide two examples where we identify the eigenvalues of the transition matrix and apply the theorem above.

Example 1. (The complete graph, K_n .) K_n is the graph with n vertices where every vertex is connected to every other vertex. Let X_k be a flat in L consisting of k vertices. Then $\mu(X_k, V(G)) = (-1)^{n-k}$. By Theorem 2, the eigenvalue corresponding to flat X_k is $\binom{k}{2} / \binom{n}{2}$ and the number of flats of size k is $\binom{n}{k}$. Hence,

$$\begin{aligned} d_{TV}(t) &\leq - \sum_{X \in L^*} \mu(X, V(G)) \lambda_X^t = - \sum_{k=2}^{n-1} \sum_{X_k} \mu(X_k, V(G)) \left(\frac{\binom{k}{2}}{\binom{n}{2}} \right)^t \\ &= - \sum_{k=2}^{n-1} (-1)^{n-k} \binom{n}{k} \left(\frac{\binom{k}{2}}{\binom{n}{2}} \right)^t \\ &\leq n \left(1 - \frac{2}{n} \right)^t + \sum_{i=2}^{n-2} (-1)^{i-1} \binom{n}{i} \left(1 - \frac{i}{n} \right)^{2t} + o(n) \\ &= \mathcal{O} \left(n \left(1 - \frac{2}{n} \right)^t \right), \end{aligned}$$

provided that $t \geq \frac{1}{2} n \log n$.

Example 2. (The complete bipartite graph, $K_{m,n}$.) $K_{m,n}$ is a graph whose vertex set is partitioned into two sets of sizes m and n , where every vertex of one set is connected to every vertex of the other set. Take a flat $X_{k,l} \in L$, where k and l denote the numbers of vertices from the set of size m and the set of size n respectively. The reader can verify using Theorem 2 that $\lambda_{X_{k,l}} = kl/mn$ with multiplicity $\binom{m}{k} \binom{n}{l}$. We can also show that $\mu(X_{k,l}, V(G)) = (-1)^{m+n-k-l}$. So, by Theorem 3,

$$\begin{aligned} d_{TV}(t) &\leq - \sum_{X \in L^*} \mu(X, V(G)) \lambda_X^t = n \left(\frac{(n-1)m}{nm} \right)^t + m \left(\frac{n(m-1)}{nm} \right)^t \\ &\quad - \sum_{k=1}^{m-1} \sum_{l=1}^{n-1} (-1)^{m+n-k-l} \binom{m}{k} \binom{n}{l} \left(\frac{kl}{mn} \right)^t \\ &= n \left(1 - \frac{1}{n} \right)^t + m \left(1 - \frac{1}{m} \right)^t \\ &\quad + \sum_{j=1}^{m-1} (-1)^j \binom{m}{j} \left(1 - \frac{j}{m} \right)^t \sum_{i=1}^{n-1} (-1)^i \binom{n}{i} \left(1 - \frac{i}{n} \right)^t. \end{aligned}$$

Suppose $n \geq m$. Then, for $t \geq n \log n$, we have $d_{TV}(t) = \mathcal{O}(n(1 - (1/n))^t)$.

A computationally tractable bound was obtained in [5] by a coarser version of the coupling argument in the proof of Theorem 3 in [3]:

$$d_{TV}(t) \leq - \sum_{X \in L^*} \mu(X, V(G)) \lambda_X^t \leq \sum_{\{M: M \text{ co-maximal in } L\}} \lambda_M^t, \tag{5}$$

where M is a *co-maximal* flat in L if $M < X$ implies X is maximal in L . In edge flipping, co-maximal flats are simply obtained by removing a vertex from the flat of all vertices, since any flat is a subset of vertices.

The precision of the alternating bound compared to (5) is discussed in [8] considering various examples of random hyperplane arrangements. In the examples that we study, the latter bound is as good as the former.

Now we can provide an upper bound for the rate of convergence of edge flipping in general connected graphs.

Theorem 4. Consider edge flipping on a connected graph G with n vertices and m edges, and let δ be the degree of the vertex with the minimum degree in G . For $c > 0$,

$$c \frac{m}{\delta} \leq t_{\text{mix}} \left(qe^{-2c} \right) \leq \frac{m}{\delta} (\log n + 2c - \log q).$$

Proof. Observe that the co-maximal flats obtained by removing any vertex with the minimum number of edges have the smallest eigenvalue among all co-maximal flats. Let M^* denote this flat. Then the eigenvalue indexed by M^* is

$$\lambda_{M^*} = \sum_{\text{supp } F \subseteq M^*} w(F) = \frac{m - \delta}{m} = 1 - \frac{\delta}{m}$$

by Theorem 2. By (5), we have

$$d_{TV}(t) \leq \sum_{\{M: M \text{ co-maximal in } L\}} \lambda_M^t \leq n \lambda_{M^*}^t \leq n \left(1 - \frac{\delta}{m} \right)^t \leq e^{-tm/\delta} \leq qe^{-2c}$$

if $t \geq (m/\delta)(\log n + 2c - \log q)$.

For the lower bound, let us label a vertex with the minimal degree by x and take its color to be blue in the initial configuration without loss of generality. Let R be the set of all configurations where x is colored red. We have $\pi(R) = q$ because whenever an edge with an endpoint x is chosen, it is colored blue with probability p and red with probability q . Now let us look at $P^t(R)$. The probability that the edge connecting x to the rest of the graph is not chosen before step t is

$$\left(1 - \frac{\delta}{m} \right)^t \geq \exp \left\{ -t \left(\frac{m}{\delta} + \frac{m^2}{\delta^2} \right) \right\}.$$

If we take $t \leq cm/\delta$, then we have $P^t(R) \leq q(1 - e^{-2c})$. Finally, using the second inequality in (4), we have $\|P^t - \pi\|_{TV} \geq |P^t(R) - \pi(R)| \geq qe^{-2c}$. \square

We note that this bound is not optimal in either direction, as shown in the following section.

3.2. Rate of convergence for regular graphs

A graph G is called *regular* if every vertex of G has the same degree. It is called k -*regular* if the degree of the vertices is k . We show the following bounds on the mixing time of edge flipping for regular graphs, which are independent of the degree of the vertices.

Theorem 5. *Let G be a connected, regular graph with n vertices. Then, for $c > 0$, the mixing time of edge flipping satisfies $\frac{1}{4}n \log n - \mathcal{O}(n) \leq t_{\text{mix}}(e^{-c}) \leq \frac{1}{2}n \log n + \mathcal{O}(n)$.*

The smallest-degree regular graph on n vertices is the cycle C_n , and the largest-degree regular graph on n vertices is K_n . We remark that the stationary distributions of edge flipping for these graphs are studied in [6, 7].

The derivation of these bounds shows that, as long as the number of vertices of the same degree is of order n , the principal terms in the bounds above remain the same.

Proof of the upper bound in Theorem 5. Suppose the degree of the vertices of G is equal to k . Let M denote a co-maximal flat. Then the corresponding eigenvalue is

$$\lambda_M = \sum_{\text{supp } F \subseteq M} w(F) = \frac{(kn/2) - k}{kn/2} = 1 - \frac{2}{n}$$

by Theorem 2. The bound (5) gives

$$\|P_C^t - \pi\|_{\text{TV}} \leq n \left(1 - \frac{2}{n}\right)^t \leq e^{-c}$$

if we take $t \geq n \log n/2 + cn/2$. □

For the proof of the lower bound in Theorem 5 we use the Wilson’s method: [18] showed that an eigenvector is a good candidate for a test statistic as the variance of the associated eigenfunction can be estimated inductively from the transition probabilities of the Markov chain. Assuming the second-order estimate for the eigenfunction, a lower bound can be obtained as follows.

Lemma 1. ([11, Theorem 13.5].) *Let X_t be an irreducible, aperiodic, time-homogenous Markov chain with state space Ω . Let Φ be an eigenfunction associated with eigenvalue $\lambda > \frac{1}{2}$. If, for all $x \in \Omega$, $\mathbb{E}((\Phi(X_1) - \Phi(x))^2 | X_0 = x) \leq R$ for some $R > 0$, then*

$$t_{\text{mix}}(\varepsilon) \geq \frac{1}{2 \log(1/\lambda)} \left(\log \left(\frac{(1 - \lambda)\Phi(x)^2}{2R} \right) + \log \left(\frac{1 - \varepsilon}{\varepsilon} \right) \right)$$

for any $x \in \Omega$.

The maximum eigenvalue is 1 with right eigenfunction $\phi_0(C) = 1$ for all $C \in \mathcal{C}$. The eigenvectors of eigenvalues corresponding to co-maximal flats in hyperplane arrangements are identified by Pike as follows.

Theorem 6. ([14, Theorem 3.1.1].) *For each $i \in \{1, \dots, n\}$, the Markov chain defined on the chambers by the transition probabilities in (2) has the right eigenfunction*

$$\phi_i(C) = \begin{cases} - \sum_{\substack{F \in \mathcal{F} \\ \sigma_i(F) = +}} w(F), & \sigma_i(C) = -, \\ \sum_{\substack{F \in \mathcal{F} \\ \sigma_i(F) = -}} w(F), & \sigma_i(C) = +, \end{cases}$$

corresponding to eigenvalue $\lambda_i = \sum_{\sigma_1(F)=0} w(F)$. Moreover, $\phi_0, \phi_1, \dots, \phi_n$ are linearly independent, where ϕ_0 is the eigenvector for the largest eigenvalue 1.

In our problem, the eigenvectors have a simpler expression as

$$\phi_i(C) = \begin{cases} -p, & \sigma_i(C) = -, \\ q, & \sigma_i(C) = +, \end{cases} \tag{6}$$

for $i = 1, \dots, n$. Each i represents a vertex, and we can read off the color of i in a chamber by the eigenfunction ϕ_i . These eigenfunctions are not quite distinguishing on their own. In fact, Wilson’s lemma applied to any ϕ_i gives a lower bound of order n only. Nevertheless, if an eigenvalue corresponding to a co-maximal flat has algebraic multiplicity larger than one, linear combinations of eigenvectors can be used in Wilson’s lemma. In the case of regular graphs, where each vertex has the same degree, we have the second largest eigenvalue

$$\lambda = \sum_{\sigma_1(F)=0} w(F) = \frac{kn/2 - k}{kn/2} = 1 - \frac{2}{n}$$

associated with the co-maximal flat obtained by removing the vertex indexed by 1. Of course, it has multiplicity n considering all other co-maximal flats and equality of degrees. Therefore, from Theorem 6, the eigenspace for the second largest eigenvalue is spanned by eigenvectors (6) for $i = 1, \dots, n$. If we consider the linear combination $\Phi = \sum_{i=1}^n \phi_i$, it is easy to see that $\Phi(C) = \Phi(C')$ if and only if they have the same number of blue vertices. If the color configuration C has k blue vertices, then $\Phi(C) = k - pn$. So, Φ is a statistic that counts the number of blue vertices up to an additive term which guarantees $\mathbb{E}(\Phi(\pi)) = 0$, where the expectation is taken with respect to the stationary distribution π , i.e. $\mathbb{E}(\Phi(X_t | X_0 = C)) = \lambda^t \Phi(C) \rightarrow 0$. We can use this statistic to obtain the lower bound by Wilson’s method. Since, at each step, at most two vertices can change color, we have

$$\mathbb{E}[(\Phi(X_1) - \Phi(C))^2 | X_0 = C] \leq 4, \tag{7}$$

so that the condition in Lemma 1 is satisfied with $R = 4$.

Proof of the lower bound in Theorem 5. For the lower bound, since the second eigenvalue has multiplicity n given that the graph is regular, the sum of eigenvectors is an eigenvector of the second largest eigenvalues above. Take the initial state to be the blue monochromatic graph, for which $\Phi(C) = n - pn = qn$. Then, using Lemma (1),

$$t_{\text{mix}}(\varepsilon) \geq \frac{n}{4} \left(\log \left(\frac{q^2 n^2}{4n} \right) + \log \left(\frac{1 - \varepsilon}{\varepsilon} \right) \right) \geq \frac{1}{4} n \log n - Cn$$

for some constant C depending only on p and ε . We take $\varepsilon = e^{-c}$ to conclude the proof. \square

We also note, with thanks to the anonymous referee who pointed this out, that the construction of eigenfunctions of edge flipping is analogous to the construction of eigenfunctions for the simple random walk on the hypercube. This also justifies the lower bound on the leading order of $\frac{1}{4} n \log n$.

3.3. Exact rate of convergence for complete graphs

In Section 3.2 we established a lower bound of $\frac{1}{4}n \log n$ for the mixing time of edge flipping in regular graphs with n vertices. However, the upper bound obtained from eigenvalues differs by a factor of two. So we can only give an interval for the exact rate of convergence. Yet some earlier studies on similar processes suggest that the lower bound captures the correct rate of convergence; see [12], for instance. We can verify that in the case of complete graphs. First, we formally define what we mean by the ‘exact rate’.

Definition 1. A sequence of Markov chains shows a cutoff at the *mixing time* $\{t_n\}$ with a *window of size* $\{w_n\}$ if

- (i) $\lim_{n \rightarrow \infty} w_n/t_n = 0,$
- (ii) $\lim_{\gamma \rightarrow -\infty} \liminf_{n \rightarrow \infty} d_{\text{TV}}(t_n + \gamma w_n) = 1,$
- (iii) $\lim_{\gamma \rightarrow \infty} \limsup_{n \rightarrow \infty} d_{\text{TV}}(t_n + \gamma w_n) = 0.$

We will use a coupling argument to show that edge flipping in complete graphs shows a cutoff at time $\frac{1}{4}n \log n$ with a window of size n . The coupling time of two copies of a process is defined as $\tau := \min\{t: X_t = Y_t\}$. Writing $d(\cdot)$ for $d_{\text{TV}}(\cdot)$, the coupling time is related to the mixing time by

$$d(t) \leq \max_{x,y \in \Omega} \mathbb{P}(\tau > t \mid X_0 = x, Y_0 = y). \tag{8}$$

See, for example, [11, Theorem 5.4].

Proof of Theorem 1. Consider the linear combination of eigenvectors denoted by Φ in Section 3.2. We already showed that it counts the number of blue vertices. Now, if we view edge flipping on K_n as a random walk on the hypercube $(\mathbb{Z}/2\mathbb{Z})^n$ where at each step two coordinates are replaced either by 1s with probability p or by 0s with probability q , then Φ is just the *Hamming weight*, the sum of the 0–1 coordinates. See [11, Section 2.3] for the definition of a hypercube random walk. An $n \log n$ upper bound is obtained for a lazy random walk on the hypercube by a coupling and a strong stationary time argument, which can be found in [11, Sections 5.3 and 6.4], respectively. This bound is later improved by monotone coupling in Section 18.2 of the same book following the observation that the problem can be reduced to convergence of Hamming weights, which can be considered as a lazy Ehrenfest urn problem. The same reduction argument is valid in the case of edge flipping in complete graphs, as follows. We claim that if the initial configuration is monochromatic, then

$$\|P^t - \pi\|_{\text{TV}} = \|\Phi(X_t) - \pi_\Phi\|_{\text{TV}}, \tag{9}$$

where X_t is the chamber at step t , and π_Φ is the stationary distribution for the number of blue vertices. To argue for this, let us denote the set of chambers with k blue vertices by $C_k = \{C \in \mathcal{C} : \Phi(C) = k\}$. Starting from a monochromatic configuration, X_t is equally likely to be any of the chambers with the same number of blue vertices by symmetry. Thus, we have

$$\sum_{C \in C_k} |P(X_t = C) - \pi(C)| = \left| \sum_{C \in C_k} P(X_t = C) - \pi(C) \right| = |\mathbb{P}(\Phi(X_t) = k) - \pi_\Phi(k)|.$$

The equality in (9) shows that we can study the convergence of $\Phi(X_t)$ to its stationary distribution to obtain an upper bound for the mixing time of the edge flipping on K_n . We

define a coupling with three stages to study the mixing time. Let X_t and Y_t be two edge flipping processes where the initial states are monochromatic of opposite colors. We denote the number of blue vertices in each process by $W_t = \Phi(X_t)$, $Z_t = \Phi(Y_t)$, and take $W_0 = 0$ and $Z_0 = n$. From now on, we will only deal with W_t and Z_t . Let us first define $\Delta_t := Z_t - W_t$ for $t = 0, 1, \dots$

First stage: Coupon collector’s problem For the first stage of the coupling $(W_t, Z_t)_{t=0}^\infty$ we define the stopping time τ_1 as $\tau_1 = \min_t \{ \Delta_t \leq \lceil \sqrt{n} \rceil \}$; then we define the rule of coupling up to this time. If $t < \tau_1$, the two chains move identically in the sense that we choose the same two vertices in the underlying graphs and color them the same.

The first part suggests that the process up to time τ_1 is twice as fast as the coupon collector’s problem where you stop collecting more coupons if the remaining number of coupons is $\lceil \sqrt{n} \rceil$. Let us define a process $\tilde{\Delta}_t$ for all t with transition probabilities

$$\tilde{\Delta}_{t+1} = \tilde{\Delta}_t + \begin{cases} 0 & \text{with probability } \binom{n-\tilde{\Delta}_t}{2} / \binom{n}{2}, \\ -1 & \text{with probability } \tilde{\Delta}_t(n - \tilde{\Delta}_t) / \binom{n}{2}, \\ -2 & \text{with probability } \binom{\tilde{\Delta}_t}{2} / \binom{n}{2} \end{cases}$$

and $\tilde{\Delta}_0 = n$. Note that $\tilde{\Delta}_t = \Delta_t$ for $t \leq \tau_1$. Calculating the conditional expectation, we have $\mathbb{E}[\tilde{\Delta}_{t+1} | \tilde{\Delta}_t] = (1 - (2/n))\tilde{\Delta}_t$. Therefore,

$$\mathbb{E}[\tilde{\Delta}_t] = \left(1 - \frac{2}{n}\right)^t \tilde{\Delta}_0 \leq ne^{-2t/n}.$$

By Markov’s inequality, $\mathbb{P}(\tau_1 > s) = \mathbb{P}[\tilde{\Delta}_s > \sqrt{n}] \leq \sqrt{ne^{-2s/n}}$. If we take $s = \frac{1}{4}n \log n + \gamma_1 n$, we obtain

$$\mathbb{P}\left(\tau_1 > \frac{1}{4}n \log n + \gamma_1 n\right) < e^{-\gamma_1}. \tag{10}$$

Second stage: Coupling with lazy simple random walks Note that if we continued with the first coupling until the two chains meet, this expression would give an upper bound of order $\frac{1}{2}n \log n$, which is the same as we obtained from eigenvalues. To improve this rate, whenever $t \geq \tau_1$, we let the chains move according to the second rule where we choose edges independently for two graphs but color them with the same color. We will show that it only needs $\mathcal{O}(n)$ steps to hit zero with positive probability. Given that $t \geq \tau_1$,

$$\Delta_{t+1} = \Delta_t + \begin{cases} -2 & \text{with probability } \binom{n-W_t}{2} \binom{Z_t}{2} / \binom{n}{2}^2, \\ -1 & \text{with probability } \left[\binom{n-W_t}{2} Z_t(n - Z_t) + W_t(n - W_t) \binom{Z_t}{2} \right] / \binom{n}{2}^2, \\ 1 & \text{with probability } \left[W_t(n - W_t) \binom{n-Z_t}{2} + \binom{W_t}{2} Z_t(n - Z_t) \right] / \binom{n}{2}^2, \\ 2 & \text{with probability } \binom{W_t}{2} \binom{n-Z_t}{2} / \binom{n}{2}^2, \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

We need to be more precise and show that the terms in the difference Δ_t are concentrated around pn , recalling that p was the probability of coloring a chosen edge blue.

Lemma 2. *If $t > \frac{1}{4}n \log n$, then $\mathbb{P}(|V_t - pn| > n^{1/2+\alpha}) \leq 2n^{-2\alpha}(1 + o(1))$ for any $\alpha > 0$, where V_t can be replaced by W_t or Z_t , which are defined in the beginning of the proof.*

Proof. It is not difficult to see that $\mathbb{E}[W_t] \rightarrow pn$ as $t \rightarrow \infty$. To show the concentration result, we will use a second-order estimate. First, consider

$$W_{t+1} = W_t + \begin{cases} -2 & \text{with probability } q\binom{W_t}{2}/\binom{n}{2}, \\ -1 & \text{with probability } qW_t(n - W_t)/\binom{n}{2}, \\ 1 & \text{with probability } pW_t(n - W_t)/\binom{n}{2}, \\ 2 & \text{with probability } p\binom{n - W_t}{2}/\binom{n}{2}, \\ 0 & \text{otherwise,} \end{cases} \tag{12}$$

where we take $\binom{z}{n} = 0$ unless z is a positive integer. Let

$$\mathcal{F}_t = \sigma(W_0, W_1, \dots, W_t) \tag{13}$$

be the σ -field generated by W_0, W_1, \dots, W_t . Computation gives $\mathbb{E}[W_{t+1} | \mathcal{F}_t] = 2p + (1 - (2/n))W_t$. We can evaluate this sum recursively to obtain

$$\mathbb{E}[W_t] = pn - \left(1 - \frac{2}{n}\right)^{t+1} \quad pn > pn - p\sqrt{n}, \tag{14}$$

provided that $t > \frac{1}{4}n \log n$ and $W_0 = 0$ as assumed earlier.

Secondly, we bound the variance of W_t by applying a crucial step in the proof of Wilson’s Lemma 1 (see [11, Section 13.2]), which gives

$$\text{Var}(W_t) \leq \frac{R}{1 - \lambda} = \frac{4}{1 - (1 - (2/n))} = 2n, \tag{15}$$

by (7). Then, by (14) and using (15) in Chebyshev’s inequality,

$$\mathbb{P}(|W_t - pn| \geq (p + r\sqrt{2})\sqrt{n}) \leq \frac{1}{r^2}$$

for $r > 0$. This proves the result for $V_t = W_t$. The proof for $V_t = Z_t$ follows exactly the same line of argument. By symmetric replacement of parameters, we have

$$\mathbb{E}[Z_t] = pn - \left(1 - \frac{2}{n}\right)^{t+1} \quad pn + \left(1 - \frac{2}{n}\right)^t Z_0 < pn + q\sqrt{n},$$

where we take $Z_0 = n$. Since the bound in (7) is true for all initial states, we have the same variance bound. The rest follows from Chebyshev’s inequality. \square

Next, we present a maximal inequality for W_t in the second stage of the coupling.

Lemma 3. *If $t \geq \frac{1}{4}n \log n$ and k is a positive real number independent of n , then*

$$\mathbb{P}\left(\max_{\tau_1 \leq t \leq \tau_1 + kn} |V_t - pn| > n^{1/2+\beta}\right) \leq 4n^{-\alpha}$$

for any $\beta > 0$ and $\alpha \in (0, \beta)$, where V_t can be replaced by W_t or Z_t .

Proof. Since W_t and Z_t are symmetric cases as in Lemma 2, we only prove the result for W_t . By Doob–Meyer decomposition, we construct a martingale from $\{W_{\tau_1+s}\}_{s \geq 0}$ with respect to the σ -fields: $\mathcal{G}_s := \mathcal{F}_{\tau_1+s}$ for all $s = 0, 1, 2, \dots$, where \mathcal{F}_t is as defined in (13) for all $t = 0, 1, 2, \dots$. Since $\mathbb{P}(\tau_1 < \infty) = 1$ by (10), $M_1 = W_{\tau_1+1} - W_{\tau_1}$ is well-defined. For $s \geq 2$, we let

$$M_s = W_{\tau_1+s} + \frac{2}{n} \sum_{i=1}^{s-1} W_{\tau_1+i} - W_{\tau_1} - 2sp. \tag{16}$$

From (12) and (14), we can prove the martingale property: $\mathbb{E}[M_{s+1} | \mathcal{G}_s] = M_s$. Observe that the transition probabilities of the martingale depend on W_{τ_1} . We will fix W_{τ_1} later in the proof.

In order to bound the variance of the martingale, we consider the variance of the differences. The computations give

$$\mathbb{E}[(M_{s+1} - M_s)^2] = C_1(p) \frac{\mathbb{E}[W_{\tau_1+s}^2]}{n^2} + C_2(p) \frac{\mathbb{E}[W_{\tau_1+s}]}{n} + C_3(p) \leq C(p),$$

where $C(p)$ and the other constants in the expression are independent of n and s . Therefore, by the orthogonality of martingale differences $\mathbb{E}[M_s^2] = \text{Var}(M_s) \leq C(p)s$. Then, by the Doob–Kolmogorov maximal inequality [10, Theorem 2 in Section 7.8],

$$\mathbb{P}\left(\max_{1 \leq t \leq kn} |M_t| \geq a\right) \leq \frac{\mathbb{E}[M_{kn}]}{a^2} = \frac{C(p)kn}{a^2} \tag{17}$$

for $a > 0$.

Next, we show that, for a of order $n^{1/2+\beta}$ and large enough n , $|W_{\tau_1+s} - pn| \geq a$ implies $|M_s| \geq n^{-\beta/2}a$ (not optimal) with high probability. Let A be the event that

$$|W_{\tau_1} - pn| \leq n^{1/2+\beta/2}, \tag{18}$$

which has probability greater than $1 - 4n^{-\beta}$ by Lemma 2. The following argument is conditioned on A . Let us fix an outcome $W_{\tau_1} = W^*$ in A . Note that both the stopping time τ_1 and W_{τ_1} are determined. We take $n > 2^{4/\beta}$. Suppose

$$|W_{\tau_1+s} - pn| > n^{1/2+\beta}. \tag{19}$$

From (18) and (19), we get $|W_{\tau_1+s} - W_{\tau_1}| > \frac{1}{2}n^{1/2+\beta}$. There are two cases for (19): $W_{\tau_1+s} > pn + n^{1/2+\beta}$ or $W_{\tau_1+s} < pn - n^{1/2+\beta}$. Let us assume the former. It will be apparent below that the other case can be treated similarly. From (16), we have one of

$$M_s > \frac{1}{4}n^{1/2+\beta}, \quad \frac{2}{n} \sum_{i=1}^{s-1} W_{\tau_1+i} - 2sp \leq -\frac{1}{4}n^{1/2+\beta}.$$

If the latter is the case, by (18) and (19), there must exist a time s' such that

$$s' = \max \left\{ 0 \leq t < s : W_{\tau_1+t} \leq pn + 2n^{1/2+\beta/2} \right\}.$$

Then observe that

$$\frac{2}{n} \sum_{i=1}^{s'-1} W_{\tau_1+i} - 2s'p < \frac{2}{n} \sum_{i=1}^{s-1} W_{\tau_1+i} - 2sp,$$

since $pn < W_{\tau_1+t}$ for $s' \leq t < s$. Therefore,

$$M_{s'} = W_{\tau_1+s'} - W_{\tau_1} + \frac{2}{n} \sum_{i=1}^{s'-1} W_{\tau_1+i} - 2s'p \leq 3n^{1/2+\beta/2} - \frac{1}{4}n^{1/2+\beta} \leq -n^{1/2+\beta/2}.$$

In either case, we have $\max_{1 \leq t \leq s} |M_t| \geq n^{1/2+\beta/2}$. Therefore, by (17),

$$\begin{aligned} \mathbb{P}\left(\max_{\tau_1 \leq t \leq \tau_1+s} |W_t - pn| > n^{1/2+\beta} \mid W_{\tau_1} = W^*\right) &< \mathbb{P}\left(\max_{1 \leq t \leq s} |M_t| \geq n^{1/2+\beta/2} \mid W_{\tau_1} = W^*\right) \\ &< C(p)kn^{-\beta}. \end{aligned}$$

Now let us take $s = kn$. Since the established upper bound is uniform for W^* in A , it follows via a conditioning argument that

$$\begin{aligned} \mathbb{P}\left(\max_{\tau_1 \leq t \leq \tau_1+kn} |W_t - pn| > n^{1/2+\beta}\right) &< \mathbb{P}\left(\max_{1 \leq t \leq kn} |M_t| \geq n^{1/2+\beta/2} \mid A\right)\mathbb{P}(A) + \mathbb{P}(A^c) \\ &< C(p)kn^{-\beta}\mathbb{P}(A) + \mathbb{P}(A^c) \\ &\leq C(p)kn^{-\beta}(1 - 4n^{-\beta}) + 4n^{-\beta} < 4n^{-\alpha}, \end{aligned}$$

where n is chosen large enough to satisfy the last inequality. □

We return to (11). We use the lemmas above to bound the transition probabilities of Δ_t uniformly. We first need the following corollary.

Corollary 1. *Let $t \in [\tau_1, \tau_1 + kn]$ for some $k > 0$ independent of n . For every $\varepsilon > 0$, there exist $\alpha \in (0, \frac{1}{2})$ and $\beta \in (\alpha, \frac{1}{2})$ such that*

$$\begin{aligned} p_t &:= \frac{\sqrt{2}W_t}{n} \in [p - \varepsilon, p + \varepsilon], \\ \delta &:= \sqrt{2}\frac{Z_t - W_t}{n} = \mathcal{O}(n^{\beta-1/2}), \end{aligned} \tag{20}$$

with probability $1 - 8n^{-\alpha}$.

Proof. Both statements follow from Lemma 3. For the latter, since W_t and Z_t are correlated by coupling, we use the union bound. □

Therefore, under the assumptions of Corollary 1 we can rewrite (11) as

$$\Delta_{t+1} = \Delta_t + \begin{cases} -2 & \text{with probability } p_t^2 q_t^2 + 2p_t q_t^2 \delta + o(n^{2\beta-1}), \\ -1 & \text{with probability } 2p_t q_t (p_t^2 + q_t^2) + q_t((q_t - p_t)^2 + p_t)\delta + o(n^{2\beta-1}), \\ 1 & \text{with probability } 2p_t q_t (p_t^2 + q_t^2) - p_t((p_t - q_t)^2 + q_t)\delta + o(n^{2\beta-1}), \\ 2 & \text{with probability } p_t^2 q_t^2 - 2p_t^2 q_t \delta + o(n^{2\beta-1}), \\ 0 & \text{otherwise.} \end{cases}$$

Then, examining the probabilities, we see that it is more likely for $\{\Delta_t\}_{t>\tau_1}$ to go to the left if $\Delta_t > 0$. In fact,

$$\begin{aligned} \mathbb{P}(\Delta_{t+1} - \Delta_t = -2) &\geq p_t^2 q_t^2 \geq \mathbb{P}(\Delta_{t+1} - \Delta_t = 2), \\ \mathbb{P}(\Delta_{t+1} - \Delta_t = -1) &\geq 2p_t q_t (p_t^2 + q_t^2) \geq \mathbb{P}(\Delta_{t+1} - \Delta_t = 1), \end{aligned} \tag{21}$$

provided that $Z_t \geq W_t$. Thus, we can compare it to the following random walk, which is equally likely to go in either direction at any time:

$$S_{t+1} = S_t + \begin{cases} -2 & \text{with probability } p_t^2 q_t^2, \\ -1 & \text{with probability } 2p_t q_t (p_t^2 + q_t^2), \\ 0 & \text{with probability } (p_t^2 + q_t^2)^2 + 2p_t^2 q_t^2, \\ 1 & \text{with probability } 2p_t q_t (p_t^2 + q_t^2), \\ 2 & \text{with probability } p_t^2 q_t^2. \end{cases}$$

Let us take $S_0 = \Delta_{\tau_1} = \lceil \sqrt{n} \rceil$. We define the stopping times $\tau_2 = \min_t \{ \Delta_{\tau_1+t} \in I \}$ and $\tau_S = \min_t \{ S_t \in I \}$, where $I = \{-1, 0, 1\}$. Comparing the transition probabilities of the two random walks by (21), we have $\Delta_{\tau_1+t} \leq S_t$ since $Z_t \geq W_t$. Therefore, we can find a coupling for two processes such that

$$\tau_2 \leq \tau_S. \tag{22}$$

Furthermore, observe that $S_{t+1} - S_t$ is a sum of independent and identically distributed random variables which take values -1 or 1 with probability $p_t q_t$ and 0 with probability $p_t^2 + q_t^2$. But then $S_t = S'_{2t}$, where S'_t is none other than a lazy simple random walk on the integers, so we may use the results on the lazy simple random walk to bound $\mathbb{P}(\tau_S > s)$, such as [11, Corollary 2.28]. The only difficulty is that S_t is not time-homogeneous as the parameter p_t varies, which is dealt with next.

Given $r \in [0, 1]$, we let $\tau_S(r) = \min_t \{ S_t(r) \in I \}$ where $S_t(r)$ is the lazy simple random walk with holding probability $r^2 + (1-r)^2$. Considering the bound in (20), we define $p^* = (p - \varepsilon)\mathbf{1}_{\{p \leq 1/2\}} + (p + \varepsilon)\mathbf{1}_{\{p > 1/2\}}$. The choice of p^* ensures that $S_t(p^*)$ has the maximal holding probability among all values in $[p - \varepsilon, p + \varepsilon]$.

Lemma 4. $\mathbb{P}(\tau_S(p^*) > s) \geq \mathbb{P}(\tau_S > s)$.

Proof. Let us take $q^* = 1 - p^*$. We define a Bernoulli random variable for each t as follows:

$$\xi_t = \begin{cases} 0 & \text{with probability } [1 - (p^{*2} + q^{*2})]/[1 - (p_t^2 + q_t^2)], \\ 1 & \text{otherwise.} \end{cases}$$

ξ is well-defined by the choice of p^* . Now observe that $S_t(p^*)$ has the same distribution as the chain which moves according to the law of S_t (which is the same as $S_t(p_t)$) if $\xi_t = 1$, and stays at the same state if $\xi_t = 0$. Then, it is obvious that we can find a coupling satisfying $\tau_S(p^*) \geq \tau_S$. The result follows. \square

Finally, by Lemmas 3 and 4, (22), and modifying the numbers in [11, Corollary 2.28] accordingly, we have

$$\mathbb{P}(\tau_2 > s) \leq 4n^{-2\alpha} + \frac{4\sqrt{n}}{\sqrt{2p^*(1-p^*)}s}.$$

So, if we take $s = \gamma_2 n$,

$$\mathbb{P}(\tau_2 > \gamma_2 n) \leq 4n^{-2\alpha} + \frac{4}{\sqrt{2p^*(1-p^*)\gamma_2}} \leq \frac{4 \max\{p^{-1/2}, (1-p)^{-1/2}\}}{\sqrt{\gamma_2}} \tag{23}$$

by choosing ε small enough, say $\varepsilon = \min\{p, 1-p\}/100$, in (20).

Third stage: Number of sign changes For the last stage of the coupling, let us define $\tau_3 = \min_t\{\Delta_{\tau_1+\tau_2+t} = 0\}$, which is to account for the case that $Z_{\tau_2} - W_{\tau_2} = 1$ and Δ_t does not hit zero in the next step. By symmetry, we see that it is more likely for $\{\Delta_t\}_{t>\tau_1}$ to go to the right if $\Delta_t < 0$. So, we can again bound the stopping time by the time associated with S_t . In fact, for the simple random walk, we have the probability that ‘the number of sign changes is equal to k up to $t = 2n + 1$ ’ is equal to the probability that ‘the walk is at position $2k + 1$ at $t = 2n + 1$ ’; see [9, Theorem 1 in Chapter III.6]. This allows us to use the normal approximation for the sign changes [9, Theorem 2 in Chapter III.6]. Let F_Z be the distribution function of the standard normal distribution. Let us use $S_t(p^*)$ again, the laziest simple random walk in the range. We have

$$\mathbb{P}\left(|t \in \{1, 2, \dots, K(p^*)n\} : S_{\tau_1+\tau_2+t}(p^*) \in I \leq \frac{\sqrt{n}}{\gamma_3}\right) \leq 2F_Z\left(\frac{1}{\gamma_3}\right) - 1 \leq \frac{1}{\sqrt{2\pi}\gamma_3},$$

where $K(p) = 1 + [(p^2 + q^2)/2pq]$. Since

$$\mathbb{P}(|t \in \{1, 2, \dots, k\} : S_{\tau_1+\tau_2+t} \in I \leq x) \leq \mathbb{P}(|t \in \{1, 2, \dots, k\} : S_{\tau_1+\tau_2+t}(p^*) \in I \leq x)$$

for all k and x , Δ_t visits I on an order larger than \sqrt{n} times a probability bounded away from zero. At every visit to I , it has a positive probability to hit zero at the next stage. Let $\kappa = \mathbb{P}(S_{t+1}(p^*) = 0 \mid S_t = -1 \text{ or } 1)$. Thus,

$$\mathbb{P}(\tau_3 > K(p^*)n) \leq (1 - \kappa)^{\sqrt{n}/\gamma_3} + \frac{1}{\sqrt{2\pi}\gamma_3} < \gamma_3^{-1} \tag{24}$$

for any $\gamma_3 > 0$ independent of n for large n . Let us take $\gamma_3 > K(p^*)$ for convenience. Finally, let $\tau = \tau_1 + \tau_2 + \tau_3$, which is the stopping time, $\alpha > 0$, and $\gamma = \gamma_1 + \gamma_2 + \gamma_3$. By (10), (23), (24), and using the strong Markov property iteratively,

$$\mathbb{P}(\tau > \frac{1}{4}n \log n + \gamma n) < 1 - (1 - e^{-\gamma_1})(1 - 4 \max\{p^{-1/2}, (1-p)^{-1/2}\}\gamma_2^{-1/2})(1 - \gamma_3^{-1}).$$

Since we chose the opposite monochromatic configurations for the different initial states of the coupled chains, among all choices for the starting pair chambers the coupling time above is clearly the maximum, so is the probability in (8). Therefore,

$$d\left(\frac{1}{4}n \log n + \gamma n\right) \leq \frac{C}{\sqrt{\gamma}}$$

by choosing γ_1, γ_2 , and γ_3 large enough. Combining with the lower bound in Theorem 5, we have $t_n = \frac{1}{4}n \log n$ and $w_n = n$ in Definition 1. □

Acknowledgements

The second and the third authors would like to thank Jason Fulman for the suggestion of this subject of study. We would also like to thank the anonymous referees whose suggestions and corrections improved the paper significantly.

Funding information

The second author is supported partially by BAP grant 20B06P1.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

- [1] ATHANASIADIS, C. A. AND DIACONIS, P. (2010). Functions of random walks on hyperplane arrangements. *Adv. Appl. Math.* **45**, 410–437.
- [2] BASU, R., HERMON, J. AND PERES, Y. (2017). Characterization of cutoff for reversible Markov chains. *Ann. Prob.* **45**, 1448–1487.
- [3] BIDIGARE, P., HANLON, P. AND ROCKMORE, D. (1999). A combinatorial description of the spectrum for the Tsetlin library and its generalization to hyperplane arrangements. *Duke Math. J.* **99**, 135–174.
- [4] BROWN, K. (2000). Semigroups, rings, and Markov chains. *J. Theoret. Prob.* **13**, 871–928.
- [5] BROWN, K. AND DIACONIS, P. (1998). Random walks and hyperplane arrangements. *Ann. Prob.* **26**, 1813–1854.
- [6] BUTLER, S., CHUNG, F., CUMMINGS, J. AND GRAHAM, R. (2015). Edge flipping in the complete graph. *Adv. Appl. Math.* **69**, 46–64.
- [7] CHUNG, F. AND GRAHAM, R. (2012). Edge flipping in graphs. *Adv. Appl. Math.* **48**, 37–63.
- [8] CHUNG, F. AND HUGHES, J. (2014). A note on an alternating upper bound for random walks on semigroups. *Discrete Appl. Math.* **176**, 24–29.
- [9] FELLER, W. (1968). *Introduction to Probability and its Applications*, Vol. 17. Wiley, Chichester.
- [10] GRIMMETT, G. AND STIRZAKER, D. (2020). *Probability and Random Processes*. Oxford University Press.
- [11] LEVIN, D. A., PERES, Y. AND WILMER, E. L. (2009). *Markov Chains and Mixing Time*. AMS, Providence, RI.
- [12] NESTORIDI, E. (2017). A non-local random walk on the hypercube. *Adv. Appl. Prob.* **49**, 1288–1299.
- [13] NESTORIDI, E. (2019). Optimal strong stationary times for random walks on the chambers of a hyperplane arrangement. *Prob. Theory Relat. Fields* **174**, 929–943.
- [14] PIKE, J. (2013). Eigenfunctions for random walks on hyperplane arrangements. PhD thesis, University of Southern California.
- [15] SALEZ, J. (2021). Cutoff for non-negatively curved Markov chains. Preprint, [arXiv:2102.05597](https://arxiv.org/abs/2102.05597).
- [16] SALIOLA, F. (2012). Eigenvectors for a random walk on a left-regular band. *Adv. Appl. Math.* **48**, 306–311.
- [17] STANLEY, R. P. (2007). An introduction to hyperplane arrangements. *Geom. Combinatorics* **13**, 389–496.
- [18] WILSON, D. B. (2004). Mixing times of lozenge tiling and card shuffling Markov chains. *Ann. Appl. Prob.* **14**, 274–325.