Advances in Applied Mathematics 48 (2012) 37–63

Contents lists available at ScienceDirect

Advances in Applied Mathematics

www.elsevier.com/locate/yaama

# Edge flipping in graphs

Fan Chung [1], Ron Graham *

*Department of Mathematics, University of California, San Diego, La Jolla, CA 92037, United States*

ARTICLE INFO

*Article history:*
Received 30 August 2010
Accepted 30 April 2011
Available online 22 June 2011

*MSC:*
05C81

*Keywords:*
Random walk
Left-regular band
Tsetlin library
Edge flipping
Graphs

ABSTRACT

In this paper we investigate certain random processes on graphs which are related to the so-called Tsetlin library random walk as well as to some variants of a classical voter model. A specific example of what we study is the following. Suppose we begin with some finite graph $G$ in which each vertex of $G$ is initially arbitrarily colored red or blue. At each step in our process, we select a random edge of $G$ and (re-)color both its endpoints *blue* with probability $p$, or *red* with probability $q = 1 - p$. This "edge flipping" process generates a random walk on the set of all possible color patterns on $G$. We show that the eigenvalues for this random walk can be naturally indexed by subsets of the vertices of $G$. For example, in the uniform case (where $p = \frac{1}{2}$), for each subset $T$ of vertices of $G$ there is an eigenvalue $\lambda_T$ (with multiplicity 1) which is equal to the number of edges in the subgraph induced by $T$ divided by the number of edges of $G$.

We also carry out a fairly detailed analysis of the stationary distribution of this process for several simple classes graphs, such as paths and cycles. Even for these graphs, the asymptotic behavior can be rather complex.



## 1. Introduction

Let us consider the following coloring process (or game) on a given graph $G$, first suggested to us by Persi Diaconis (also see [2]). The vertex set $V$ of $G$ will represent the players. The edge set $E$ of $G$ consists of pairs of players who can interact with one another. We will always assume $G$ has no isolated vertices. Initially each vertex of $G$ is colored either blue or red. At each step, an edge

* Corresponding author.
*E-mail address:* graham@ucsd.edu (R. Graham).

[1] Research supported in part by ONR MURI N000140810747, and FA9550-09-1-0900.

0196-8858/$ – see front matter 
doi:10.1016/j.aam.2011.06.002

is uniformly chosen at random and the two players associated with its endpoints will recolor their corresponding vertices. However, the two players tend to influence each other and are likely to choose the same color. Although many possible more complicated formulations are possible for the rules of the game, here we consider a relatively simple version. Namely, with probability $p$, both endpoints select blue and with probability $q = 1 - p$, both endpoints select red, independent of any other vertices or the actions taken before. This process is then repeated some large number of times. The color configuration (or *pattern*) of $G$ (usually) changes at each step. Many questions arise. For example, "What does a typical color configuration look like?", "What is the stationary distribution for the color patterns?", "How long will it take for before you are close to the stationary distribution?", or "What are most likely and least likely patterns that occur in the long run"?

Facing these questions, a natural reaction is to run simulations on some examples. However, as one might suspect, simulation seems to reveal very little since the number of possible color patterns is exponentially large and any particular one is very unlikely to occur (also see [3,13]). Although general questions of this type can be rather difficult, we will focus on a version for which we can make some progress (although there are still many more questions than answers, even here).

While we are considering vertex colorings of a graph $G$, there is in fact an associated state graph $H$ in play. The nodes of the state graph $H$ are the color patterns of vertices of $G$. If one color pattern $B$ can result from "flipping" an edge (i.e., choosing an edge at random and changing the colors of both endpoints to the same red or blue color) from another configuration $A$, then we will say there is a *directed edge* in $H$ from $A$ to $B$. Each time we flip an edge in $G$ (with some probability), we have a random walk moving from one node to another in the state graph $H$. If $G$ has $n$ vertices, the corresponding state graph can be quite large, e.g., having on the order of $2^{\binom{n}{2}}$ nodes. Nevertheless, a random walk on the state graph turns out to have some amazing properties which can be deduced by using the connection with the spectral theory for random walks on semigroups. The transition probability matrices for random walks on these state graphs have real eigenvalues which are of a surprisingly simple form. For example, for the case of $p = 1/2$, the eigenvalues of the state graph $H$ can be shown to all be of the form $k/m$ for some positive integer $k$ where $m = |E(G)|$, and with multiplicities that can also be explicitly determined. This result is a consequence of known results for random walks on certain semigroups [4–6]. In Sections 2–4, we will apply this theory in the context of the edge flipping process and its state graph. In particular, we will derive a relatively simple description for the spectrum of the walk on the state graph. Our description of eigenvalues involves only terminology from graph theory although in Section 2 we will illustrate the connection with the various structures (e.g., ideals, flats, chambers, lattices, etc.) that have been well studied in connection with semigroups.

From the spectrum, we can estimate with high probability the number of steps required for the random walk on the state graph to converge to the stationary distribution. However, the spectral theory for semigroups tells us very little about what the stationary distribution for a random walk on $H$ actually looks like. For example, for the stationary distribution, how will the maximum value of a vertex in $H$ compare with minimum value? In other words, which color patterns of $G$ are most likely (or least likely) to be reached? For two given configurations, which one is more likely to appear? In particular, Persi Diaconis raised the question of determining the stationary distribution for the state graph associated with edge flipping on an $n$-cycle.

In Sections 5–8, we investigate the stationary distribution on the state graph of edge flipping for several simple families of graphs. For example, we will show that when we flip edges in a cycle $C_n$ with flipping probability $1/2$, the stationary distribution of the associated state graph for the state with all vertices colored blue converges to $(2/\pi)^n$ as $n$ approaches infinity. We will also consider the solutions with a general flipping probability $p$ for several families of graphs, such as paths, ball-and-chains (i.e., a path together with a star at one end), and cycles, as well as some cases where there pattern under consideration consists of alternating red and blue blocks.

In addition to flipping edges, we also consider the process of flipping vertices in a graph. Vertex flipping can be viewed as the analog of the situation that one player can influence the decisions of others who are nearby neighbors. Thus, for the vertex flipping process, in each round, we choose a vertex $v$ of $G$ randomly and (re-)color all vertices within distance $d$ from $v$ (including $v$ itself) blue with probability $p$ and red with probability $q = 1 - p$. In Sections 6 and 7, some partial results and further questions for vertex flipping are discussed. This vertex flipping game is motivated by the "democratic primary game" in [13,14] which has a somewhat different formulation.

Although the flipping processes we examine here are somewhat special, the same approach can be adapted to deal with many other similar games, especially those with moves which are *memoryless*. Namely, if for any two moves $x$ and $y$, the effect of $xyx$ is the same as $xy$, and in addition, $xx$ has the same effect as $x$, we say it satisfies the *left-regular band* property (LRB), or the moves form an LRB semigroup [15,17] under concatenation. In such a scenario, we can compute the spectrum of the state graph and therefore determine the rate of convergence for the state graph.

This paper is organized as follows. In Section 2, we give definitions and a brief overview of random walks on LRB semigroups. In Section 3 we focus on the state graph of the edge flipping process and the connection with random walks on LRB semigroups. In particular, the spectrum of a state graph is given in graph-theoretical terminology. In Section 4, we describe a more general setting for processes of this type, and how these considerations can be used to give specific information about the stationary distributions for the edge flipping process. In Section 5, we examine the generating functions for the values of monochromatic states for colorings of paths and cycles. In Section 6, we extend the methods to states with mixed colors blocks. In Section 7, we consider other families of graphs for the edge flipping problem. The vertex flipping version is examined in Section 8. In Section 9, we look at the problem of finding the most frequent and least frequent color configurations (i.e., the ones which assume the largest and smallest values in the stationary distribution of the state graph). In the last section, we discuss several generalizations of the edge flipping process and mention some related problems, such as voter models.

## 2. Random walks on semigroups

Before we proceed to give a precise formulation of the edge flipping problem, we will first mention some background material concerning random walks on semigroups. There is a special class of semigroups which was introduced in 1940 by Klein-Barem [15] and also by Schützenberger [17], called left-regular bands. A *left-regular band*, or LRB for short, is a semigroup $S$ in which every element is idempotent and, in addition, the following identity is satisfied:

$$xyx = xy \tag{1}$$

for all $x$, $y$ in $S$. This implies that in an LRB, all occurrences of any term $x$ in a product which occurs to the right of the first occurrence of $x$ can be removed without affecting the value of the product. If $S$ is an LRB and is finitely generated, then $S$ is finite. For a probability distribution $\{w_x\}$ defined on $S$, we can define a random walk with transition probability matrix $P$ defined by

$$P(s,t) = \sum_{\substack{x \in S \\ xs = t}} w_x. \tag{2}$$

A useful special case of this walk is to restrict the walk to elements of some *left ideal*, i.e., a subset which is closed under left-multiplication by arbitrary elements of $S$.

There are many interesting problems that can be formulated as random walks on various LRB semigroups including hyperplane chamber walks [4], card shuffling [7], and self-organizing search [9]. One of the simplest examples is the semigroup $\mathfrak{S}$ consisting of sequences of elements of $\{1, 2, \dots, N\} = [N]$, and the multiplication defined by

$$(x_1, \dots, x_l)(y_1, \dots, y_m) = (x_1, \dots, x_l, y_1, \dots, y_m)^{\wedge}$$

where the wedge means deletion of any element that has already occurred to its left. The random walk is on the ideal $I$ consisting of the reduced words of length $N$ (i.e., having no repeated elements), each of which can be viewed as a permutation of $N$ elements. A step in the random walk consists in choosing some element $x \in \mathfrak{S}$ with the probability (or weight) associated with $x$, and moving from the current element $y \in I$ to the new element $(xy)^\wedge \in I$. If all the weights are concentrated on the sequences of length 1, the random walk is called the *Tsetlin library* and has applications to the *random-to-top* shuffle as well as the *move-to-front* data structure.

Here we state some basic definitions and useful facts about an LRB semigroup $S$. We mainly follow the definitions in [5].

(i) There is a natural partial order defined on $S$ defined by

$$x \leqslant y \quad \Leftrightarrow \quad xy = y.$$

Thus, left multiplication by $x$ is a projection from $S$ to $S_{\geqslant x}$ where

$$S_{\geqslant x} = \{y \in S\colon\ y \geqslant x\}.$$

(ii) A semilattice $L$ can be constructed from $S$ as follows: First we define a relation $\preccurlyeq$ on $S$ as follows:

$$y \preccurlyeq x \quad \Leftrightarrow \quad xy = x.$$

The equivalence class under $\preccurlyeq$ which contains $x$ is said to be the support of $x$, denoted by $\operatorname{supp} x$. In addition, for $x$, $y$ in $S$,

$$\operatorname{supp} xy = \operatorname{supp} x \vee \operatorname{supp} y$$

and

$$x \leqslant y \quad \Rightarrow \quad \operatorname{supp} x \preccurlyeq \operatorname{supp} y.$$

In many cases, the semilattice $L$ turns out to be a lattice (as seen later in the examples). Elements in $L$ are called flats (following the terminology for semigroups associated with matroids [5]).

(iii) An element $c \in S$ is said to be a *chamber* if $cx = c$ for all $x \in S$. Therefore $c$ is maximal in the poset $S$. The set of all chambers forms an ideal.

For example, the semigroup $\mathfrak{S}$ generated by $[N]$ is associated with the Boolean lattice of subsets of $[N]$, partially ordered by containment. The support of a word $(x_1, \ldots, x_l)$ is the underlying set $\{x_1, \ldots, x_l\}$. We usually have our random walk on the set $C$ of chambers (following the terminology from hyperplane random walks). In this example, the chambers are just words of length $N$.

The eigenvalues of a random walk on chambers of semigroups have the following elegant form (see [5,7]).

**Theorem A.** *For a semigroup $S$, a random walk on $S$, as defined in* (2)*, has eigenvalues indexed by elements of the semilattice $L$, as described in* (ii)*. For each $X \in L$, there is an eigenvalue*

$$\lambda_X = \sum_{x \in X} w_x$$

*with multiplicity $m_X$ satisfying*

$$\sum_{Y \succcurlyeq X} m_Y = c_X$$

*where $c_Y$ is the cardinality of $S_{\geqslant Y} = S_{\geqslant y} = \{z \in S\colon z \geqslant y\}$ where $y$ is any element with support $Y$ (this is independent of the choice of $y$). Alternatively,*

$$m_X = \sum_{Y \succcurlyeq X} \mu(X, Y) c_Y$$

*where $\mu$ is the Möbius function of the lattice $L$.*

For the semigroup $\mathfrak{S}$ generated by $[N]$, Phatarfod [16] first determined the eigenvalues of the random walk on chambers of $\mathfrak{S}$ with $w_X = 1/N$ for $X$ with length 1. For each subset $X \subset [N]$, there is an eigenvalue $\lambda_X = |X|/N$, with multiplicity equal to the so-called derangement number $d_k$ where $k = n - |X|$.

For a random walk on the family $C$ of chambers in a semigroup $S$, we can then use eigenvalues to bound the rate of convergence to the stationary distribution $\pi$. The *total variation distance* after taking $s$ steps in the random walk is bounded above (see [7]) as follows:

$$\begin{aligned}\|P^s - \pi\|_{TV} &= \max_{A \subseteq C} \max_y \left| \sum_{x \in A} P^s(y, x) - \pi(x) \right| \\ &= \frac{1}{2} \max_y \sum_x \left| P^s(y, x) - \pi(x) \right| \\ &\leqslant \sum_{X \in L^*} m_X \lambda_X^s\end{aligned}$$

where $L^*$ denotes the subset of the lattice $L$ excluding the maximum element.

Thus, for the random walk on the semigroup $\mathfrak{S}$ generated by $[N]$ with $w_X = 1/N$ for $X$ with length 1, the total variation distance after $s$ steps to the stationary distribution satisfies

$$\begin{aligned}\|P^s - \pi\|_{TV} &\leqslant \sum_{k=1}^{N} \left(1 - \frac{k}{N}\right)^s \binom{N}{k} k! \\ &\leqslant N^2 \left(1 - \frac{1}{N}\right)^s \leqslant e^{-c}\end{aligned}$$

if $s \geqslant 2N \log N + cN$.

For many examples in [5,7], the stationary distribution $\pi$ could be determined. However, for some other problems, including our edge/vertex flipping process, the stationary distributions were not previously known. More discussion of the stationary distributions will be given in later sections.

## 3. The spectra of random walks on state graphs

For a graph $G$ on $n$ vertices and $m$ edges, our state graph has nodes each of which represents a total coloring of vertices of $G$. We consider a deck of cards, each of which represents an action we can take on a state of $G$. Each edge of $G$ is associated with two cards, one of which assigns 0 (or red) to both endpoints and the other which assigns 1 (or blue) to both endpoints. Each card $c$ acts on a state $\sigma$ by flipping the edge associated with $c$ (e.g., the colors of two vertices in $\sigma$ which are endpoints associated with $c$ to the color specified by $c$). It is easy to check that $c_1 c_2 \sigma = c_1(c_2 \sigma)$. This defines the multiplication for the semigroup $S$ generated by a set of cards. Note that for $c_1$, $c_2$ in $S$, we regard $c_1 = c_2$ if $c_1 \sigma = c_2 \sigma$ for all states $\sigma$. Each element of $S$ is just a finite sequence of distinct cards. It is easy to check that every element of $S$ is idempotent and Eq. (1) is satisfied. Therefore $S$ is an LRB.

The following lemma is basically a straightforward application of the definitions (i)–(iii) in the preceding section and the results in [5,7] as stated in Theorem A:

**Lemma 1.** *Suppose a graph $G$ contains no isolated vertices. The semigroup $S$ generated by cards in the edge flipping process on a graph $G$ as described above satisfies the following*:

(1) *For $x$, $y$ in $S$, $\operatorname{supp} x \preccurlyeq \operatorname{supp} y$ if the set of vertices in $V$ having colors determined by $x$ is contained by the set of vertices with colors determined by $y$. Thus, the support of a card $c$ is the set of two endpoints of $c$. The support of an element $s$ of $S$ is the union of supports of cards in $s$. A flat is a subset of $V$ which is the support of some element in $S$. Therefore, $V$ is a flat and any subset $T$ of $V$ is a flat if the induced subgraph on $T$ in $G$ contains no isolated vertices. The set of flats forms a sublattice $L$ of the Boolean lattice on $V$.*

(2) *For $x$, $y$ in $S$, we say $x \leqslant y$ if the color configuration determined by $x$ is the same as the color configuration determined by $y$ restricted to the support of $x$. The partial order on $S$ is transitive, reflective and anti-symmetric (i.e., $x \leqslant y$ and $y \leqslant x$ imply $x = y$). Furthermore, a maximum element $c$ of $S$, i.e., with $\operatorname{supp} c = V$, determines a coloring configuration for all vertices in $G$. A chamber consists of all maximum elements of $S$ that determine the same coloring configuration. A random walk on chambers is thus associated with a random walk on the state graph.*

(3) *For each card $x$, we assign a probability $w_x$ with $\sum_x w_x = 1$. For the random walk on the state graph of $G$ with the probability of moving from a chamber $y$ to $xy$ defined to be $w_x$, the transition probability matrix can be diagonalized. The eigenvalues are indexed by elements of the associated lattice $L$. For each flat $X$ in $L$, the eigenvalue $\lambda_X$ is*

$$\lambda_X = \sum_{\operatorname{supp} y \subseteq X} w_y$$

*with multiplicity $m_X$ satisfying*

$$\sum_{Y \succcurlyeq X} m_Y = c_X$$

*and*

$$c_X = |S_{\geqslant X}| = |S_{\geqslant x}| = \left|\{y \in S:\ y \geqslant x\}\right|$$

*where $X = \operatorname{supp} x$.*

For example, if $G$ is a path $P_4$ on three edges, it can be easily checked that the spectrum of the random walk on the state graph of $P_4$ is 1, 2/3 (with multiplicity 2), 1/3 (with multiplicity 5), and 0 (with multiplicity 8). Another example is the case of 4-cycle $C_4$. The spectrum of the random walk on the state graph of $C_4$ is 1, 1/2 (with multiplicity 4), 1/4 (with multiplicity 4) and 0 (with multiplicity 7). For the case of $P_5$, the spectrum of the random walk on the state graph is 2, 3/4 (with multiplicity 2), 1/2 (with multiplicity 6), 1/4 (with multiplicity 10) and 0 (with multiplicity 13). In Fig. 1, the lattice of $P_5$ is illustrated. The vertices of $P_5$ are labeled consecutively by a, b, c, d, e. For each flat $F$, the associated eigenvalue is attached as well as the multiplicity $m(F)$. For example, for the empty flat, the associated eigenvalue 0 has multiplicity 13. There are exactly 13 subsets of $V$ which are independent sets (and so, have eigenvalue 0).

We are ready to describe the spectrum of the transition probability matrix for the random walk on the state graph of the edge flipping process:

**Theorem 1.** *For a graph $G$ with $n$ vertices and $m$ edges, let $w_x$ denote the probability we move from a state $\sigma$ to $x\sigma$ using the card $x$, and assume that the sum of all $w_x$ satisfies $\sum_x w_x = 1$. The associated random walk of*

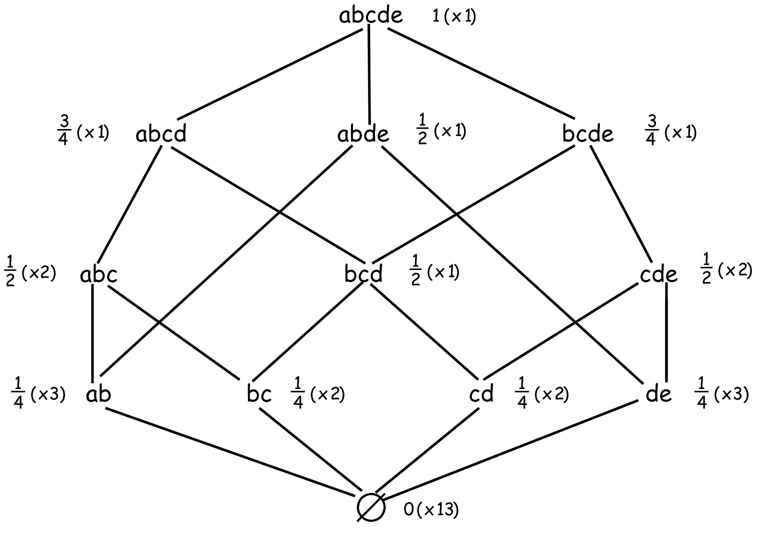


**Fig. 1.** The lattice for $P_5$ and the associated eigenvalues with multiplicities.

*edge flipping on the state graph of G has an eigenvalue*

$$\lambda_T = \sum_{\operatorname{supp} x \subseteq T} w_x$$

*of multiplicity* 1 *for each subset T of the vertex set V. For the special case of* $w_x = 1/(2m)$, *we have*

$$\lambda_T = \frac{e(T)}{m}$$

*where* $e(T)$ *denotes the number of edges in the induced subgraph on T in G.*

**Proof.** Instead of indexing the eigenvalues by flats, we will index eigenvalues by subsets of the vertices which we will see amounts to the same thing.

**Claim.** *For a flat T, the multiplicity* $m_T$ *for the eigenvalue* $\lambda_T$ *is equal to the number of supersets* $T'$ *containing T with the property that the induced subgraph on* $T'$ *has the same edge set as the induced subgraph on T in G.*

**Proof.** We will prove this by induction. For $T = V$, the largest flat of $G$, $m_T = 1 = |\{T'\colon\ T' \supseteq V\}|$. Now suppose $T$ is not the largest flat of $G$, but that the statement of the Claim is true for any flat $F$ with $|F| > |T|$.

From Lemma 1, we see that for a flat $T$, $c_T$ is exactly the number of states in which the colors of vertices in $T$ are already determined by an element $x$ of $S$ with support $T$. Namely,

$$c_T = |S_{\geqslant T}| = \left|\{y \in S\colon\ y \geqslant x\}\right| = 2^{n-|T|}.$$

Therefore

$$\sum_{Y \succcurlyeq T} m_Y = 2^{n-|T|} = \left|\left\{T'\colon\ T' \supseteq T\right\}\right|. \tag{3}$$

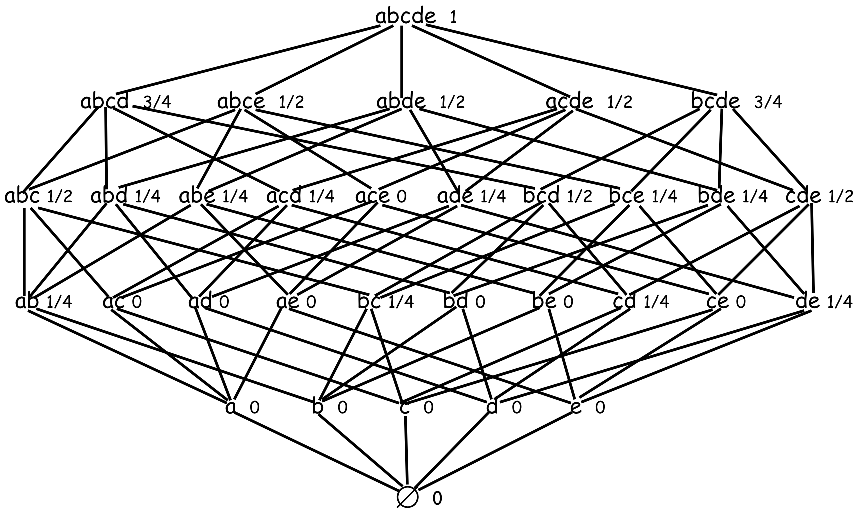


**Fig. 2.** The Boolean lattice for $P_5$ and the associated eigenvalues (each with multiplicity 1).

By induction, for a flat $Y \supset T$, we have

$$m_Y = \left|\left\{T'\colon\ T' \supseteq Y \text{ and } e\left(T'\right) = e(Y)\right\}\right|.$$

Substituting into (3), we have

$$\begin{aligned} m_T &= 2^{n-|T|} - \sum_{Y \succ T} m_Y \\ &= 2^{n-|T|} - \sum_{Y \succ T} \left|\left\{T'\colon\ T' \supseteq Y \text{ and } e\left(T'\right) = e(Y)\right\}\right| \\ &= \left|\left\{T'\colon\ T' \supseteq T\right\}\right| - \sum_{Y \succ T} \left|\left\{T'\colon\ T' \supseteq Y \text{ and } e\left(T'\right) = e(Y)\right\}\right| \\ &= \left|\left\{T'\colon\ T' \supseteq T \text{ and } e\left(T'\right) = e(T)\right\}\right|. \end{aligned}$$

The Claim is proved. □

Now we index the eigenvalues by subsets of $V$. If a subset $T'$ is not a flat, it contributes 1 toward the multiplicity of the eigenvalue $\lambda_T$ where $T'$ contains a flat $T$ and the induced subgraph on $T'$ has the same edge set as that of $T$. This completes the proof of Theorem 1. □

As an example, we consider the Boolean lattice of $P_5$ in Fig. 2. Next to each subset $T$ of vertices, the associated eigenvalue (i.e., the number of edges in the induced subgraph on $T$ divided by the total number of edges) is illustrated. It is easy to check that there are exactly 13 subsets of $V$ which are independent sets in Fig. 2.

We remark that the eigenvalues stay the same if the condition $w_x = 1/(2m)$ for all colored cards $x$ is replaced by the condition that $w_x + w_y = 1/m$ for two cards $x$, $y$ associated with the same edge. Hence the eigenvalues of the edge flipping random walk on the state graph on $G$ are independent of the selection probability $p$.

**Theorem 2.** *The random walk on the state graph of edge flipping process with $p = 1/2$ on a connected graph $G$ on $n$ vertices and $m$ edges converges to the stationary distribution in $2m \log n$ steps. In other words, the total variation distance of the transition probability matrix $P$ satisfies*

$$\left\| P^s - \pi \right\|_{TV} \leqslant e^{-c}$$

*if $s \geqslant 2m \log n + cn$.*

**Proof.** Using Theorem 1, the total variation distance of the random walk on the state graph of $G$ after $s$ steps satisfies

$$\begin{aligned} \left\| P^s - \pi \right\|_{TV} &\leqslant \sum_{F \in L^*} m_F \lambda_F^s \\ &\leqslant \sum_{k=1}^{n} \left(1 - \frac{\delta_k}{m}\right)^s \binom{n}{k} \end{aligned}$$

where $\delta_k$ denotes the minimum number of edges incident to $k$ vertices in $G$. Since $G$ is connected, we have

$$\begin{aligned} \left\| P^s - \pi \right\|_{TV} &\leqslant n^2 \left(1 - \frac{1}{m}\right)^s \\ &\leqslant e^{-c} \end{aligned}$$

if $t \geqslant 2m \log n + cn$. Theorem 2 is proved. □

## 4. A general setting

Here, we describe a more abstract version of our process. We begin with a finite set $X$ of points and a set $[r] := \{1, 2, \ldots, r\}$ of "colors". A *coloring* $F$ is defined to be a mapping $F : X \to [r]$. We will denote by $\mathbf{F}$ the set of all possible colorings of $X$. We also are given a collection of (not necessarily distinct) subsets (or *edges*) $X_k$ of $X$, for $1 \leqslant k \leqslant t$, where we assume that $\bigcup_k X_k = X$. With each subset $X_k$ is associated a mapping $c_k : X_k \to [r]$, and a probability $p_k > 0$. We assume that $\sum_k p_k = 1$. Thus, each $c_k$ "colors" the points of $X_k$ but leaves points not in $X_k$ untouched. We will call the $c_k$ *cards* in a $t$-card deck $\mathbf{C}$.

Our random walk proceeds as follows. At each step a card $c_k$ is selected from $\mathbf{C}$ with probability $p_k$ and is applied to the current coloring $F$ of $X$. Thus, $c_k$ (re-)colors the points of $X_k \subseteq X$ but leaves the colors of points of $X$ not in $X_k$ unchanged. In this way a new coloring $F'$ of $X$ is produced. Standard arguments in Markov chain theory show that this process has a unique stationary distribution $\pi$ on $\mathbf{F}$.

For example, a special case is our edge flipping process on a graph $G$, where $X$ is the set of vertices of $G$, the set of colors is $\{0, 1\}$, the $X_k$ are (ordinary) edges (= 2-element subsets of $X$) of $G$, and for each edge $X_k$ there are two cards, one which assigns the color 0 to both endpoints of $X_k$, and one which assigns the color 1 to these endpoints. In fact, for the remainder of the paper, we will usually use 0 and 1 as colors, rather than red and blue.

We will consider the semigroup $S = \langle c_1, c_2, \ldots, c_t \rangle$ generated by all the cards $c_k$, with the semigroup operation defined to be composition of maps. Let us call a word $c_{i_1} c_{i_2} \ldots c_{i_r} \in S$ *full* if $\bigcup_{j=1}^{r} X_{i_j} = X$. We will let $I$ denote the (left) ideal of full words in $S$. It is now easy to check that $S$ is an LRB to which the results of [7] can be applied.

Observe that in a full word $c_{i_1}c_{i_2}\dots c_{i_r}$, if for some $u$, $\bigcup_{j=1}^{u-1} X_{i_j} \supseteq X_{i_u}$, then the card $c_{i_u}$ can be removed from the product without affecting the final coloring of $X$ this product produces (since the earlier terms assign colors later to all the vertices that $c_{i_u}$ colors). When all such "redundant" cards are removed from a full word, we will say that the word is *reduced*. Thus, we can replace any word from $S$ by the equivalent reduced word which has the same coloring action on $X$.

In particular, the unique stationary distribution $\pi$ for this walk is the same as the distribution obtained by the following process. Namely, select all $t$ cards *without replacement* from our deck **C** according the probability distribution $p_k$, thereby producing an ordering $(c'_1, c'_2, \dots, c'_t)$ of the cards which in turn produces some coloring of $X$. (Remember that we apply this map from the left, i.e., first apply $c'_t$, then $c_{t-1}$, etc., and finally $c_1$ to get the final coloring of $X$.) Then the distribution given by this process is also $\pi$. This was first observed in [7].

It is not difficult to see why this is so. For example, let us compute the probability that in the original (unbounded) process (with replacement), the last occurrence of card $c_1$ occurs *before* the last occurrence of card $c_2$. Thus, wherever the last occurrence of $c_1$ is (which occurs with probability $p_1$), the remaining $s$ cards must have no $c_1$ but at least one $c_2$. The probability that this happens is $p_1((1-p_1)^t - (1-p_1-p_2)^t)$. Summing this over $s \geqslant 1$ gives:

$$\sum_{s\geqslant 1} p_1\bigl((1-p_1)^s - (1-p_1-p_2)^s\bigr) = p_1\left(\frac{1-p_1}{p_1} - \frac{1-p_1-p_2}{p_1+p_2}\right) = \frac{p_2}{p_1+p_2}.$$

Of course, this is just the probability that in our selection process *without replacement*, $c_2$ is selected before $c_1$ (and therefore occurs to the left of $c_1$ in the ordering).

As another example, let us compute the probability that the order of the last occurrences of cards $c_1$, $c_2$ and $c_3$ is $(c_1, c_2, c_3)$ in the unbounded process. Using the previous computation, we see that the probability that $c_3$ occurs later than the last occurrences of $c_1$ and $c_2$ is $\frac{p_3}{p_1+p_2+p_3}$. If we now consider the cards occurring before the last occurrence of $c_3$, then the probability that $c_2$ occurs later than the last occurrence of $c_1$ is $\frac{p_2}{p_1+p_2}$. Thus, the last occurring order $(c_1, c_2, c_3)$ occurs with probability $\frac{p_3}{p_1+p_2+p_3}\frac{p_2}{p_1+p_2}$. However, this is just the probability that for the selection process without replacement, $c_3$ is chosen before $c_2$ which in turn is chosen before $c_1$.

Similar arguments can be used to show that for any set of cards, the probability that the *last occurrences* of these cards occur in some particular order in the unbounded process with replacement is exactly the same as the probability for this ordering of these cards to occur in the selection process without replacement. We should keep in mind that when coloring $X$ with the process without replacement, the *first* color assigned by a card to a vertex is the final color of that vertex.

This fact will allow us to compute (recursively) the stationary distributions for several families of simple graphs (see Sections 5–8).

As an example of a variant of our edge flipping process, we can consider a "vertex flipping" process on a graph $G$. Here, each vertex $v$ in $G$ will be associated with two cards: $c_v(0)$ which assigns 0 to $v$ and all its neighbors, and $c_v(1)$ which attempts to assign 1 to $v$ and all its neighbors. (Of course, if a vertex has already been assigned a value, then the card does nothing to that vertex.) The semigroup $S$ is generated by this set of cards. Let $p_v(i)$ be the probability of choosing the card $c_v(i)$ for our random walk on the state graph of $G$ where we assume $\sum_{v,i} p_v(i) = 1$. We can apply arguments very similar to those used earlier to prove the following:

**Theorem 3.** *Let $G$ be a graph on $n$ vertices with the deck of $2n$ cards as described above. Thus, the probability we move from a state $\sigma$ to $c_v(i)\sigma$ (which means we assign the color $i$ to the vertex $v$ and all its neighbors) is $p_v(i)$. Then the random walk associated with vertex flipping on the state graph of $G$ has the eigenvalue*

$$\lambda_T = \sum_{\operatorname{supp} c_v(i) \subseteq T} p_v$$

*for each subset $T$ of the vertex set $V$, where* $\operatorname{supp} c_v(i)$ *for the card* $c_v(i)$ *consists of the vertex $v$ and all the neighbors of $v$, and* $p_v = p_v(0) + p_v(1)$. *For each $T$, the multiplicity of $\lambda_T$ is 1. For the special case of* $p_v(i) = 1/(2n)$ *for all $v$ and $i$, we have*

$$\lambda_T = \frac{\delta(T)}{n}$$

*where $\delta(T)$ denotes the number of vertices $v$ with all its neighbors in $T$.*

In general, for a graph $G$, our deck $\mathbf{C}$ could consist of cards $c_A$ which color the endpoints of *arbitrary subsets* $A$ of edges of $G$, and which are selected with probability $p_A$, where $\sum_A p_A = 1$. We can then consider the semigroup $S$ which is generated by such a set of cards.

**Theorem 4.** *The random walk on the state graph of G as specified above has the eigenvalue*

$$\lambda_T = \sum_{\operatorname{supp} c_A \subseteq T} p_A$$

*for each subset $T$ of the vertex set $V$, where* $\operatorname{supp} c_A$ *denotes the set of vertices that are involved in the card $c_A$, and* $p_A = p_A(0) + p_A(1)$.

Note that when $T = \emptyset$, the corresponding eigenvalue is 0, and its multiplicity is just the number of independent sets of vertices in $G$. Since determining the number of independent sets in a graph in general is a $\#P$-complete problem (even for planar bipartite graphs with maximum degree 4 (see [18])), it is usually not easy to determine exactly the eigenvalue multiplicities for large unstructured graphs. However, for special families of graphs, such as a path, this can be done fairly easily.

**Theorem 5.** *For a path $P_n$ on $n > 1$ vertices, we consider the edge flipping process with $p = 1/2$. The eigenvalues of the state graph are $k/(n-1)$ with multiplicity $S(n,k)$ satisfying the following generating function*:

$$\mathbf{S}(x,y) := \sum_{n,k \geqslant 0} S(n,k) x^n y^k = \frac{1 + x - xy}{1 - x - xy - x^2 + x^2 y}$$
$$= 1 + 2x + 3x^2 + x^2 y + 5x^3 + 2x^3 y + 8x^4 + 5x^4 y + x^3 y^2 + 13x^5 \ldots.$$

**Proof.** By the results of the preceding theorems, we need to count for each $k$, the number of subsets of the vertex set of $P_n$ which induce a subgraph having exactly $k$ edges. Let us denote this number by $S(n,k)$. Thus, $S(n,k)$ is just the number of binary $n$-tuples which have exactly $k$ pairs of consecutive 1's. Let $A(n,k)$ denote the number of such sequences which begin with 0, and let $B(n,k)$ denote the number which begin with 1. Then it is easy to see that

$$A(n+1,k) = A(n,k) + B(n,k),$$
$$B(n+1,k) = A(n,k) + B(n,k-1)$$

for $n, k \geqslant 0$. Eliminating $B$ from the above and shifting indices, we obtain the recurrence

$$A(n,k) = A(n-1,k) + A(n-1,k-1) + A(n-2,k-1) - A(n-2,k-2)$$

where for convenience, we define $A(n,-1) = 0$ for all $n$ and $A(0,0) = 1$. Since $A(1,0) = 1$ and $A(1,1) = 0$, then $A(n,k)$ is well defined for all $n \geqslant 0$ and $k \geqslant -1$. If we now define the two-variable generating function

$$\mathbf{A}(x,y)=\sum_{n,k\geqslant 0}A(n,k)x^n y^k$$

then by standard "manipulatorics" (see [11]), it follows that $\mathbf{A}(x,y)$ is given by the generating function

$$\mathbf{A}(x,y)=\frac{1-xy}{1-x-xy-x^2+x^2y}.$$

Since $S(n,k)=A(n+1,k)$ then we have

$$\mathbf{S}(x,y)=\sum_{n,k\geqslant 0}S(n,k)x^n y^k=\frac{1+x-xy}{1-x-xy-x^2+x^2y}. \qquad \square$$

Notice that for $k=0$, we have $S(n,0)=F_{n+2}$, the familiar Fibonacci numbers. Also, notice these coefficients are just the values shown in Fig. 1.

## 5. Edge flipping on a cycle

We will first focus on the edge flipping process on the *augmented* path $\tilde{P}_n$ (which will be useful for dealing with cycles and paths later). By augmented, we mean a path on $n-1$ vertices $1,2,\ldots,n-1$ which in addition to the usual edges $\{i,i+1\}$ for $1\leqslant i<n-1$, also has two additional (half) edges attached to 1 and $n-1$. These two additional edges are treated just like all the other edges, except that the assignments to their missing endpoints are ignored. We do this for the purpose of making the induction argument cleaner.

Our first task will be to determine the probability $u(n)$ that all of the $n-1$ vertices of $\tilde{P}_n$ have the value 0 assigned to them, where we will set $u(0)=u(1)=1$ (and we will always take $u(m)=0$ for $m<0$). As usual, we assume $p_e(0)=p$ and $p_e(1)=q=1-p$ for each edge $e$.

**Theorem 6.** *$u(n)$ is given by*

$$\begin{aligned}U(z)=\sum_{n\geqslant 0}u(n)z^n&=1+z+pz^2+\frac{1}{2}p(2p+1)z^3+\cdots\\&=\sqrt{\frac{q}{p}}\tan\left(z\sqrt{pq}-\arctan\left(\sqrt{\frac{p}{q}}\right)\right).\end{aligned}$$

**Proof.** First, consider $\tilde{P}_2$ which has one vertex and two (half) edges. The probability that either edge is selected first is $\frac{1}{2}$, so that the probability that the vertex is assigned the value 0 is just $u(2)=\frac{1}{2}p+\frac{1}{2}p=p$.

Now consider the path $\tilde{P}_n$ for some $n\geqslant 3$. Each of the $n$ edges is equally likely to be selected first. When it is selected, it determines the values of its endpoints (which in this case, must be 0). If this first edge selected is one of the two end (half) edges then the resulting graph is a shorter path $\tilde{P}_{n-1}$, which by induction, has probability of getting the all 0 assignment equal to $u(n-1)$. On the other hand, if the first edge selected is one of the (internal) $n-2$ edges, then it splits $\tilde{P}_n$ into two smaller paths $\tilde{P}_k$ and $\tilde{P}_{n-2-k}$ for some $k$ with $1\leqslant k\leqslant n-3$. Adding all these cases up, and using the definition of $u(k)$, we find that $u(n)$ satisfies the recurrence

$$u(n)=\frac{p}{n}\left(\sum_{k=0}^{n-1}u(k)u(n-1-k)\right). \tag{4}$$

Let us set

$$U(z)=\sum_{n\geqslant 0}u(n)z^n.$$

Multiplying (4) by $nz^{n-1}$ and summing over $n\geqslant 2$, we have

$$\sum_{n\geqslant 2}n\,u(n)z^{n-1}=p\sum_{n\geqslant 2}\sum_{k=0}^{n-1}u(k)u(n-1-k)z^{n-1}. \tag{5}$$

However, the left-hand side of (5) is simply $\frac{d}{dz}U(z)-1$ while the right-hand side is $p(U(z)^2-1)$. Hence, we can rewrite (5) as

$$\frac{d}{dz}U(z)=pU(z)^2+q. \tag{6}$$

It is straightforward to check that the solution to this differential equation satisfying the boundary condition $U(0)=1$ is given by

$$U(z)=\sqrt{\frac{q}{p}}\tan\left(z\sqrt{pq}+\arctan\left(\sqrt{\frac{p}{q}}\right)\right), \tag{7}$$

and the proof is complete. □

Note that for the complementary probability $\tilde{u}(n)$ that all the vertices are assigned 1, we can simply interchange $p$ and $q$ in the corresponding expressions. Thus, we have

$$\tilde{U}(z)=\sqrt{\frac{p}{q}}\tan\left(z\sqrt{pq}+\arctan\left(\sqrt{\frac{q}{p}}\right)\right)=\sum_{n\geqslant 0}\tilde{u}(n)z^n. \tag{8}$$

The asymptotic growth of the coefficients $u(n)$ of $U(z)$ can be obtained by examining the singularities of $U(z)$ closest to the origin (e.g., see [10]). It is not hard to check that this is the singularity at $\frac{1}{\sqrt{pq}}\arctan(\sqrt{\frac{q}{p}})$. Hence, we find (after some calculation) that

$$u(n)\sim\frac{1}{p}r^{n+1} \tag{9}$$

where

$$r=\frac{\sqrt{pq}}{\arctan(\sqrt{\frac{q}{p}})}. \tag{10}$$

Similarly,

$$\tilde{u}(n)\sim\frac{1}{q}\tilde{r}^{n+1} \tag{11}$$

where

$$\tilde{r} = \frac{\sqrt{pq}}{\arctan(\sqrt{\frac{p}{q}})}. \tag{12}$$

It is now easy to compute the (asymptotic) probability $C(n)$ that this process on the $n$-cycle has all its vertices assigned 0. Namely, the probability that any particular edge is the first one selected is $\frac{1}{n}$ and then the value 0 must be chosen (this happens with probability $p$). After this selection is made, we have left an augmented path with $n-2$ vertices, and this reaches the correct state with probability $u(n-1)$. Summing over all $n$ edges of the cycle, we obtain

**Theorem 7.**

$$C(n) \sim r^n.$$

Note that when $p = \frac{1}{2}$ then $r = \frac{2}{\pi} = 0.6366\ldots$ which is much larger than $\frac{1}{2}$.

## 6. The stationary distribution of paths and cycles with mixed colors

We now suppose that we will carry out the same process on a general undirected graph $G = (V, E)$. Thus, we select edges uniformly at random without replacement, and then assign a value of 0 (with probability $p$) or 1 (with probability $q = 1 - p$) to each unassigned endpoint of the selected edge. As before, once a vertex is assigned a value then nothing further happens to it. Let $\lambda : V \to \{0, 1\}$ denote some assignment to the vertex set $V$ of $G$. For an arbitrary fixed vertex $x \in V$, let $\lambda_x : V \setminus \{x\} \to \{0, 1\}$ denote the assignment which is just $\lambda$ restricted to $V \setminus \{x\}$. Let $E$ denote the event that terminal assignment is given by $\lambda_x$ on $V \setminus \{x\}$, where it doesn't matter what gets assigned to $x$. For $i = 0, 1$, let $E_i$ denote the event that the terminal assignment $\lambda$ agrees with $\lambda_x$ on $V \setminus \{x\}$ *and* has $\lambda(x) = i$.

**Lemma 2** *(Reduction Lemma).*

$$\Pr(E) = \Pr(E_0) + \Pr(E_1). \tag{13}$$

Eq. (13) follows from the definition of probability.

As an example of an application of (13), let $C(n-1, 1)$ denote the probability that the terminal assignment on an $n$-cycle is all 0 *except* on the single vertex $n$, on which it is 1. By (13), we see that $C(n-1, 1) + C(n) = u(n)$ (where the vertex $n$ is taken for $x$ in the lemma). Hence,

$$C(n-1, 1) = u(n) - C(n) \sim \frac{1}{p} r^{n+1} - r^n \sim \left(\frac{r}{p} - 1\right) r^n. \tag{14}$$

Now let us return to (augmented) paths. Let $v(k, l)$ denote the probability that the final assignment on a path of $k + l$ vertices has the first $k$ assigned 0 and the last $l$ assigned 1. This path has $k + l + 1$ edges, two of which are the so-called half edges. Thus, by (13), we have

$$v(a, b-1) + v(a-1, b) = u(a)\tilde{u}(b) \tag{15}$$

for $a \geqslant 0$, $b \geqslant 0$ where, as usual, if $v$ has a negative argument, then its value is 0. Applying this recursively leads to the expression

$$v(a, b-1) = \sum_{i \geqslant 0} (-1)^i u(a-i)\tilde{u}(b+i). \tag{16}$$

Using (7) and (8), we have the following:

**Theorem 8.** *For positive integers a and b,*

$$v(a,b) = \big(1+o(1)\big)\frac{1}{\sqrt{pq}}\left(\frac{2}{\pi}\right)^{a+b-1} r^{a+1}\tilde{r}^{b+1}$$

*where the* $o(1)$ *term goes to zero if* $n = a+b$ *approaches infinity and the ratios* $a/n$ *and* $b/n$ *are bounded away from* 0.

Letting $\tilde{v}(b,a)$ denote the probability for a path on $b+a$ vertices with the first $b$ having the value 1 and the last $a$ having the value 0, then by (15) we have the symmetric form

$$\tilde{v}(b,a-1) = v(a-1,b) = \sum_{i\geqslant 0}(-1)^i\tilde{u}(b-i)u(a+i). \tag{17}$$

Specific examples of (15) are:

$$\begin{aligned}
v(n-1,0) &= u(n),\\
v(n-2,1) &= u(n-1)-u(n),\\
v(n-3,2) &= qu(n-2)-u(n-1)+u(n),\\
v(n-4,3) &= \frac{1}{3}q(2q+1)u(n-3)-qu(n-2)+u(n-1)-u(n).
\end{aligned}$$

If we let $C(a,b)$ denote the probability that a cycle on $a+b$ vertices has the first $a$ vertices assigned the value 0, and the remaining $b$ vertices assigned the value 1, then by (13) we have:

$$C(a,b)+C(a+1,b-1) = v(a,b-1) \tag{18}$$

for $b\geqslant 1$ (where $C(a,0)=C(a)$).

By using (8), we have

**Theorem 9.** *For positive integers a and b,*

$$C(a,b) = \big(1+o(1)\big)\left(\frac{2}{\pi}\right)^{2} r^{a}\tilde{r}^{b}$$

*where the* $o(1)$ *term goes to zero if* $n = a+b$ *approaches infinity and the ratios* $a/n$*, and* $b/n$ *are bounded away from* 0.

Also, we have

$$\begin{aligned}
C(n,0) &= pu(n-1),\\
C(n-1,1) &= u(n)-pu(n-1),\\
C(n-2,2) &= (1+p)u(n-1)-2u(n),\\
C(n-3,3) &= qu(n-2)-(2+p)u(n-1)+3u(n),\\
C(n-4,4) &= \frac{1}{3}q(2q+1)u(n-3)-2qu(n-2)+(3+p)u(n-1)-4u(n).
\end{aligned} \tag{19}$$

The same inductive argument used to show (16) can be used to establish the value $v(a_1, a_2, a_3, \ldots, a_{m-1}, a_m - 1)$ for the path with alternating blocks which has the first $a_1$ vertices assigned 0, the next $a_2$ vertices assigned 1, the next $a_3$ vertices assigned 0, and so forth, with the $m$-th (final) block of $a_m - 1$ vertices assigned 0 or 1, according to whether $m$ is odd or even. In this case we get:

$$v(a_1, a_2, \ldots, a_{m-1}, a_m - 1) = \sum_{i_1, i_2, \ldots, i_{m-1} \geqslant 0} (-1)^{i_1 + i_2 + \cdots + i_{m-1}} \prod_{j=1}^{m} u_j(a_j - i_j + i_{j-1})$$

where $i_k = 0$ if $k \leqslant 0$ or $k \geqslant m$, and

$$u_j(x) = \begin{cases} u(x) & \text{if } j \text{ is even,} \\ \tilde{u}(x) & \text{if } j \text{ is odd.} \end{cases}$$

This in turn implies the following asymptotic expression for $v(a_1, a_2, a_3, \ldots, a_{m-1}, a_m)$ when each ratio $\frac{a_i}{\sum_i a_i}$ is bounded away from 0 as the sum goes to infinity.

$$v(a_1, a_2, a_3, \ldots, a_m) \sim \gamma_m \left(\frac{2}{\pi}\right)^{A + \tilde{A} - 1} r^A \tilde{r}^{\tilde{A}} \tag{20}$$

where

$$A = \sum_{i \text{ odd}} a_i, \qquad \tilde{A} = \sum_{i \text{ even}} a_i$$

and

$$\gamma_m = \begin{cases} \frac{r^2}{p} & \text{if } m \text{ is odd,} \\ \frac{r\tilde{r}}{\sqrt{pq}} & \text{if } m \text{ is even.} \end{cases}$$

Intuitively, we can think of this expression as being formed in the following way. As we move along the path from left to right (starting with the first block of $a_1$ 0's), we are charged a factor of $r$ for each vertex with value 0, and a factor of $\tilde{r}$ for each vertex with value 1. In addition, whenever we *transition* from a block of 0's to a block of 1's, we accumulate an additional factor of $\frac{2}{\pi} \frac{\tilde{r}\sqrt{p}}{r\sqrt{q}}$. Symmetrically, when we transition from a block of 1's to a block of 0's, we accumulate a factor of $\frac{2}{\pi} \frac{r\sqrt{q}}{\tilde{r}\sqrt{p}}$. The cumulative effect of these factors over $r$ blocks is to accumulate the "change" factors shown in (20).

Using this same approach, the value of $C(a_1, a_2, a_3, \ldots, a_{2m})$ for a cycle of alternating blocks (with a block of $a_1$ 0's, followed by a block of $a_2$ 1's, etc.) is given by

$$C(a_1, a_2, a_3, \ldots, a_{2m}) \sim \left(\frac{2}{\pi}\right)^{2m} r^A \tilde{r}^{\tilde{A}} \tag{21}$$

where $A$ and $\tilde{A}$ are defined as above.

## 7. Edge flipping on the ball-and-chain graph

In this section we will determine the probability that all the vertices of a particular graph $BC(n, d)$ (which we have flippantly named the ball-and-chain graph) end up with value 0. The graph $BC(n, d)$ consists of an augmented path on $n$ vertices, say the set $\{v_1, v_2, \dots, v_n\}$, in which $v_1$ has a half edge attached to it, and the other terminal vertex $v_n$ has $d$ additional vertices attached to it. Thus, $BC(n, d)$ has $n+d$ vertices and $n+d$ edges in which one end has a vertex of degree $d+1$. We let $u_d(n)$ denote the probability that all the vertices of $BC(n, d)$ have value 0 at the end of the process. We show the following:

**Theorem 10.** *For a fixed integer $d > 0$, as n goes to infinity, we have*

$$u_d(n) \sim \left((-1)^{\frac{d-1}{2}} d! \left(\frac{p}{q}\right)^{\frac{d}{2}} sin\left(\frac{\sqrt{pq}}{r}\right) + +T_d\left(\frac{1}{r}\right)\right) u(n+d)$$

*where*

$$T_d(z) = p^d \sum_{k=1}^{\lfloor \frac{d}{2} \rfloor} \frac{(-1)^{k-1}}{(pq)^k} \binom{d}{2k} (2k)! z^{d-2k}. \tag{22}$$

**Proof.** We let

$$U_d(z) = \sum_{n \geqslant 0} u_d(n) z^{n+d} \tag{23}$$

denote the corresponding generating function. Arguing inductively, we see that $u_d(n)$ satisfies that recurrence

$$u_d(n) = \frac{p}{n+d} \left( \sum_{k=0}^{n-1} u(k) u_d(n-1-k) \right) + dp^{n-1} u(n), \quad n \geqslant 1, \tag{24}$$

where we define $u_d(0) = p^d$. Multiplying (24) by $(n+d)z^{n+d-1}$ and summing over all $n \geqslant 1$, we obtain

$$\sum_{n \geqslant 1} (n+d) u_d(n) z^{n+d-1} = p \sum_{n \geqslant 1} \sum_{k=0}^{n-1} u(k) u_d(n-1-k) z^{n+d-1} + dp^d \sum_{n \geqslant 1} u(n) z^{n+d-1}. \tag{25}$$

The left-hand side of (25) is just $U_d(z)' - 1$. The first term on the right-hand side of (25) is exactly $pU(z)U_d(z)$, while the second term is $dp^d z^{d-1}(U(z) - 1)$. Thus, $U_d(z)$ satisfies the differential equation

$$U_d(z)' = \left(pU_d(z) + dp^d z^{d-1}\right) U(z). \tag{26}$$

Let us assume that the solution to (26) has the form

$$U_d(z) = U(z) R_d(z) + S_d(z)$$

for suitable functions $R_d$ and $S_d$. Letting $A_d$ denote the quantity $dp^d z^{d-1}$, we have

$$\begin{aligned} U_d' &= R_d' U + R_d U' + S_d' \\ &= R_d' U + R_d\big(pU^2 + q\big) + S_d' \\ &= \big(pR_d U + R_d'\big)U + R_d q + S_d'. \end{aligned}$$

Thus, if $R_d$ and $S_d$ satisfy

$$R_d' = pS_d + A_d, \qquad R_d q + S_d' = 0 \tag{27}$$

then we have

$$(pR_d U + pS_d + A_d)U + R_d q + S_d' = (pU_d + A_d)U,$$

as desired. Eliminating $S_d$ from (27) leads to the second-order differential equation

$$\frac{d^2 R_d}{dz^2} + pqR_d = A_d'. \tag{28}$$

The general solution to (28) is

$$R_d = c_1 \sin(z\sqrt{pq}) + c_2 \cos(z\sqrt{pq}) + T_d \tag{29}$$

where $T_d$ is the polynomial which satisfies (28) and is given by

$$\begin{aligned} T_d &= \sum_{k=1}^{\lfloor \frac{d}{2} \rfloor} \frac{(-1)^{k-1}}{(pq)^k} A_d^{(2k-1)} \\ &= p^d \sum_{k=1}^{\lfloor \frac{d}{2} \rfloor} \frac{(-1)^{k-1}}{(pq)^k} \binom{d}{2k} (2k)! z^{d-2k}, \end{aligned} \tag{30}$$

where $A_d^{(j)}$ denotes the $j$-th derivative of $A_d$ with respect to $z$.

Observe that if $d$ is even, then $A_d'$ is an even function of $z$ and so in the solution to (29), $c_1$ must be 0. Similarly, if $d$ is odd, then $c_2 = 0$. Let us first consider the case that $d$ is even. The constant term of $U_d(z)$ is 0, so by extracting the constant term of $R_d U + S_d$, we obtain the equation

$$c_2 = (-1)^{\frac{d}{2}} d! \left(\frac{p}{q}\right)^{\frac{d}{2}}.$$

Substituting this value into (28) gives us the explicit form for $U_d(z)$ when $d$ is even:

$$\begin{aligned} U_d(z) &= R_d U + S_d \\ &= \left((-1)^{\frac{d}{2}} d! \left(\frac{p}{q}\right)^{\frac{d}{2}} d \cos(z\sqrt{pq}) + T_d\right) U + \frac{1}{p}\big(R_d' - A_d\big). \end{aligned}$$

Applying a similar argument for $d > 1$ odd, so that $c_2 = 0$, we find that

$$c_1 = (-1)^{\frac{d-1}{2}} d! \left(\frac{p}{q}\right)^{\frac{d}{2}},$$

so that in this case, we have

$$\begin{aligned} U_d(z) &= R_d U + S_d \\ &= \left((-1)^{\frac{d-1}{2}} d! \left(\frac{p}{q}\right)^{\frac{d}{2}} \sin(z\sqrt{pq}) + T_d\right) U + \frac{1}{p}\left(R'_d - A_d\right). \end{aligned}$$

From these expressions we can derive the asymptotic behavior of $u_d(n)$ as $n \to \infty$. In particular, since the singularity of $U(z)$ closest to the origin is the simple pole at $\frac{1}{\sqrt{pq}} \arctan(\sqrt{\frac{q}{p}}) = \frac{1}{r}$, then

$$u_d(n) \sim \left(R_d\left(\frac{1}{r}\right) + T_d\left(\frac{1}{r}\right)\right) u(n+d). \qquad \square$$

Note that, since $u(n) \sim \frac{r^{n+1}}{p}$, then

$$\begin{aligned} u_2(n) &\sim \left(\frac{1-\sqrt{p}}{q}\right) r^{n+3}, \\ u_3(n) &\sim \frac{6p}{q}\left(\frac{1}{r} - \frac{1}{\sqrt{p}}\right) r^{n+4}. \end{aligned}$$

## 8. Vertex flipping in graphs

In this section we will consider a modified version on our edge flipping process on a path $P_n$ with $n$ vertices. We call this a "vertex flipping" process on a graph. The *vertices* of the path are selected uniformly at random without replacement. If at some stage, the vertex $v$ is selected, then it and any of its unmarked neighbors are given the value 0 with probability $p$, and the value 1 with probability $q = 1 - p$. We will let $P(n)$ denote the probability that at the end of this process, all the vertices of the path have the value 0. As before, we will find a recurrence for $P(n)$, derive a first-order differential equation for the generating function for $P(n)$, solve the differential equation and from that deduce the asymptotic behavior for $P(n)$ as $n$ tends to infinity.

**Theorem 11.** *The generating function* $\mathbf{P}(z) = \sum_{n\geqslant 0} P(n) z^n$ *satisfies*:

$$\mathbf{P}(z) = \frac{\sqrt{p}}{\tanh(z\sqrt{p} + \frac{1}{2}\ln(\frac{1+\sqrt{p}}{1-\sqrt{p}})) - z\sqrt{p}}$$

*and*

$$P(n) \sim \frac{1}{\sqrt{p}} r_0^{n-1}$$

*where* $r_0$ *is the unique pole of* $\mathbf{P}$.

**Proof.** To begin, define $P(0)=1$, $P(1)=P(2)=p$. By examining what happens when each of the $n$ vertices of $P_n$ is selected first, we find by induction that $P(n)$ satisfies the recurrence:

$$P(n)=\frac{p}{n}\left(2P(n-2)+\sum_{k=0}^{n-3}P(k)P(n-3-k)\right),\quad n\geqslant 3. \tag{31}$$

To derive the generating function $\mathbf{P}(z)$, we multiply both sides of (31) by $nz^{n-1}$, and sum over $n\geqslant 3$.

$$\sum_{n\geqslant 3}nP(n)z^{n-1}=p\sum_{n\geqslant 3}2P(n-2)+p\sum_{n\geqslant 0}\sum_{k=0}^{n-3}P(k)P(n-3-k). \tag{32}$$

The left-hand side of (32) is just $\mathbf{P}'(z)-P(1)-2P(2)z$ (where $\mathbf{P}'(z)$ denotes the derivative of $\mathbf{P}(z)$ with respect to $z$). On the other hand, the right-hand side of (32) is easily seen to be

$$2pz\big(\mathbf{P}(z)-1\big)+pz^2\mathbf{P}(z)^2.$$

Combining these, we see that $\mathbf{P}(z)$ satisfies the differential equation

$$\mathbf{P}'(z)=p\big(z^2\mathbf{P}(z)^2+2z\mathbf{P}(z)+1\big)=p\big(z\mathbf{P}(z)+1\big)^2 \tag{33}$$

with the boundary condition $\mathbf{P}(0)=P(0)=1$.

Fortunately, (33) can be solved explicitly. The solution is

$$\mathbf{P}(z)=\frac{\sqrt{p}}{\tanh(z\sqrt{p}+\frac{1}{2}\ln(\frac{1+\sqrt{p}}{1-\sqrt{p}}))-z\sqrt{p}}. \tag{34}$$

The nearest pole of $\mathbf{P}(z)$ to the origin is the root $\gamma_0$ of the denominator of (34). Since $\gamma_0$ is a simple pole of $\mathbf{P}(z)$, then letting $r_0=\gamma_0^{-1}$, standard arguments in complex function theory imply that the coefficient $P(n)$ of $z^n$ in $\mathbf{P}(z)$ has the asymptotic behavior $P(n)\sim cr_0^n$ for some constant $c$. To determine the value of $c$, we multiply $\mathbf{P}(z)$ by $1-r_0z$ and take the limit as $z\to\gamma_0$. Thus,

$$\begin{aligned}
c&=\lim_{z\to\gamma_0}\frac{\sqrt{p}(1-zr_0)}{\tanh(z\sqrt{p}+\frac{1}{2}\ln(\frac{1+\sqrt{p}}{1-\sqrt{p}}))-z\sqrt{p}}\\
&=\lim_{z\to\gamma_0}\frac{\sqrt{p}(-r_0)}{\sqrt{p}\cosh^{-2}(z\sqrt{p}+\frac{1}{2}\ln(\frac{1+\sqrt{p}}{1-\sqrt{p}}))-\sqrt{p}}\\
&=\lim_{z\to\gamma_0}\frac{-r_0}{\tanh^2(z\sqrt{p}+\frac{1}{2}\ln(\frac{1+\sqrt{p}}{1-\sqrt{p}}))}\\
&=\frac{r_0}{(\gamma_0\sqrt{p})^2}=\frac{1}{r_0p}.
\end{aligned}$$

Hence, we conclude that

$$P(n)\sim\frac{1}{\sqrt{p}}r_0^{n-1}\quad\text{as } n\to\infty.\qquad\square \tag{35}$$

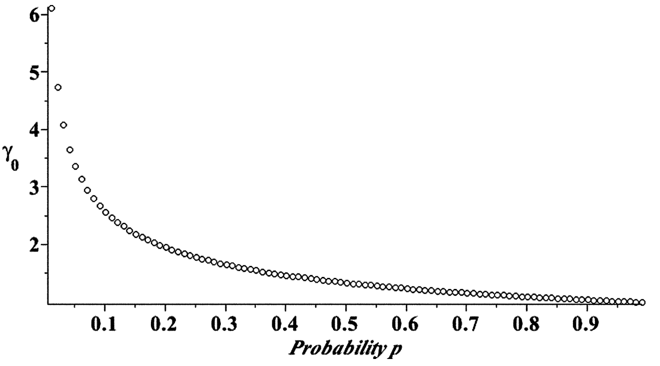


**Fig. 3.** The value of the root $\gamma_0$ as a function of $p$.

An easy corollary to this result is the asymptotic behavior for this process on the cycle $C_n$ on $n$ vertices. Each choice of a vertex of $C(n)$ leaves a path with $n-3$ vertices. By the symmetry of $C(n)$ and the previous result for paths, we see that the asymptotic probability that this vertex flipping process on $C(n)$ ends up with all vertices assigned the value 0 is just

$$n\frac{p}{n}\frac{r_0^n}{p}=r_0^n. \tag{36}$$

You couldn't ask for a nicer answer than that! Below, we show a plot of the behavior of $\gamma_0$ as a function of $p$ in Fig. 3.

For example, when $p=1/2$, $\gamma_0=1.34345\ldots$ and $r_0=0.74435\ldots$.

We should remark that we can generalize the vertex flipping process above so each selected vertex $v$ determines the values on all unmarked vertices at a distance of at most $t$ (the previous case is when $t=1$). If $P^{(t)}(n)$ denotes the probability that all $n$ vertices on a path $P_n$ end up with the value 0, and $\mathbf{P}^{(t)}(z)=\sum_{n\geqslant 0}P^{(t)}(n)z^n$, then it can be shown that $\mathbf{P}^{(t)}(z)$ satisfies the differential equation

$$\frac{d}{dz}\mathbf{P}^{(t)}(z)=p\left(\frac{z^t-1}{z-1}+z^t\mathbf{P}^{(t)}(z)\right)^2$$

with the initial condition $\mathbf{P}^{(t)}(0)=1$. Unfortunately, we can't write down a closed form solution to this differential equation when $t\geqslant 2$. However, for any particular value of $p$, we can numerically bound the behavior of the coefficients of the Taylor series for the solutions of this equation. For example, take $p=\frac{1}{2}$ and $t=2$, and let $y(z)$ denote the function $\mathbf{P}^{(2)}(z)$. Our equation is then

$$y'=\frac{1}{2}\left(1+z+z^2y\right)^2.$$

Let $z_\infty$ denote the least positive singularity of $y$. We can compute a lower bound for $z_\infty$ by subdividing the $x$-axis into intervals of size $\frac{1}{n}$ for some large $n$ and recursively upper bounding $y$ as follows. Suppose $y(\frac{k-1}{n})\leqslant b(\frac{k-1}{n})$ for some $k\geqslant 1$. Then for $\frac{k-1}{n}\leqslant z\leqslant\frac{k}{n}<z_\infty$, we have

$$y'\leqslant\frac{1}{2}\left(1+\frac{k}{n}+\frac{k^2}{n^2}y\right)^2$$

which implies

$$\frac{y'}{(1+\frac{k}{n}+\frac{k^2}{n^2}y)^2} \leqslant \frac{1}{2}.$$

Integrating this over the interval $\frac{k-1}{n} \leqslant z \leqslant \frac{k}{n}$, we obtain

$$\frac{1}{1+\frac{k}{n}+\frac{k^2}{n^2}y(\frac{k-1}{n})} - \frac{1}{1+\frac{k}{n}+\frac{k^2}{n^2}y(\frac{k}{n})} \leqslant \frac{k^2}{n^3}.$$

Hence,

$$\begin{aligned}\frac{1}{1+\frac{k}{n}+\frac{k^2}{n^2}y(\frac{k}{n})} &\geqslant \frac{1}{1+\frac{k}{n}+\frac{k^2}{n^2}y(\frac{k-1}{n})} - \frac{k^2}{n^3}\\ &\geqslant \frac{1}{1+\frac{k}{n}+\frac{k^2}{n^2}b(\frac{k-1}{n})} - \frac{k^2}{n^3}\end{aligned}$$

and

$$y\left(\frac{k}{n}\right) \leqslant \frac{n^2}{k^2}\left(\left(\frac{1}{1+\frac{k}{n}+\frac{k^2}{n^2}b(\frac{k-1}{n})} - \frac{k^2}{n^3}\right)^{-1} - 1 - \frac{k}{n}\right) = b\left(\frac{k}{n}\right).$$

Thus, we can iterate this recurrence for $b$ (starting with $b(0) = 1$) to get upper bounds on $y$ and consequently, lower bounds on $z_\infty$. For example, using $n = 10\,000$ and $p = 0.5$, we get the bound $y(1.2057) < 34\,129$, so that $z_\infty > 1.2057$.

To get an upper bound on $z_\infty$, we do the following. Suppose $\alpha$ satisfies $0 < \alpha < z_\infty$. Then

$$y' \geqslant \frac{1}{2}\left(1+\alpha+\alpha^2 y\right)^2$$

for $\alpha \leqslant z \leqslant z_\infty$ where as before, we are taking $p = \frac{1}{2}$. Thus, dividing and integrating we have

$$\int_{\alpha}^{z_\infty} \frac{y'\,dz}{(1+\alpha+\alpha^2 y)^2} \geqslant \frac{1}{2}\int_{\alpha}^{z_\infty} dz$$

which implies

$$\frac{1}{\alpha^2(1+\alpha+\alpha^2 y(\alpha))} \geqslant \frac{1}{2}(z_\infty - \alpha),$$

i.e.,

$$z_\infty \leqslant \alpha + \frac{2}{\alpha^2(1+\alpha+\alpha^2 y(\alpha))}. \tag{37}$$

Now, define $Q_n(z) = \sum_{k=0}^{n} P^{(2)}(k)z^k$. Hence, for any $n \geqslant 0$ and $0 \leqslant z \leqslant z_\infty$,

$$y(z) = \mathbf{P}^{(2)}(z) = \sum_{k \geqslant 0} P^{(2)}(k)z^k > Q_n(z).$$

Thus, by (37), we have

$$z_\infty < \alpha + \frac{2}{\alpha^2(1 + \alpha + \alpha^2 Q_n(\alpha))}. \tag{38}$$

Now taking $n = 500$ and $\alpha = 1.193$, we deduce from (38) that $z_\infty < 1.20598$. By taking larger values of $n$, we can get arbitrarily close to the (computed) value of $z_\infty = 1.20577584\ldots$.

## 9. The most likely and least likely color patterns

In this section, we focus on the problem "Which color pattern on the graph $G$ appears most frequently?" In other words, in the stationary distribution of the state graph associated with colorings of $G$, which state has the largest probability? A natural inclination would be to guess that the monochromatic configuration (having all vertices in the same color) is the best candidate. We first show that is indeed the case for the edge flipping process.

Assume now that $G$ is a graph (as usual, with no isolated vertices) on which there is some color pattern $P$. Observe that we can associate a unique labeling $\lambda = \lambda(P)$ of the edges of $G$ by assigning the label $\lambda(e) = D$ if the endpoints of $e$ have *different* colors, and $\lambda(e) = S$ if the endpoints have the *same* colors. In any such labeling, the number of $D$ labels in any cycle must be even, and in fact, this condition is also sufficient for the existence of a valid $D/S$ labeling of the edges of a graph (i.e., one that comes from some color pattern $P$ on the vertices of $G$). Note that for each valid labeling $\lambda$ there are in fact exactly two color patterns $P$ and $\tilde{P}$ on $G$ which generate $\lambda$, and these two color patterns differ only by a red/blue color (i.e., $0/1$) interchange. (If $G$ has $t$ connected components, then there are $2^t$ color patterns with generate $\lambda$.) Let us partition the vertex set $X$ of $G$ into two sets $V_0$ and $V_1$, where $V_i$ consists of the vertices of $G$ which have color $i$ in the pattern $P$, for $i = 0, 1$. Thus, all the edges of $G$ which lie entirely in some $V_i$ have the label $S$, while all those connecting points in different $V_i$ have the value $D$.

As we have seen, the probability that $P$ occurs is just the sum over all permutations $\pi$ of the deck which result in $P$ of the probability that $\pi$ occurs. It is clear that for any edge $e = \{u, v\}$ labeled $D$, the colors of both its endpoints must have been colored by cards in $\pi$ *before* either of the two cards associated with $e$ were selected. Let us now modify the coloring of $P$ to form a monochromatic coloring $P_0$ in which all the points are colored with 0, where we assume without loss of generality that $p \geqslant \frac{1}{2}$, so that for any edge, its 0-card is always as least likely as its 1-card to be selected. For each permutation $\pi$ that generates $P$, form a new permutation $\pi_0$ by replacing every card which gave a vertex in $V_1$ its color 1 (and so, must be a 1-card) by its 0-card mate. The result of applying $\pi_0$ will then colors all the points 0 and consequently, change all the $D$ labels to $S$. Since $p \geqslant \frac{1}{2}$, then $\pi_0$ is as least as likely to occur as $\pi$. Furthermore, there are additional permutations of the deck that generate $P_0$, since the 0-card for any of the cards initially labeled $D$ can now be selected much earlier than before, for example, as the first card. Hence, the probability that the pattern $P_0$ occurs is strictly larger than the probability that $P$ occurs. Hence, we have proved the following result.

**Theorem 12.** *The most probable color pattern for the edge flipping process on any graph $G$ is the monochromatic pattern in which every vertex is colored with the color that is most likely (i.e., if $p \geqslant \frac{1}{2}$ then this is color 0). When $p = \frac{1}{2}$ then each of the two monochromatic patterns is equally likely.*

One might now ask what the *least* likely pattern for a graph $G$ is. This seems to be a more difficult question in general. However, for one class of graphs, cycles of odd lengths, we can answer the question.

**Theorem 13.** *For $p = \frac{1}{2}$, the least likely color pattern for the edge flipping process on an odd cycle $C_{2n+1}$ consists of alternating* 0*'s and* 1*'s, except for one occurrence of two adjacent vertices with the same color.*

**Proof.** Suppose $P$ is some color pattern on $C_{2n+1}$. Label the edges by $D$ and $S$, as in the previous theorem, to form the edge labeling $\lambda$. Since $C_{2n+1}$ must have an even number of edges labeled $D$, there is at least one edge labeled $S$. If there is only one such edge, we are done. So we can assume that $P$ has at least three edges labeled $S$. We now are going to reverse the process we used in the previous result. Namely, we are going to replace two adjacent $S$'s by $D$'s to form a new edge labeling $\lambda'$. It follows from the preceding argument that the probability that $\lambda$ occurs is strictly greater than the probability that $\lambda'$ occurs. We now continue this process until we reach the desired color pattern, i.e., one with a single pair of identically colored vertices. This proves the theorem. □

It is not hard to show that the probability of reaching such a state on $C_{2n+1}$ is $\frac{1}{2(2n+1)!}$. Similarly, for a path with $n$ edges, the least likely pattern consists of alternating colors except for one pair of adjacent vertices at the end of the path colored the same. In this case, the probability of this occurring is $\frac{1}{2^n n!}$.

For general graphs $G$, it might seem that the least likely patterns would come from finding an appropriate maximum cut in $G$, and assigning the two colors to the two opposite sides of the cut. For example, for trees, this would result in a coloring which has just a single pair of adjacent vertices with the same color. (Of course, any possible pattern must always have at least one such pair.) Strictly speaking, patterns with *no* monochromatic pairs are the least likely, since they have probability 0 of occurring! However, there is still much to be done to clarify the situation.

In the remainder of this section, we will illustrate by examples that for general values of $p \neq \frac{1}{2}$, the situation for paths and cycles becomes more complicated. In particular, there is an interesting parity effect which takes place. Given some vertex $x_i$ in a large $n$-cycle, the probability that $x_i$ is assigned the color red at the end of the process is $p$. One might ask how does the color of $x_i$ affect the color assigned to a neighboring point $x_{i+1}$. In other words, if we know $x_i$ is red, what is the probability that $x_{i+1}$ is red? Thus, we want the conditional probability

$$\begin{aligned}
\Pr(x_{i+1} \text{ is red} \mid x_i \text{ is red}) &= \frac{\Pr(x_i \text{ is red and } x_{i+1} \text{ is red})}{\Pr(x_i \text{ is red})} \\
&= \frac{u(3)}{u(2)} \\
&= \frac{1}{3}(2p+1).
\end{aligned}$$

Hence, if $x_i$ ends up red, then the probability that its neighbor $x_{i+1}$ also ends up red is always greater than $1/3$, no matter how small $p$ is. This makes sense since one of the ways that $x_{i+1}$ ends up red is because of a card which assigned red to both $x_i$ and $x_{i+1}$. In fact when $p$ is very small, this is by far the most likely way that this can happen.

However, the situation changes dramatically if we know that the *two* left neighbors $x_i$ and $x_{i+1}$ of $x_{i+2}$ are red. Now we compute

$$\begin{aligned}
\Pr(x_{i+2} \text{ is red} \mid x_i \text{ and } x_{i+1} \text{ are red}) &= \frac{\Pr(x_i, x_{i+1} \text{ and } x_{i+2} \text{ are red})}{\Pr(x_i \text{ and } x_{i+1} \text{ are red})} \\
&= \frac{u(4)}{u(3)} \\
&= \frac{p(p+2)}{2p+1}
\end{aligned}$$

**Table 1**
Ratio of $u(k+2)$ to $u(k+1)$, $1 \leqslant k \leqslant 8$.

| $k$ | $\frac{u(k+2)}{u(k+1)}$ |
|---|---|
| 1 | 0.66666... |
| 2 | 0.62500... |
| 3 | 0.64000... |
| 4 | 0.63541... |
| 5 | 0.63700... |
| 6 | 0.63648... |
| 7 | 0.63666... |
| 8 | 0.63660... |

which goes to 0 as $p \to 0$. In particular, it is always more likely that a vertex $x$ is red if we know that its nearest neighbor is red than if we know that its two (left) neighbors are red!

But now with three neighbors, the situation flips back. In this case,

$$\Pr(x_{i+3} \text{ is red} \mid x_i,\ x_{i+1} \text{ and } x_{i+2} \text{ are red}) = \frac{u(5)}{u(4)} > \frac{1}{5}$$

for all $p$.

In general, it can be shown that

$$\Pr(x_{i+k} \text{ is red} \mid x_i, x_{i+1}, \dots x_{i+k-1} \text{ are red}) = \frac{u(k+2)}{u(k+1)}$$

where

$$\frac{u(k+2)}{u(k+1)} = \begin{cases} > \frac{1}{k+2} & \text{for all } p > 0 \text{ if } k \text{ is odd}, \\ \to 0 & \text{as } p \to 0 \text{ if } k \text{ is even}. \end{cases}$$

For example, when $p = 1/2$, we list in Table 1 the conditional probabilities $\frac{u(k+2)}{u(k+1)}$ for small values of $k$. The limiting ratio alternates about the limit $r = \frac{2}{\pi} = 0.6336197\dots$ depending on the parity of $k$.

Recall that $C(a,b)$ denotes the probability that a cycle on $a+b$ vertices has the first $a$ vertices assigned the value 0 (or red), and the remaining $b$ vertices assigned the value 1 (or blue). From (9) and (19) we find:

$$\begin{aligned} C(n,0) &\sim r^n, \\ C(n-1,1) &\sim \frac{1}{p}(-p+r)r^n, \\ C(n-2,2) &\sim \frac{1}{p}(1+p-2r)r^n, \\ C(n-3,3) &\sim \frac{1}{p}\big(q-(2+p)r+3r^2\big)r^{n-1}, \\ C(n-4,4) &\sim \frac{1}{p}\left(\frac{1}{3}q(2q+1)-2qr+(3+p)r^2-4r^3\right)r^{n-2}. \end{aligned}$$

This shows the impact of a small block of 1's in a cycle of otherwise all 0's. In particular, it illustrates some rather counterintuitive behavior in this situation. In particular, depending on the value of $p$, the relative order of the probabilities of these various patterns occurring changes quite a lot. In Fig. 4, we

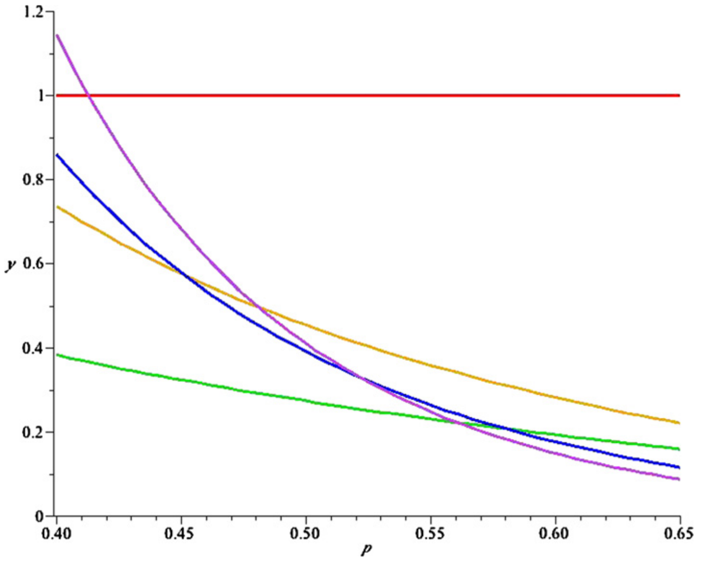


**Fig. 4.** Relative order of $C(n-k,k)$, $0 \leqslant k \leqslant 4$. The straight horizontal line represents $C(n,0)$ while the curved lines represent the quantities (from top to bottom) $C(n-4,4)$, $C(n-3,3)$, $C(n-2,2)$ and $C(n-1,1)$ in that order.

**Table 2**
Relative order of various $C(n-k,k)$.

| 0 | 0.155 | 0.340 | 0.381 | 0.413 | 0.451 | 0.481 | 0.522 | 0.561 | 0.582 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|
| 4 | 4 | 4 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 3 | 3 | 3 | 0 | 4 | 4 | 2 | 2 | 2 | 2 | |
| 2 | 2 | 0 | 3 | 3 | 2 | 4 | 3 | 3 | 1 | |
| 1 | 0 | 2 | 2 | 2 | 3 | 3 | 4 | 1 | 3 | |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 4 | 4 | |

show the relative asymptotic probabilities of these five configurations as $p$ varies from 0.3 to 0.65 (normalized so that the value $C(n,0)$ for the monochromatic pattern is set to 1).

In Table 2, we show 9 values of $p$ (going from 0 to 1), where transitions in the relative order of values of some of the $C(n-k,k)$ occur. The values on the top of Table 2 are the values of $p$ at which the transitions occur. The columns show the relative order of the values of $C(n-k,k)$ with smaller values below larger values. The integer $k$ in Table 2 denotes the pattern $C(n-k,k)$. Notice that $C(n-2,2)$ is always more probable than $C(n-1,1)$. This again is a parity effect. However, $C(n-4,4)$ can be more probable or less probable than $C(n-3,3)$, depending on the value of $p$.

## 10. A voter model and more related problems

We remark that a special case of our vertex-flipping problem is related to a well-studied voter model. Introduced by Clifford and Sudbury [8] and Holley and Liggett [12] independently, the voter model assumes that initially in a graph $G$ on $n$ vertices, each vertex (or player) is in one of $k$ colors, and in each round a player is chosen at random who then chooses a random neighbor and changes his/her color to the color of the neighbor. In [1], several proofs were given using coalescing random walks to show that it takes $O(n^3 \log n)$ rounds to converge to the state that every player has the same color.

In our interpretation, we have a set of stubborn, influential voters (for example, think of various local television news commentators). Whenever they broadcast their opinions, they influence the set of voters who happen to be listening. We assume that (most of) our (gullible?) voters are persuaded by what they last heard, perhaps with some probability $p$ (or $p_v$). Then it is not hard to see that this situation can be modeled by a random walk on an appropriate LRB.

It would be of interest to consider the following generalization of our process. For a given selection probability and a positive $\epsilon > 0$, if we execute the edge flipping process on a graph $G$ with $n$ vertices, what is the probability that a random state (according to the stationary distribution of the state graph) has at least $(1-\epsilon)n$ vertices of the same color?

Of course, we have barely scratched the surface in this paper for numerous problems suggested by the edge/vertex flipping process. In general, it would be of interest to understand the situation for general graphs $G$. For example, what structures within $G$ (such as having large cutsets) have the greatest effect on the stationary distributions in the state graph for various patterns? What happens in the hypergraph case, i.e., when the cards affect more than just two vertices? How does the situation change if we allow the selection probabilities to depend on the vertices (instead of being constant, as we have assumed in this paper)? For example, some voters might be more persuasive than others in changing their neighbors opinions. There are many attractive questions on this topic that remained unresolved, and we hope to return to some of them in the near future.

## Acknowledgment

The authors wish to thank an excellent referee for helpful comments in preparing the final version of the paper.

## References

[1] D. Aldous, J. Fill, Reversible Markov chains and random walks on graphs, draft, 1994.
[2] C. Athanasiatis, P. Diaconis, Functions of random walks on hyperplane arrangements, Adv. in Appl. Math. 45 (2010) 410–437.
[3] P. Balister, B. Bollobás, J.R. Johnson, M. Walters, Random majority percolation, Random Structures Algorithms 36 (2010) 315–340.
[4] T.P. Bidigare, P. Hanlon, D.N. Rockmore, A combinatorial description of the spectrum for the Tsetlin library and its generalization to hyperplane arrangements, Duke Math. J. 99 (1999) 135–174.
[5] K.S. Brown, Semigroups, rings and Markov chains, J. Theoret. Probab. 13 (2000) 871–938.
[6] K.S. Brown, Semigroup and ring theoretical methods in probability, in: Representations of Finite Dimensional Algebras and Related Topics in Lie Theory and Geometry, in: Fields Inst. Commun., vol. 40, Amer. Math. Soc., Providence, RI, 2004, pp. 3–26.
[7] K.S. Brown, P. Diaconis, Random walks and hyperplane arrangements, Ann. Probab. 26 (1998) 1813–1854.
[8] P. Clifford, A. Sudbury, A model for spatial conflict, Biometrika 60 (1973) 581–588.
[9] J.A. Fill, An exact formula for the move-to-front rule for self-organizing lists, J. Theoret. Probab. 9 (1996) 113–160.
[10] P. Flajolet, R. Sedgewick, Analytic Combinatorics, Cambridge University Press, Cambridge, 2009, xiv+810 pp.
[11] R.L. Graham, D.E. Knuth, O. Patashnik, Concrete Mathematics. A Foundation for Computer Science, second ed., Addison–Wesley Publishing Co., Reading, MA, 1994, xiv+657 pp.
[12] R.A. Holley, T.M. Liggett, Ergodic theorems for weakly interacting infinite systems and the voter model, Ann. Probab. 3 (1975) 643–663.
[13] M. Kearns, S. Suri, N. Montfort, An experimental study of the coloring problem on human subject networks, Science 313 (5788) (2006) 824–827.
[14] M. Kearns, J. Tan, Biased voting and the democratic primary problem, in: Lecture Notes in Comput. Sci., vol. 5385, Springer, New York, 2008, pp. 639–652.
[15] F. Klein-Barmen, Über eine weitere Versallgemeinerung des Verbandsbegriffes, Math. Z. 46 (1940) 472–480.
[16] R.M. Phatarfod, On the matrix occurring in a linear search problem, J. Appl. Probab. 28 (1991) 336–346.
[17] M.-P. Schützenberger, Sur certains treillis gauches, C. R. Acad. Sci. Paris 224 (1947) 776–778.
[18] S.P. Vadhan, The complexity of counting in sparse, regular and planar graphs, SIAM J. Comput. 31 (2002) 398–427.