FUNCTIONS OF RANDOM WALKS ON HYPERPLANE ARRANGEMENTS

CHRISTOS A. ATHANASIADIS AND PERSI DIACONIS

ABSTRACT. Many seemingly disparate Markov chains are unified when viewed as random walks on the set of chambers of a hyperplane arrangement. These include the Tsetlin library of theoretical computer science and various shuffling schemes. If only selected features of the chains are of interest, then the mixing times may change. We study the behavior of hyperplane walks, viewed on a subarrangement of a hyperplane arrangement. These include many new examples, for instance a random walk on the set of acyclic orientations of a graph. All such walks can be treated in a uniform fashion, yielding diagonalizable matrices with known eigenvalues, stationary distribution and good rates of convergence to stationarity.

1. INTRODUCTION

Many seemingly disparate Markov chains may be successfully studied by viewing them as random walks on the set of chambers of a hyperplane arrangement [9]. These include the Tsetlin library of theoretical computer science, a variety of walks on the hypercube and various shuffling schemes [13]. If only selected features of such a Markov chain are of interest (for instance, only a few sites on the hypercube or the relative ordering of the top few cards), then the mixing time may change. Following a suggestion of Uyemura Reyes [46], we study the behavior of hyperplane walks, viewed on subarrangements of a given hyperplane arrangement. This leads to new Markov chains which permit a full analysis. The following two examples illustrate our results and are used as running examples throughout.

Example 1.1 (Conquering Territory). Consider an $m \times n$ grid, with each node labeled with $+1$ or -1. At each stage, a node is chosen from a fixed probability distribution, then a neighborhood of this node is chosen and finally, all labels of the nodes in this neighborhood are changed to $+1$ or all are changed to -1, according to a specific distribution. As explained in Section 4, this Markov chain can be viewed as a hyperplane walk on the Boolean arrangement. Such walks were first studied in [9, 13] and include the classical Ehrenfest urn. The stationary distribution depends on the various probabilistic specifications but the theory of [9, 13], reviewed in Section 2, gives a useful description of this distribution, as well as of the eigenvalues and rates of convergence to stationarity.

Suppose now that only the labels of a few nodes (for instance, the four corners or the middle node) are of interest. Common sense suggests that the induced process on these nodes may converge to stationarity at a faster rate than the entire chain. For example, in the Ehrenfest urn with n particles, order of $n \log n$ steps are required to equilibrate on the full state space but order of n steps suffice for a few tagged particles. Further details and examples appear in the sequel.   □

Sometimes the induced chain is the object of direct interest, with the original chain opaque in the background. This is the case in our second example.

Example 1.2 (Acyclic Orientations). Let $\mathcal{G}$ be a simple undirected graph. A Markov chain on the set of acyclic orientations of $\mathcal{G}$ can be defined as follows: At each stage, a node v of $\mathcal{G}$ is chosen from a fixed probability distribution w and all edges of $\mathcal{G}$ incident to v are oriented inward, towards v. Under some mild assumptions on $\mathcal{G}$ and w, this is an ergodic Markov chain on the set of acyclic orientations of $\mathcal{G}$ with describable stationary distribution and eigenvalues and with good control on rates of convergence. It arises as the chain induced from the Tsetlin library on the braid arrangement, where the subarrangement is the graphical arrangement corresponding to $\mathcal{G}$. It also arises as a walk on the Boolean arrangement; see Section 4 for a detailed discussion.   □

This paper is organized as follows. Section 2 includes background on hyperplane walks and functions of a Markov chain, along with an overview of the basic examples of hyperplane walks on the Boolean and braid arrangements. Our main results appear in Section 3. Developing a suggestion in [46], the process induced from a hyperplane walk on the set of chambers of a subarrangement is considered. Although a function of a Markov chain is usually not Markov, it is shown that subarrangement processes are Markov chains. Moreover, the subarrangement chains are shown to be hyperplane walks in their own right. This implies that the whole tool kit of results for hyperplane walks [9, 13] is available. One striking feature of general hyperplane walks is that they have nonnegative real eigenvalues, although these chains are almost never symmetric or reversible. A purely combinatorial proof of this fact, as well as a new coupling proof of the basic theorem of [13], giving rates of convergence to stationarity, also appear in Section 3.

Sections 4 and 5 give applications of the general theory to the hyperplane walks of Section 2, treating Examples 1.1 and 1.2 in several variations. These include a variety of functions on the Tsetlin library and inverse a-shuffling Markov chains on the symmetric group, such as those assigning to a permutation the set of elements preceding a given entry in its linear representation, the descent set and the cyclic descent set. Most of the induced chains we study seem very different from their parents. As a byproduct of our considerations, we mention an interpretation for the unsigned coefficients of the chromatic polynomial of a graph as the multiplicities of the transition matrix of a certain natural Markov chain on the set of acyclic orientations of $\mathcal{G}$ (see Proposition 5.11). Section 6 briefly discusses extensions to random walks on semigroups.

2. Background

This section reviews the main results on hyperplane walks, develops needed examples (hypercube walks, Tsetlin library, inverse a-shuffles) and provides background on functions of a Markov chain.

2.1. **Hyperplane walks.** Let $\mathcal{A}$ be a hyperplane arrangement in $V = \mathbb{R}^n$, meaning a finite set of affine hyperplanes in V. The *intersection poset* of $\mathcal{A}$ is the set $\mathcal{L}_\mathcal{A} = \{\cap \mathcal{E} : \mathcal{E} \subseteq \mathcal{A}\}$, consisting of all affine subspaces of V which can be written as intersections of some of the hyperplanes of $\mathcal{A}$, partially ordered by reverse inclusion. The space V, corresponding to $\mathcal{E} = \varnothing$, is the minimum element of $\mathcal{L}_\mathcal{A}$.

The connected components of the space obtained from V by removing the union of the hyperplanes of $\mathcal{A}$ are called *chambers*. The *restriction* of $\mathcal{A}$ on an intersection subspace $W \in \mathcal{L}_\mathcal{A}$ is the hyperplane arrangement in W consisting of the intersections of W with the hyperplanes of $\mathcal{A}$ which are not parallel to (in particular, do not contain) W. The chambers of all such restricted arrangements are called *faces* of $\mathcal{A}$. Thus the chambers of $\mathcal{A}$ are exactly its faces of dimension n. We will denote by $\mathcal{C}_\mathcal{A}$ the set of chambers of $\mathcal{A}$ and by $\mathcal{F}_\mathcal{A}$ its set of faces. The elements of $\mathcal{F}_\mathcal{A}$ are the open cells of a regular cell decomposition of V; see [10, Section 2.1]. Figure 2.1 shows an arrangement of four hyperplanes (lines) in $\mathbb{R}^2$ which has ten chambers, thirteen one-dimensional faces (open line segments) and four zero-dimensional faces (points).
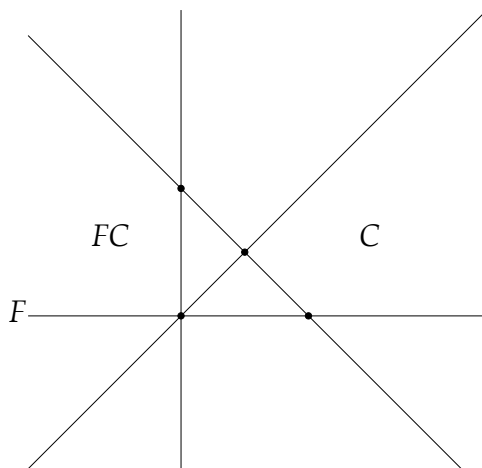
Given faces $F, G \in \mathcal{F}_\mathcal{A}$, we say that F is a face of G if F is contained in the closure of G, with respect to the standard topology on V. Given a chamber $C \in \mathcal{C}_\mathcal{A}$ and a face $F \in \mathcal{F}_\mathcal{A}$, a lemma of Tits [45] asserts that among all chambers of $\mathcal{A}$ which have F as a face, there is a unique chamber C' which is closest to C, in the sense that the set of hyperplanes in $\mathcal{A}$ separating C' from C is minimum with respect to inclusion. We will refer to this chamber C' as the *projection* of C on F and will denote it by FC. For an alternative definition, pick any points $x \in F$ and $y \in C$ and move by a small distance away from x in the direction of y, in the line segment joining these two points. The resulting point lies in a well defined

face of $\mathcal{A}$, which is the face FC. An example is given in Figure 2.1. The second recipe can be used to define the face $FG \in \mathcal{F}_\mathcal{A}$ for any two faces $F, G \in \mathcal{F}_\mathcal{A}$.

Using these ingredients, Bidigare, Hanlon and Rockmore [9] suggested the following family of Markov chains on the state space $\mathcal{C}_\mathcal{A}$. Start with a probability measure w on $\mathcal{F}_\mathcal{A}$. A step in the chain is given by choosing a face $F \in \mathcal{F}_\mathcal{A}$ from w and moving from the current chamber $C \in \mathcal{C}_\mathcal{A}$ to FC. Equivalently, we can describe this chain by defining its transition matrix K by the equation

$$(2.1) \qquad K(C, C') \;=\; \sum_{F \in \mathcal{F}_\mathcal{A}:\, FC=C'} w(F).$$

Henceforth, we will refer to this Markov chain as the *hyperplane walk* on $\mathcal{A}$ (or on $\mathcal{C}_\mathcal{A}$) associated to w.

Hyperplane walks being so general, it is surprising that there is a relatively complete theory for them. We recall that the *total variation distance* between two probability distributions P and Q on a finite set Ω is defined as

$$\|P - Q\|_{\mathrm{TV}} \;=\; \max_{A \subseteq \Omega} |P(A) - Q(A)|.$$

The measure w on $\mathcal{F}_\mathcal{A}$ is said to be *separating* [13] if for every $H \in \mathcal{A}$ there exists a face $F \in \mathcal{F}_\mathcal{A}$ such that $F \not\subseteq H$ and $w(F) > 0$. We combine the main results of [9, 13] into the following theorem.

Theorem 2.1. *Let $\mathcal{A}$ be a hyperplane arrangement in V with set of faces $\mathcal{F}_\mathcal{A}$ and intersection poset $\mathcal{L}_\mathcal{A}$ and let w be a probability measure on $\mathcal{F}_\mathcal{A}$.*

(i) *The characteristic polynomial of K is given by*

$$\det(xI - K) \;=\; \prod_{W \in \mathcal{L}_\mathcal{A}} (x - \lambda_W)^{m_W},$$

where

$$(2.2) \qquad \lambda_W \;=\; \sum_{F \in \mathcal{F}_\mathcal{A}:\, F \subseteq W} w(F)$$

is an eigenvalue,

$$m_W \;=\; |\mu_\mathcal{A}(V, W)| \;=\; (-1)^{\mathrm{codim}(W,V)} \mu_\mathcal{A}(V, W),$$

$\mu_\mathcal{A}$ is the Möbius function of $\mathcal{L}_\mathcal{A}$ and $\mathrm{codim}(W, V)$ is the codimension of W in V.

(ii) *The matrix K is diagonalizable.*

(iii) *K has a unique stationary distribution π if and only if w is separating.*

(iv) *Assume that w is separating and let K_C^l be the distribution of the chain started from the chamber C after l steps. Then its total variation distance from π satisfies*

$$(2.3) \qquad \|K_C^l - \pi\|_{\mathrm{TV}} \;\leq\; P\{F_1 F_2 \cdots F_l \notin \mathcal{C}_\mathcal{A}\},$$

where $(F_1, F_2, \dots)$ consists of independent and identically distributed picks from the measure w on $\mathcal{F}_\mathcal{A}$, and

$$(2.4) \qquad \|K_C^l - \pi\|_{\mathrm{TV}} \;\leq\; \sum_{H \in \mathcal{A}} \lambda_H^l.$$

Furthermore, there is a useful description of the stationary distribution which will not be detailed here; see [13, Theorem 2 (b)]. As noticed in [13, Section 4], the right-hand side of (2.3) is bounded from above by that of (2.4). The bounds in (2.3) and (2.4) are usually good but not perfect.

2.2. **Examples.** Our main examples involve the Boolean and braid arrangements. In what follows, we denote by $\mathfrak{S}_n$ the symmetric group of permutations of the set $[n] := \{1, 2, \dots, n\}$. We will often use the one line notation $(\tau(1), \tau(2), \dots, \tau(n))$ for a permutation $\tau \in \mathfrak{S}_n$. It will be helpful to think of τ as a linear ordering of a deck of n cards, labeled bijectively by the elements of $[n]$.

A. The Boolean arrangement. Let $\mathcal{A} = \{H_1, H_2, \dots, H_n\}$ be the *Boolean arrangement* in $V = \mathbb{R}^n$, where H_i is the coordinate hyperplane in $\mathbb{R}^n$ defined by the equation $x_i = 0$, for $1 \leq i \leq n$. The intersection poset $\mathcal{L}_\mathcal{A}$ is isomorphic to the lattice of subsets of the set $[n]$, ordered by inclusion, where the isomorphism maps a subspace $W \in \mathcal{L}_\mathcal{A}$ to the set of indices $i \in [n]$ for which $W \subseteq H_i$. The Möbius function $\mu_\mathcal{A}$ of this poset satisfies $\mu_\mathcal{A}(V, W) = (-1)^{\mathrm{codim}(W,V)}$ for $W \in \mathcal{L}_\mathcal{A}$.

The set $\mathcal{F}_\mathcal{A}$ of faces of $\mathcal{A}$ is in bijection with the set $\{-, 0, +\}^n$ of all 3^n possible sign vectors of length n (where the bijection is induced by the map assigning to each point $x \in \mathbb{R}^n$ the sequence of signs of the coordinates of x). The chambers of $\mathcal{A}$ are the orthants in $\mathbb{R}^n$; they correspond to the 2^n elements of $\{-, +\}^n$. Using these bijections, we may identify $\mathcal{C}_\mathcal{A}$ and $\mathcal{F}_\mathcal{A}$ with the sets $\{-, +\}^n$ and $\{-, 0, +\}^n$, respectively (the former may also be identified with the set of vertices of the n-dimensional cube $[-1, 1]^n$). The projection FC of a chamber $C \in \mathcal{C}_\mathcal{A}$ on a face $F \in \mathcal{F}_\mathcal{A}$ is the chamber whose sign vector is obtained from that of F by switching every zero coordinate to the corresponding coordinate of C. Thus, given a probability measure w on $\mathcal{F}_\mathcal{A}$, the hyperplane walk on $\mathcal{A}$ associated to w proceeds from the current chamber $C \in \mathcal{C}_\mathcal{A}$ by selecting a face $F \in \mathcal{F}_\mathcal{A}$ with probability $w(F)$ and replacing the ith coordinate of C by the ith coordinate of F, whenever the latter is nonzero, to reach the next chamber in the walk. Some examples of these walks appear in [9, Section 2.3] and [13, Section 3B].

It follows from Theorem 2.1 that the transition matrix K for this chain has eigenvalues

$$(2.5) \qquad \lambda_S \;=\; \sum_{F \in \mathcal{F}_S} w(F),$$

one for each $S \subseteq [n]$, where $\mathcal{F}_S$ denotes the set of faces of $\mathcal{A}$ whose sign vectors have their ith coordinate equal to zero for every $i \in S$, and that K has a unique stationary distribution π if and only if for every $1 \leq i \leq n$ there exists a face $F \in \mathcal{F}_\mathcal{A}$ with nonzero ith coordinate, such that $w(F) > 0$.

The eigenvalues which correspond to the hyperplanes H_i and which appear in the right-hand side of (2.4), are the numbers $\lambda_{\{i\}}$ for $1 \le i \le n$.

Example 2.2. For $\varepsilon \in \{-, +\}$ and $1 \le i \le n$, we denote by E_i^ε the face of $\mathcal{A}$ whose sign vector has its ith coordinate equal to ε and all other coordinates equal to zero. We choose face weights so that $w(E_i^\varepsilon) = w_i^\varepsilon$ for $\varepsilon \in \{-, +\}$ and $1 \le i \le n$, and $w(F) = 0$ for all other faces $F \in \mathcal{F}_\mathcal{A}$, where the w_i^ε are nonnegative real numbers summing to 1. The resulting Markov chain is the nearest neighbor random walk on the vertex set $\{-, +\}^n$ of the n-dimensional cube, which evolves by picking a coordinate i, at each stage, and switching it to ε, with probability w_i^ε. Thus, the transition matrix K for this chain is given by

$$
K(x, x') = \begin{cases} \sum_{i=1}^n w_i^{\varepsilon_i}, & \text{if } x' = x, \\ w_i^{-\varepsilon_i}, & \text{if } x' \text{ is obtained from } x \text{ by switching the } i\text{th coordinate to } -\varepsilon_i, \\ 0, & \text{otherwise} \end{cases}
$$

for $x, x' \in \{-, +\}^n$ with $x = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$. It has eigenvalues

$$
\lambda_S = \sum_{i \in [n] \smallsetminus S} w_i,
$$

one for each $S \subseteq [n]$, where $w_i = w_i^- + w_i^+$. Moreover, K has a unique stationary distribution π if and only if $w_i > 0$ holds for every $1 \le i \le n$. In that case, π is given by the formula

$$
(2.6) \qquad\qquad \pi(x) = \prod_{i=1}^n \frac{w_i^{\varepsilon_i}}{w_i}
$$

for $x = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) \in \{-, +\}^n$ (see, for instance, [13, Section 3B]) and (2.4) gives the bound

$$
(2.7) \qquad\qquad \|K_x^l - \pi\|_{\mathrm{TV}} \le \sum_{i=1}^n (1 - w_i)^l,
$$

where K_x^l is the distribution of the chain started from x after l steps. $\qquad\square$

B. The braid arrangement. Let $\mathcal{A}$ be the *braid arrangement* in $\mathbb{R}^n$, consisting of the $\binom{n}{2}$ hyperplanes defined by the equations $x_i - x_j = 0$ for $1 \le i < j \le n$. The intersection poset $\mathcal{L}_\mathcal{A}$ is isomorphic to the lattice of partitions of the set $[n]$, ordered by refinement. The isomorphism maps a subspace $W \in \mathcal{L}_\mathcal{A}$ to the partition of $[n]$ in which i and j are in the same block if and only if $x_i = x_j$ holds for every point $(x_1, x_2, \ldots, x_n) \in W$.

The set $\mathcal{C}_\mathcal{A}$ of chambers of $\mathcal{A}$ is in bijection with $\mathfrak{S}_n$ and the set $\mathcal{F}_\mathcal{A}$ of faces is in bijection with the set of ordered partitions of $[n]$, meaning set partitions $(B_1, B_2, \ldots, B_k)$ of $[n]$ in which the order of the blocks matters. To be specific, let us agree that the permutation $\tau \in \mathfrak{S}_n$ corresponds to the chamber

$$
x_{\tau(1)} > x_{\tau(2)} > \cdots > x_{\tau(n)}.
$$

More generally, the ordered partition $B = (B_1, B_2, \ldots, B_k)$ of $[n]$ corresponds to the face of $\mathcal{A}$ defined by the equalities $x_i = x_j$, whenever i and j lie in the same block of B, and the inequalities $x_i > x_j$, whenever the block of B which contains i has smaller index than the block of B which contains j.

The action of faces on chambers can be described as follows. Let $C \in \mathcal{C}_\mathcal{A}$ be the chamber corresponding to $\tau \in \mathfrak{S}_n$ and $F \in \mathcal{F}_\mathcal{A}$ be the face corresponding to the ordered partition B of $[n]$. One can then check that FC is the chamber which corresponds to the permutation of $[n]$ which is obtained from B by listing the elements of each block of B in the relative order in which they appear in τ. For instance, if $n = 9$, $\tau = (8, 1, 4, 9, 7, 2, 6, 3, 5)$ and $B = (\{6, 9\}, \{1, 3, 7\}, \{4\}, \{2, 5, 8\})$, then the resulting permutation is equal to $(9, 6, 1, 7, 3, 4, 8, 2, 5)$. In the sequel, we identify faces (respectively, chambers) of $\mathcal{A}$ with the corresponding ordered partitions (respectively, permutations) of the set $[n]$.

Different choices of probability measure on $\mathcal{F}_\mathcal{A}$ lead to various interesting Markov chains on $\mathfrak{S}_n$. We concentrate on the following two examples.

Example 2.3 (Tsetlin Library). Let $w_1, w_2, \ldots, w_n$ be nonnegative real numbers summing to 1. Choose face weights so that

$$w(B) = \begin{cases} w_i, & \text{if } B = (\{i\}, [n] \smallsetminus \{i\}) \text{ for some } 1 \le i \le n \\ 0, & \text{otherwise} \end{cases}$$

for an ordered partition B of $[n]$. The projection of τ on $B = (\{i\}, [n] \smallsetminus \{i\})$ removes the entry i in the one line notation of the permutation τ and places it in front. Hence the transition matrix K is the $n! \times n!$ matrix defined by

$$K(\tau, \tau') = \begin{cases} w_i, & \text{if } \tau' \text{ is obtained from } \tau \text{ by moving } i \text{ in front, for some } i \\ 0, & \text{otherwise.} \end{cases}$$

This chain has been extensively studied as a model of dynamic storage allocation; see [20] for a survey and [7] for recent results. It was shown by Phatarfod [32], and follows easily from Theorem 2.1 (see [9, 13]), that K is diagonalizable with eigenvalues

$$(2.8) \qquad \lambda_\tau = \sum_{\tau(i)=i} w_i,$$

one for each $\tau \in \mathfrak{S}_n$. Moreover, K has a unique stationary distribution π if and only if we have $w_i = 0$ for at most one index i. In that case, π is given by "sampling the weights without replacement to generate a random permutation". Thus we have

$$(2.9) \qquad \pi(\tau) = \frac{w_{\tau(1)} w_{\tau(2)} \cdots w_{\tau(n-1)}}{(1 - w_{\tau(1)})(1 - w_{\tau(1)} - w_{\tau(2)}) \cdots (1 - w_{\tau(1)} - \cdots - w_{\tau(n-2)})}$$

for $\tau \in \mathfrak{S}_n$. This stationary distribution is known as the *Luce model* in cognitive psychology; see [14, p. 174] for extensive references. Equation (2.4) of Theorem 2.1 (iv) gives the

bound

$$(2.10) \qquad \|K_\tau^l - \pi\|_{\mathrm{TV}} \ \leq \ \sum_{1 \leq i < j \leq n} (1 - w_i - w_j)^l$$

on the rate of convergence to stationarity, where K_τ^l is the distribution of the chain started at τ after l steps. Suppose, for instance, that $w_i = 1/n$ for $1 \leq i \leq n$, so that at each stage of the chain, an entry of the current permutation is selected uniformly at random and moved in front (thus the chain evolves by the "random to top" rule). Then we have

$$(2.11) \qquad \|K_\tau^l - \pi\|_{\mathrm{TV}} \ \leq \ \binom{n}{2} \left(1 - \frac{2}{n}\right)^l.$$

The expression on the right is bounded above by $e^{-2c}/2$ if $l \geq n(\log n + c)$, for $c > 0$. In this case there is a matching lower bound which shows that $n(\log n + c)$ steps are in fact necessary and sufficient for convergence to stationarity; see [16] for further details and more refined asymptotics. $\qquad \square$

Example 2.4 (Inverse a-shuffles). Ordinary riffle shuffles have received a careful analysis in [8]. A key to this analysis is a natural model on inverse riffle shuffles. Informally, begin with a deck of cards in order. Label the back of each card by one of the numbers in $\{1, 2, \ldots, a\}$, choosing the labels uniformly and independently. Then remove all cards labeled 1, keeping them in the same relative order, and start a new deck. Remove the cards labeled 2, keeping them in the same relative order, and place them below the ones labeled 1. Continue, placing the cards labeled a at the bottom. This can be seen as a random walk on the braid arrangement.

More formally, let $a \geq 2$ be an integer and $\mathcal{A}$ be the braid arrangement in $\mathbb{R}^n$, as before. A *weak ordered partition* of $[n]$ is a sequence of pairwise disjoint sets (called blocks) whose union is equal to $[n]$. From such a sequence one gets an ordered partition of $[n]$ by simply removing the empty blocks. We define a probability measure w on $\mathcal{F}_\mathcal{A}$ by first assigning weight equal to $1/a^n$ to each of the a^n weak ordered partitions $(B_1, B_2, \ldots, B_a)$ of $[n]$ with a blocks and then letting $w(B)$ equal the sum of the weights of all weak ordered partitions of $[n]$ with a blocks which correspond to the ordered partition B. For instance, if $a = 2$, then

$$w(B) \ = \ \begin{cases} 1/2^{n-1}, & \text{if } \ B = ([n]) \\ 1/2^n, & \text{if } \ B = (s, [n] \smallsetminus s) \text{ and } s \neq \varnothing, \ s \neq [n] \\ 0, & \text{otherwise} \end{cases}$$

for an ordered partition B of $[n]$.

The resulting chain on $\mathfrak{S}_n$ proceeds from a given permutation τ by selecting uniformly at random a weak ordered partition $(B_1, B_2, \ldots, B_a)$ of $[n]$ with a blocks and listing the elements of each block B_j in the relative order in which they appear in τ, to reach a new permutation τ' (such a permutation is said to be obtained from τ by an *inverse a-shuffle*). Equivalently, the transition matrix K of the chain satisfies

$$K(\tau, \tau') \ = \ \frac{\nu(\tau, \tau')}{a^n},$$

where $\nu(\tau, \tau')$ is the number of weak ordered partitions of $[n]$ with a blocks, the projection of τ on which is equal to τ'.

Let $W \in \mathcal{L}_\mathcal{A}$ be an intersection subspace of codimension $k = \text{codim}(W, V)$ and let σ be the corresponding partition of $[n]$, so that the number of blocks of σ is equal to $n - k$. Then the right-hand side of (2.2) is equal to the probability that the following holds for a random weak ordered partition B of $[n]$ with a blocks: for every pair $\{i, j\}$ of elements of $[n]$ belonging to the same block of σ, the elements i and j also belong to the same block of B. This probability is clearly equal to $1/a^k$ and hence

$$(2.12) \qquad \lambda_W = 1/a^{\text{codim}(W,V)}.$$

Thus it follows easily from Theorem 2.1 (i) (see [9, Equation (31)]) that the distinct eigenvalues of K are $1, 1/a, 1/a^2, \ldots, 1/a^{n-1}$ and that the multiplicity of the eigenvalue $1/a^i$ is equal to the number of permutations in $\mathfrak{S}_n$ which have exactly $n - i$ cycles. The stationary distribution π in this case is the uniform distribution on $\mathfrak{S}_n$ and (2.4) gives the bound

$$\|K_\tau^l - \pi\|_{\text{TV}} \leq \binom{n}{2} \left(\frac{1}{a}\right)^l.$$

The expression on the right is bounded above by $a^{-c}/2$ if $l \geq 2 \log_a n + c$ and $c > 0$. In fact $(3/2) \log_a n + c$ steps are necessary and sufficient for convergence to uniformity; see [8] for further details and asymptotics.

The chain of inverse a-shuffles converges to the uniform distribution at precisely the same rate as the chain of ordinary riffle shuffles on $\mathfrak{S}_n$. Thinking of the elements of $\mathfrak{S}_n$ as linear orderings of a deck of n cards, this chain proceeds from a given ordering as follows. The deck is cut into a (possibly empty) packets according to the multinomial distribution on their sizes. Then all a packets are riffled together, each time dropping a card from one of the a packets with probability proportional to its size, to get to a new ordering of the deck. For more information and extensive discussions, see [8, 14, 30]. $\qquad \square$

Examples 2.3 and 2.4 are two of the most interesting cases of general hyperplane walks. Other hyperplane arrangements for which the chambers are indexed by familiar combinatorial objects and the associated Markov chain has a reasonably down to earth description appear in [13, Section 3]. Further examples where the probabilistic analysis remains to be done can be found in [5, 33, 39] and [41, Lecture 5].

2.3. **Functions of a Markov chain.** Let $X_0, X_1, X_2, \ldots$ be the successive outcomes of a Markov chain on a finite state space $\mathfrak{X}$. Consider a finite set $\mathfrak{Y}$ and a surjective map $f : \mathfrak{X} \to \mathfrak{Y}$. We may think of $\mathfrak{Y}$ as a set partition of $\mathfrak{X}$ and of the map f as the canonical surjection. Thus $f(x)$ is equal to the unique block of $\mathfrak{Y}$ which contains x, for every $x \in \mathfrak{X}$. We set $Y_i = f(X_i)$ for each index i and refer to (Y_i) as the stochastic process (or chain) on the state space $\mathfrak{Y}$ which is *induced from* (X_i) by the map f.

A function of a Markov chain is usually not Markov. The following lemma gives a necessary and sufficient condition for Markovianity in the situation described above. We refer the reader to [26, Sections 6.3-6.4] for a good elementary treatment. For a more sophisticated treatment and references, see [35].

Lemma 2.5 (Dynkin's Criterion). *Let $(X_i)_{\geq 0}$ be a Markov chain on a finite state space $\mathfrak{X}$ and let $\mathfrak{Y}$ be a partition of $\mathfrak{X}$. The chain induced by the canonical surjection $f : \mathfrak{X} \to \mathfrak{Y}$ is Markov for all starting distributions for X_0 if and only if for any two distinct blocks $B_0, B_1 \in \mathfrak{Y}$, the probability $P(X_1 \in B_1 \mid X_0 = x_0)$ is constant in $x_0 \in B_0$.*

It is known that if the chain (X_i) is ergodic with stationary distribution π, then the induced chain (Y_i) has a limiting stationary distribution $\bar{\pi}$, given by

$$(2.13) \qquad \bar{\pi}(B) \;=\; \sum_{x \in B} \pi(x)$$

and one may inquire about rates of convergence to stationarity (even if the induced chain is not Markov). There has been considerable work on convergence rates in the situation of Example 2.3 (see [7, 21]) and in that of riffle shuffling (see [15] for a survey and [3] for some recent developments and references). Further work appears in Sections 4 and 5.

3. Main results

This section contains our main theoretical contribution. Following a suggestion of Uyemura Reyes [46], we show that the process which is induced from a hyperplane walk on the set of chambers of a subarrangement is a Markov chain which is itself a hyperplane walk, with transition matrix easily computable in terms of the original walk (Corollary 3.2). We also give a new proof of the description of the eigenvalues of hyperplane walks (part (i) of Theorem 2.1), which uses only basic enumerative combinatorics, and a new proof of the basic convergence theorem (part (iv) of Theorem 2.1), which is perhaps more transparent than the one given in [13].

Throughout this section, $\mathcal{A}$ is a hyperplane arrangement in the vector space $V = \mathbb{R}^n$ with set of chambers $\mathcal{C}_\mathcal{A}$ and set of faces $\mathcal{F}_\mathcal{A}$, $\mathcal{B} \subseteq \mathcal{A}$ is a subarrangement with set of chambers $\mathcal{C}_\mathcal{B}$ and set of faces $\mathcal{F}_\mathcal{B}$ and K is the transition matrix of the hyperplane walk on $\mathcal{A}$ associated to a given probability measure w on $\mathcal{F}_\mathcal{A}$. Our starting point is the observation that every chamber $C \in \mathcal{C}_\mathcal{A}$ is contained in a unique chamber of $\mathcal{B}$, which we denote by $\overline{C}$. Moreover, every chamber of $\mathcal{B}$ contains at least one chamber of $\mathcal{A}$. Thus there is a surjective map $f : \mathcal{C}_\mathcal{A} \to \mathcal{C}_\mathcal{B}$ defined by $f(C) = \overline{C}$ for $C \in \mathcal{C}_\mathcal{A}$ and hence the hyperplane walk on $\mathcal{A}$ associated to w induces a stochastic process on the state space $\mathcal{C}_\mathcal{B}$, in the sense of Section 2.3. The following proposition verifies Dynkin's criterion in this situation.

Proposition 3.1. *Let $D, D' \in \mathcal{C}_\mathcal{B}$ be chambers. If $C \in \mathcal{C}_\mathcal{A}$ is any chamber with $\overline{C} = D$, then the sum*

$$(3.1) \qquad Q(C, D') \;=\; \sum_{C' \in \mathcal{C}_\mathcal{A}:\ \overline{C'} = D'} K(C, C')$$

depends only on D and D' and not on the choice of C.

Proof. Replacing $K(C, C')$ by the right-hand side of (2.1), we find that

$$(3.2) \qquad Q(C, D') \;=\; \sum_{C' \in \mathcal{C}_\mathcal{A}:\ \overline{C'} = D'} \ \sum_{F \in \mathcal{F}_\mathcal{A}:\ FC = C'} w(F) \;=\; \sum_{F \in \mathcal{F}_\mathcal{A}:\ \overline{FC} = D'} w(F).$$

Let us denote by $\overline{F}$ the unique face of $\mathcal{B}$ which contains $F \in \mathcal{F}_{\mathcal{A}}$. It is easy to check that $\overline{FC} = \overline{F}\,\overline{C}$ holds for every $F \in \mathcal{F}_{\mathcal{A}}$. This observation and (3.2) imply that

$$(3.3) \qquad Q(C, D') \;=\; \sum_{F \in \mathcal{F}_{\mathcal{A}} : \overline{F}D = D'} w(F).$$

Clearly, the right-hand side of (3.3) is independent of the choice of C. $\qquad \square$

Corollary 3.2. *For every starting distribution on $\mathcal{C}_{\mathcal{A}}$, the stochastic process induced on $\mathcal{C}_{\mathcal{B}}$ from the hyperplane walk on $\mathcal{A}$ associated to w is Markov. Moreover, such an induced chain is itself a hyperplane walk on $\mathcal{B}$, with associated probability measure w^* on $\mathcal{F}_{\mathcal{B}}$ defined by*

$$(3.4) \qquad w^*(G) \;=\; \sum_{F \in \mathcal{F}_{\mathcal{A}} : F \subseteq G} w(F)$$

for $G \in \mathcal{F}_{\mathcal{B}}$.

Proof. The first statement follows from Proposition 3.1 and Lemma 2.5. The transition matrix K^* of the induced Markov chain on $\mathcal{C}_{\mathcal{B}}$ is given by the right-hand side of (3.1), so that

$$(3.5) \qquad K^*(D, D') \;=\; \sum_{C' \in \mathcal{C}_{\mathcal{A}} :\; C' \subseteq D'} K(C, C')$$

holds for $D, D' \in \mathcal{C}_{\mathcal{B}}$, where $C \in \mathcal{C}_{\mathcal{A}}$ is any of the chambers of $\mathcal{A}$ contained in D. Finally, we note that (3.3) can be rewritten as

$$(3.6) \qquad K^*(D, D') \;=\; \sum_{G \in \mathcal{F}_{\mathcal{B}} : GD = D'} w^*(G),$$

where $w^*(G)$ is as in (3.4). This proves the second statement in the corollary. $\qquad \square$

The next statement summarizes the main conclusions of our discussion.

Theorem 3.3. *Let $\mathcal{A}$ be a hyperplane arrangement in V with set of chambers $\mathcal{C}_{\mathcal{A}}$ and let w be a probability measure on its set of faces $\mathcal{F}_{\mathcal{A}}$. Let $\mathcal{B} \subseteq \mathcal{A}$ be a subarrangement with set of chambers $\mathcal{C}_{\mathcal{B}}$ and set of faces $\mathcal{F}_{\mathcal{B}}$ and let K^* be the transition matrix of the Markov chain on $\mathcal{B}$ induced from the hyperplane walk on $\mathcal{A}$ associated to w.*

(i) *The characteristic polynomial of K^* is given by*

$$\det(xI - K^*) \;=\; \prod_{W \in \mathcal{L}_{\mathcal{B}}} (x - \lambda_W)^{m_W^*},$$

where $\mathcal{L}_{\mathcal{B}}$ is the intersection poset of $\mathcal{B}$, λ_W is as in (2.2),

$$m_W^* \;=\; |\mu_{\mathcal{B}}(V, W)| \;=\; (-1)^{\mathrm{codim}(W,V)} \mu_{\mathcal{B}}(V, W)$$

and $\mu_{\mathcal{B}}$ is the Möbius function of $\mathcal{L}_{\mathcal{B}}$.
(ii) *The matrix K^* is diagonalizable.*

(iii) K^* has a unique stationary distribution $\bar{\pi}$ if and only if for every $H \in \mathcal{B}$ there exists a face $F \in \mathcal{F}_\mathcal{A}$ such that $F \not\subseteq H$ and $w(F) > 0$. Moreover, if w is separating, so that the stationary distribution π of the hyperplane walk on $\mathcal{A}$ also exists, then we have

(3.7)
$$\bar{\pi}(D) \ = \ \sum_{C \in \mathcal{C}_\mathcal{A}: C \subseteq D} \pi(C)$$

for every chamber $D \in \mathcal{C}_\mathcal{B}$.

(iv) Assume that $\bar{\pi}$ exists and let $(K_D^*)^l$ be the distribution of the induced chain started from the chamber $D \in \mathcal{C}_\mathcal{B}$ after l steps. Then its total variation distance from $\bar{\pi}$ satisfies

(3.8)
$$\|(K_D^*)^l - \bar{\pi}\|_{\mathrm{TV}} \ \leq \ P\left(F_1 F_2 \cdots F_l \subseteq \bigcup_{H \in \mathcal{B}} H\right),$$

where $(F_1, F_2, \dots)$ consists of independent and identically distributed picks from the measure w on $\mathcal{F}_\mathcal{A}$, and

(3.9)
$$\|(K_D^*)^l - \bar{\pi}\|_{\mathrm{TV}} \ \leq \ \sum_{H \in \mathcal{B}} \lambda_H^l.$$

Proof. Let w^* be as in Corollary 3.2. By Corollary 3.2 and Theorem 2.1, the characteristic polynomial of K^* is given by the expression suggested in part (i), provided that λ_W is replaced by

(3.10)
$$\lambda_W^* \ = \ \sum_{G \in \mathcal{F}_\mathcal{B}: G \subseteq W} w^*(G)$$

for every $W \in \mathcal{L}_\mathcal{B}$. Since every face $G \in \mathcal{F}_\mathcal{B}$ is partitioned by the faces $F \in \mathcal{F}_\mathcal{A}$ contained in G, it follows from (3.4) that the right-hand sides of (2.2) and (3.10) coincide. Hence we have $\lambda_W^* = \lambda_W$ for every $W \in \mathcal{L}_\mathcal{B}$ and part (i) follows. The remaining parts are direct consequences of the combination of Corollary 3.2 with Theorem 2.1.   □

We now turn to our new proofs of parts (i) and (iv) of Theorem 2.1. The proof of part (i) is motivated by the combinatorial method used in [4] to determine the eigenvalues of some matrices, with rows and columns indexed by the oriented rooted spanning trees of a directed graph. A related argument which proves diagonalizability is given in [11, Section 8.1] [12, Section 3.4]. We denote by $\mathrm{tr}(A)$ the trace of a $p \times p$ matrix $A = (a_{ij})$, so that

(3.11)
$$\mathrm{tr}(A^l) \ = \ \sum_{i=1}^{p} \sum_{1 \leq i_1, \dots, i_{l-1} \leq p} a_{ii_1} a_{i_1 i_2} \cdots a_{i_{l-1} i}$$

holds for every positive integer l. The method of [4] is based on the following elementary lemma.

Lemma 3.4. *Let $A = (a_{ij})$ be a $p \times p$ matrix with complex entries and let $\lambda_1, \lambda_2, \dots, \lambda_p$ be complex numbers. If $\mathrm{tr}(A^l) = \lambda_1^l + \lambda_2^l + \cdots + \lambda_p^l$ holds for every positive integer l, then $\lambda_1, \lambda_2, \dots, \lambda_p$ are the eigenvalues of A.*

Proof. We note that $\mathrm{tr}(A^l) = \mu_1^l + \mu_2^l + \cdots + \mu_p^l$ holds for every positive integer l, where $\mu_1, \mu_2, \ldots, \mu_p$ are the eigenvalues of A. It follows from this fact, our hypothesis and [4, Lemma 2.1] that the λ_i are a permutation of the μ_j. This proves the lemma. $\qquad\square$

Proof of Theorem 2.1 (i). By Lemma 3.4, it suffices to show that

$$\mathrm{tr}(K^l) = \sum_{W \in \mathcal{L}_\mathcal{A}} m_W (\lambda_W)^l$$

holds for every positive integer l. Using the definition of K, we see that for this matrix (3.11) can be rewritten as

$$(3.12) \qquad \mathrm{tr}(K^l) = \sum_{C \in \mathcal{C}_\mathcal{A}} \sum_{F_1 F_2 \cdots F_l C = C} w(F_1) w(F_2) \cdots w(F_l),$$

where the inner sum ranges over all sequences $(F_1, F_2, \ldots, F_l)$ of elements of $\mathcal{F}_\mathcal{A}$ of length l satisfying $F_1 F_2 \cdots F_l C = C$. Next we claim that for every $F \in \mathcal{F}_\mathcal{A}$ we have

$$(3.13) \qquad \#\{C \in \mathcal{C}_\mathcal{A} : FC = C\} = \sum_{W \in \mathcal{L}_\mathcal{A}:\, F \subseteq W} |\mu_\mathcal{A}(V, W)| = \sum_{W \in \mathcal{L}_\mathcal{A}:\, F \subseteq W} m_W.$$

Indeed, for a chamber $C \in \mathcal{C}_\mathcal{A}$ we have $FC = C$ if and only if F lies in the closure of C. The chambers of $\mathcal{A}$ with this property are in a one to one correspondence with the chambers of the subarrangement of $\mathcal{A}$ consisting of those hyperplanes which contain F. Thus (3.13) follows from Zaslavsky's formula [41, Theorem 2.5] [47] for the number of chambers of this subarrangement. Using equations (3.12) and (3.13) we find that

$$
\begin{aligned}
\mathrm{tr}(K^l) &= \sum_{C \in \mathcal{C}_\mathcal{A}} \sum_{F \in \mathcal{F}_\mathcal{A}:\, FC = C} \sum_{F_1 F_2 \cdots F_l = F} w(F_1) w(F_2) \cdots w(F_l) \\
&= \sum_{F \in \mathcal{F}_\mathcal{A}} \#\{C \in \mathcal{C}_\mathcal{A} : FC = C\} \sum_{F_1 F_2 \cdots F_l = F} w(F_1) w(F_2) \cdots w(F_l) \\
&= \sum_{F \in \mathcal{F}_\mathcal{A}} \sum_{W \in \mathcal{L}_\mathcal{A}:\, F \subseteq W} m_W \sum_{F_1 F_2 \cdots F_l = F} w(F_1) w(F_2) \cdots w(F_l) \\
&= \sum_{W \in \mathcal{L}_\mathcal{A}} m_W \sum_{F_1 F_2 \cdots F_l \subseteq W} w(F_1) w(F_2) \cdots w(F_l) \\
&= \sum_{W \in \mathcal{L}_\mathcal{A}} m_W \sum_{F_1 \cup \cdots \cup F_l \subseteq W} w(F_1) w(F_2) \cdots w(F_l) \\
&= \sum_{W \in \mathcal{L}_\mathcal{A}} m_W \left(\sum_{F \subseteq W} w(F) \right)^l = \sum_{W \in \mathcal{L}_\mathcal{A}} m_W (\lambda_W)^l,
\end{aligned}
$$

as desired. $\qquad\square$

Theorem 2.1 (iv) is proved in [13] by considering backward iteration. The following coupling proof is perhaps more transparent. For background on coupling we refer the reader to [14, p. 84] [29, Chapter 5]. We recall that the probability measure w on the set of faces of $\mathcal{A}$ is assumed to be separating. As before, $\mathcal{C}_\mathcal{A}$ is the set of chambers of $\mathcal{A}$.

Proof of Theorem 2.1 (iv). Consider two Markov chains (X_i) and (Y_i) evolving on $\mathcal{C}_\mathcal{A}$ as follows. The first chain starts at $X_0 = C$ and the second starts with Y_0 chosen from the stationary distribution π. At time i the face F_i is chosen from w and is used to upgrade both chains; thus $X_i = F_i X_{i-1}$ and $Y_i = F_i Y_{i-1}$. Let T be the first time l that the hyperplanes of $\mathcal{A}$ have been separated by $\{F_1, F_2, \ldots, F_l\}$, meaning that for every $H \in \mathcal{A}$ there exists an index $1 \leq i \leq l$ such that $F_i \not\subseteq H$. We claim that at this time we have $X_T = Y_T$. It is then clear that $X_j = Y_j$ has to hold for all $j \geq T$. Thus the two chains are coupled and (2.3) follows from the basic coupling inequality [14, p. 84] [29, Chapter 5]. Since (2.4) follows easily from (2.3) (see [13, p. 1839]), it remains to prove the claim.

Consider any hyperplane $H \in \mathcal{A}$ and choose an index $1 \leq i \leq T$ so that $F_i \not\subseteq H$. Then both chambers $X_i = F_i X_{i-1}$ and $Y_i = F_i Y_{i-1}$ lie in the same open half-space of V determined by H as F_i. Therefore these chambers lie in the same open half-space of V determined by H. It follows by induction on j that the same holds for X_j and Y_j for all $j \geq i$ and thus for $j = T$ as well. We have shown that for every $H \in \mathcal{A}$, the chambers X_T and Y_T lie in the same open half-space of V determined by H. Clearly any two such chambers must be equal. This proves the claim and completes the proof. $\square$

Remark 3.5. As was the case in [13], the argument in the previous proof does not require that faces are chosen independently from the same distribution. Any stationary process works as well. Nonstationary choices of face weights may be similarly handled. Then there may not be a stationary distribution and one needs to study "merging" [36].

4. APPLICATIONS TO HYPERCUBE WALKS

Throughout this section, $\mathcal{A}$ stands for the Boolean arrangement in $\mathbb{R}^N$ for some N, to be specified in each case. Specializing the choice of face weights and subarrangement gives a variety of natural examples. Part A treats the Ehrenfest urn of statistical mechanics. A spatial process driven by neighborhood attacks is studied in Part B. Part C gives a first treatment of the acyclic orientations chain (Example 1.2 in the introduction); the results are summarized in Corollary 4.1.

A. Ehrenfest Urn. Consider the Markov chain of Example 2.2 with weights $w_i^\varepsilon = 1/2n$ for all $\varepsilon \in \{-, +\}$ and $1 \leq i \leq n$. This is the usual nearest neighbor random walk on the n-dimensional cube with holding $1/2$, also known as Ehrenfests' urn. The transition matrix K has eigenvalues j/n with multiplicity $\binom{n}{j}$, for $0 \leq j \leq n$, and a uniform stationary distribution π. This walk has a small literature of its own, reviewed in [14, p. 19] [29, Section 2.3]. As explained there, the mixing time is $\frac{1}{2} n \log n$. The slightly less accurate bound

$$\|K_x^l - \pi\|_{\text{TV}} \; \leq \; n \left(1 - \frac{1}{n}\right)^l$$

follows from (2.7) and shows that the total variation distance on the left is bounded above by e^{-c} if $l \geq n(\log n + c)$. To illustrate the speedup possible for a subarrangement walk in this case, consider the subarrangement $\mathcal{B} = \{H_1, H_2, \ldots, H_k\}$ of the Boolean arrangement $\mathcal{A}$ in $\mathbb{R}^n$. The induced walk is a Markov chain on the set $\{-,+\}^k$. Theorem 3.3 implies that its transition matrix K^* has eigenvalues $(n-j)/n$ with multiplicity $\binom{k}{j}$, for $0 \leq j \leq k$, and a uniform stationary distribution $\bar{\pi}$. Equation (3.9) gives

$$\|(K_y^*)^l - \bar{\pi}\|_{\mathrm{TV}} \;\leq\; k \left(1 - \frac{1}{n}\right)^l$$

and hence the expression on the left is bounded above by e^{-c} if $l \geq n(\log k + c)$.

B. Neighborhood Attacks. Let $\mathcal{G}$ be a (finite, undirected) simple graph on the node set $[n]$. Each node of $\mathcal{G}$ is labeled with either $+$ or $-$. A Markov chain on the set $\{-,+\}^n$ of all 2^n possible labelings proceeds as follows. At each stage, a node of $\mathcal{G}$ is chosen uniformly at random. The labels of this node and of its neighbors are all changed to $+$ or all changed to $-$, with probability $1/2$. The transition matrix K for this chain satisfies

$$K(x, x') \;=\; \frac{\mu(x, x')}{2n}$$

for $x, x' \in \{-,+\}^n$, where $\mu(x, x')$ is the number of pairs (i, ε) of nodes $i \in [n]$ and signs $\varepsilon \in \{-,+\}$ for which x' is obtained from x by changing the labels of i and its neighbors in $\mathcal{G}$ to ε. Clearly, this is the chain defined by the hyperplane walk on the Boolean arrangement $\mathcal{A}$ in $\mathbb{R}^n$ for the following choice of face weights. For each node $i \in [n]$ and $\varepsilon \in \{-,+\}$ we denote by F_i^ε the face of $\mathcal{A}$ whose sign vector has j-coordinate equal to ε, if j is a neighbor of i in $\mathcal{G}$ or $j = i$, and equal to 0 otherwise. We define $w(F)$ as $1/2n$ times the number of pairs (i, ε) of nodes $i \in [n]$ and signs $\varepsilon \in \{-,+\}$ for which $F_i^\varepsilon = F$ (note that we may have $F_i^\varepsilon = F_j^\varepsilon$ for distinct nodes $i, j \in [n]$). Ehrenfests' urn occurs as the special case in which $\mathcal{G}$ has no edges.

For $S \subseteq [n]$ we denote by $\alpha(S)$ the number of nodes of $\mathcal{G}$ which are neither equal nor adjacent to any of the nodes in S. It follows from (2.5) that K has eigenvalues j/n, with multiplicity equal to the number of subsets $S \subseteq [n]$ with $\alpha(S) = j$, for $0 \leq j \leq n$ and that for $1 \leq i \leq n$, the eigenvalue contributed by the hyperplane H_i of $\mathcal{A}$ is equal to $1 - (d_i + 1)/n$, where d_i is the degree of node i in $\mathcal{G}$. The stationary distribution π for this example exists for every graph $\mathcal{G}$ but is hard to compute in general. Inequality (2.3) bounds the total variation distance $\|K_x^l - \pi\|_{\mathrm{TV}}$ from above by the probability that

$$\bigcup_{i=1}^{l} N(v_i) \;\neq\; [n],$$

where nodes $v_1, v_2, \ldots, v_l$ are picked independently and uniformly from $[n]$ and $N(v)$ stands for the set of nodes of $\mathcal{G}$ which are either adjacent or equal to v. To compute this probability is a classical problem, called the "coverage problem"; see, for instance, [2, 6, 27]. Similarly,

the eigenvalue bound (2.4) becomes

$$(4.1) \qquad \|K_x^l - \pi\|_{\mathrm{TV}} \ \leq\ \sum_{i=1}^{n} \left(1 - \frac{d_i + 1}{n} \right)^l.$$

For instance, if $\mathcal{G}$ is the complete graph on the node set $[n]$, then $d_i = n - 1$ for all i and the walk becomes random after exactly one step.

The eigenvalue bound is not perfect. For instance, consider a "star graph", having one central node of degree $n - 1$, and $n - 1$ leaves of degree one. The right-hand side of (4.1) becomes $(n-1)(1-2/n)^l$ and shows that order of $n \log n$ steps suffice. On the other hand, the coverage bound is bounded above by $(1 - 1/n)^l$, which is the chance of missing the central node in the first l steps. This implies that order of n steps suffice. An elementary argument shows that this is the correct answer. For a general graph $\mathcal{G}$, (4.1) implies that $\|K_x^l - \pi\|_{\mathrm{TV}} \leq e^{-c}$ if $l \geq \frac{n}{d+1}(\log n + c)$, where d is the largest of the degrees d_i.

To estimate the time it takes for a subset of nodes, say $\{1, 2, \ldots, k\}$, to equilibriate, consider the subarrangement $\mathcal{B} = \{H_1, H_2, \ldots, H_k\}$ of $\mathcal{A}$ and note that (3.9) becomes

$$\|(K_y^*)^l - \bar\pi\|_{\mathrm{TV}} \ \leq\ \sum_{i=1}^{k} \left(1 - \frac{d_i + 1}{n} \right)^l.$$

This offers a range of possibilities to illustrate the speedup possible; we leave further details and examples to the interested reader. One can also deduce easily from Theorem 3.3 that the transition matrix K^* of the induced chain has eigenvalues $\alpha(S)/n$, one for each $S \subseteq [k]$.

The previous situation can be easily varied; the nodes can be chosen with varying probability, the size and shape of the neighborhood can be allowed to fluctuate and the chance of $+$ or $-$ need not be symmetric. With such freedom, the stationary distribution becomes intractable but it is still staightforward to give upper bounds for the total variation distance to stationarity. Lower bounds are harder to achieve, due to our lack of knowledge of the stationary distribution.

C. Orientations. Let $\mathcal{G}$ be a (finite, undirected) simple graph on the node set $[n]$ with m edges. An *orientation* of $\mathcal{G}$ is an assignment of a direction $i \to j$ or $j \to i$ to each edge $\{i, j\}$ of $\mathcal{G}$. We will denote by $\mathrm{O}(\mathcal{G})$ the set of all orientations of $\mathcal{G}$. This set is in bijection with $\{-, +\}^m$ and hence with the set of chambers of the Boolean arrangement $\mathcal{A}$ in $\mathbb{R}^m$. To be more specific, let $E_{\mathcal{G}} = \{e_1, e_2, \ldots, e_m\}$ be the set of edges of $\mathcal{G}$, equipped with a fixed linear ordering of its elements, and let us identify an orientation $\mathfrak{o} \in \mathrm{O}(\mathcal{G})$ with the sign vector $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m) \in \{-, +\}^m$ for which

$$\varepsilon_k \ = \ \begin{cases} -, & \text{if } e_k \text{ is directed as } i \to j \text{ in } \mathfrak{o} \text{ and } i < j \\ +, & \text{if } e_k \text{ is directed as } i \to j \text{ in } \mathfrak{o} \text{ and } i > j \end{cases}$$

for $1 \leq k \leq m$, where $e_k = \{i, j\}$. Thus any hyperplane walk on $\mathcal{A}$ defines a Markov chain on $\mathrm{O}(\mathcal{G})$. A choice of face weights which gives Example 1.2 of the introduction is the following. Let $w_1, w_2, \ldots, w_n$ be nonnegative real numbers summing to 1. For $1 \leq i \leq n$, we denote by F_i the face of $\mathcal{A}$ whose sign vector has kth coordinate equal to $+$, if $e_k = \{i, j\}$

with $i < j$, to $-$, if $e_k = \{i, j\}$ with $i > j$ and to 0, if e_k is not incident to i. We let $w(F_i) = w_i$ for each node $i \in [n]$ which is not isolated in $\mathcal{G}$ and $w(F) = 0$ for all other nonzero faces of $\mathcal{A}$, and define $w(F)$ as the sum of w_i over all isolated nodes $i \in [n]$ of $\mathcal{G}$, if F is the zero face of $\mathcal{A}$.

The resulting chain on $\mathrm{O}(\mathcal{G})$ proceeds from a given orientation by selecting the node i of $\mathcal{G}$ with probability w_i and reorienting all edges of $\mathcal{G}$ incident to this node towards itself, to reach a new orientation of $\mathcal{G}$, leaving the orientations of all other edges of $\mathcal{G}$ unchanged. Equivalently, the transition matrix K of this chain on $\mathrm{O}(\mathcal{G})$ satisfies

$$(4.2) \qquad K(\mathfrak{o}, \mathfrak{o}') = \sum_{i \in [n]: \, \rho_i(\mathfrak{o}) = \mathfrak{o}'} w_i$$

for $\mathfrak{o}, \mathfrak{o}' \in \mathrm{O}(\mathcal{G})$, where $\rho_i(\mathfrak{o})$ denotes the orientation of $\mathcal{G}$ obtained from $\mathfrak{o}$ by reorienting towards i, as just described. We collect the consequences of Theorem 2.1 for this example in the following statement. We denote the stationary distribution by Π to avoid confusion with the notation of Section 5, where acyclic orientations of $\mathcal{G}$ are considered and π has a different meaning. A subset T of the set of nodes of $\mathcal{G}$ is said to be *dominating* in $\mathcal{G}$ if every edge of $\mathcal{G}$ is incident to at least one node in T.

Corollary 4.1. *Let $\mathcal{G}$ be a simple graph on the node set $[n]$ and let $E_{\mathcal{G}} = \{e_1, e_2, \ldots, e_m\}$ be the set of edges of $\mathcal{G}$. The following hold for the chain (4.2) on the set $\mathrm{O}(\mathcal{G})$ of orientations of $\mathcal{G}$:*

(i) *The matrix K is diagonalizable with eigenvalues*

$$(4.3) \qquad \lambda_S = \sum_{i \in N_S} w_i,$$

one for each $S \subseteq [m]$, where N_S is the set of nodes $i \in [n]$ which do not belong to any of the edges $e_k \in E_{\mathcal{G}}$ with $k \in S$.

(ii) *K has a unique stationary distribution Π if and only if there is no edge $\{i, j\} \in E_{\mathcal{G}}$ such that $w_i = w_j = 0$.*

(iii) *Assume that Π exists and let $K_{\mathfrak{o}}^l$ be the distribution of the chain started from the orientation $\mathfrak{o} \in \mathrm{O}(\mathcal{G})$ after l steps. We have*

$$(4.4) \qquad \|K_{\mathfrak{o}}^l - \Pi\|_{\mathrm{TV}} \leq P\left(\{v_1, v_2, \ldots, v_l\} \text{ is not dominating in } \mathcal{G}\right),$$

where $(v_1, v_2, \ldots)$ consists of independent and identically distributed picks from w, and

$$(4.5) \qquad \|K_{\mathfrak{o}}^l - \Pi\|_{\mathrm{TV}} \leq \sum_{\{i,j\} \in E_{\mathcal{G}}} (1 - w_i - w_j)^l.$$

In particular, we have

$$(4.6) \qquad \|K_{\mathfrak{o}}^l - \Pi\|_{\mathrm{TV}} \leq m\left(1 - \frac{2}{n}\right)^l$$

if $w_1 = \cdots = w_n = 1/n$.

Proof. Parts (i) and (ii) follow directly from Theorem 2.1 and the relevant discussion in Section 2.2. Part (iii) follows from Theorem 2.1 (iv), since a product of faces of $\mathcal{A}$ of the form F_v is a chamber if and only if the corresponding set of nodes v is dominating in $\mathcal{G}$ and since $\lambda_{\{e_k\}} = 1 - w_i - w_j$ is the eigenvalue corresponding to the hyperplane $x_k = 0$ of $\mathcal{A}$ associated to the edge $e_k = \{i, j\}$ of $\mathcal{G}$. $\square$

Example 4.2. Let $n = 2d$ be even and consider the graph $\mathcal{G}$ with edges $\{1, d + 1\}$, $\{2, d + 2\}, \ldots, \{d, 2d\}$. The set of orientations of $\mathcal{G}$ can be identified with $\{-, +\}^d$, where the ith coordinate of a sign vector is equal to $+$ or $-$ if the edge $\{i, d+i\}$ is directed towards i or towards $d + i$, respectively, in the corresponding orientation. The chain proceeds, at each stage, from the current sign vector by picking a coordinate i and switching it to $+$ (respectively, $-$) with probability w_i (respectively, w_{i+d}). Clearly, this chain concides with the nearest neighbor random walk of Example 2.2 on the vertex set of the d-dimensional cube, where w_i and w_{i+d} have the roles played by w_i^+ and w_i^-, respectively, in that example.
                                                                                                $\square$

We postpone the description of the stationary distribution for a general graph $\mathcal{G}$ until Section 5 (see Proposition 5.4), where more examples also appear.

5. Applications to permutation walks

Throughout this section, $\mathcal{A}$ stands for the braid arrangement in $\mathbb{R}^n$. A subarrangement of $\mathcal{A}$ is specified by a simple graph $\mathcal{G}$ on the node set $[n]$. It is first shown that every hyperplane walk on $\mathcal{A}$ induces a walk on the set of acyclic orientations of $\mathcal{G}$ (Proposition 5.1). Specializing to the Tsetlin library walk in Part A gives again the walk on acyclic orientations of Example 1.2. We give a detailed discussion, determining the eigenvalues, stationary distribution and rates of convergence. A birth and extinction example shows that the coupling bound (3.8) can be much better than the eigenvalue bound (3.9). Part B shows how various aspects of a permutation behave after successive riffle shuffles. This yields a probabilistic interpretation for the coefficients of the chromatic polynomial of a graph. Descents of permutations of various types are also treated.

Acyclic Orientations. Let $\mathcal{G}$ be a simple (undirected) graph on the node set $[n]$. Given an orientation $\mathfrak{o}$ of $\mathcal{G}$, a *directed cycle* in $\mathfrak{o}$ is a sequence of nodes $(i_0, i_1, \ldots, i_k)$ such that $i_0 \to i_1 \to \cdots \to i_k \to i_0$ in $\mathfrak{o}$. An orientation of $\mathcal{G}$ is *acyclic* if it contains no directed cycles. An acyclic orientation $\mathfrak{o}$ of $\mathcal{G}$ specifies a partial order $\preceq_{\mathfrak{o}}$ on the set $[n]$ by letting $a \preceq_{\mathfrak{o}} b$ if there exists a directed walk in $\mathfrak{o}$ with initial node a and final node b. Conversely, every partial order $\preceq$ on $[n]$ comes from a graph on the node set $[n]$ in this way (for instance, from the Hasse diagram of $\preceq$).

The acyclic orientations of $\mathcal{G}$ can be modeled by the chambers of a hyperplane arrangement as follows. The *graphical arrangement* corresponding to $\mathcal{G}$ is the subarrangement $\mathcal{B}$ of the braid arrangement $\mathcal{A}$ in $\mathbb{R}^n$, consisting of all hyperplanes of the form $x_i - x_j = 0$ for which $\{i, j\}$ is an edge of $\mathcal{G}$. A chamber $D \in \mathcal{C}_{\mathcal{B}}$ defines an acyclic orientation of $\mathcal{G}$ by assigning the direction $j \to i$ to the edge $\{i, j\}$ of $\mathcal{G}$ if $x_i > x_j$ holds for every point $(x_1, x_2, \ldots, x_n) \in D$. The resulting map is a bijection from the set of chambers $\mathcal{C}_{\mathcal{B}}$ to the

set of acyclic orientations of $\mathcal{G}$, henceforth denoted by $\mathrm{AO}(\mathcal{G})$; see [41, Section 2.3] for a proof and further information. Using this bijection, we may identify chambers of $\mathcal{B}$ with the corresponding acyclic orientations of $\mathcal{G}$.

It follows from the previous discussion that every hyperplane walk on $\mathcal{A}$ induces a Markov chain on the set $\mathrm{AO}(\mathcal{G})$, as described in Section 3. We record this conclusion in the following proposition. We recall that a permutation $\tau \in \mathfrak{S}_n$ is said to be a *linear extention* of a partial order $\preceq$ on $[n]$, if for all $a, b \in [n]$ with $a \prec b$ we have $\tau^{-1}(a) < \tau^{-1}(b)$, meaning that a appears before b in the linear ordering $(\tau(1), \tau(2), \ldots, \tau(n))$ associated to τ.

Proposition 5.1. *Every hyperplane walk on the braid arrangement in $\mathbb{R}^n$ induces a Markov chain on the set $\mathrm{AO}(\mathcal{G})$ of acyclic orientations of $\mathcal{G}$. If the original walk has a unique stationary distribution π, then the stationary distribution $\bar{\pi}$ of the induced chain is given by*

$$(5.1) \qquad \bar{\pi}(\mathfrak{o}) \;=\; \sum_{\tau \in \mathcal{E}(\mathfrak{o})} \pi(\tau)$$

for $\mathfrak{o} \in \mathrm{AO}(\mathcal{G})$, where $\mathcal{E}(\mathfrak{o})$ is the set of linear extensions of the partial order on $[n]$ defined by $\mathfrak{o}$. In particular, if π is the uniform distribution on $\mathfrak{S}_n$, then

$$(5.2) \qquad \bar{\pi}(\mathfrak{o}) \;=\; \frac{\#\mathcal{E}(\mathfrak{o})}{n!}$$

for every $\mathfrak{o} \in \mathrm{AO}(\mathcal{G})$.

Proof. Let $\mathcal{A}$ denote the braid arrangement in $\mathbb{R}^n$ and $\mathcal{B}$ denote the graphical arrangement corresponding to $\mathcal{G}$, as before. As already mentioned, the first statement follows from the previous discussion and Corollary 3.2. The second statement follows from (3.7) and the observation that for chambers $C \in \mathcal{C}_\mathcal{A}$ and $D \in \mathcal{C}_\mathcal{B}$ corresponding to the permutation $\tau \in \mathfrak{S}_n$ and the acyclic orientation $\mathfrak{o} \in \mathrm{AO}(\mathcal{G})$, respectively, we have $C \subseteq D$ if and only if $\tau \in \mathcal{E}(\mathfrak{o})$. $\square$

Remark 5.2. Acyclic orientations are of importance in various areas of applied mathematics, such as computer science, automata theory and statistics. In statistical applications they appear as part of the machinery of "Bayes nets" and "casual models", where they are used to model casual implication in complex data sets; some useful references are [22, 28, 24]. Searching for an appropriate model is often done by a random walk on acyclic orientations. We hope that our analysis will contribute to the understanding of these algorithms. For an introduction to the literature relating acyclic orientations to factoring noncommutative polynomials, see [34]. $\square$

In the remainder of this section we investigate further the Markov chain of Proposition 5.1 in the special cases of Examples 2.3 and 2.4. Note that the case of Example 2.3 is also treated by the Markov chain (4.2).

A. Tsetlin Library. Let $w_1, w_2, \ldots, w_n$ be nonnegative real numbers summing to 1 and let w be the probability measure on $\mathcal{F}_\mathcal{A}$ of Example 2.3. Thus the associated hyperplane walk on $\mathcal{A}$ is the Markov chain on $\mathfrak{S}_n$ which selects, at each stage, the entry i in the one

line notation of the current permutation with probability w_i and moves it in front. To describe the induced chain of Proposition 5.1 on the set $\mathrm{AO}(\mathcal{G})$, we observe the following: if $C \in \mathcal{C}_\mathcal{A}$ is the chamber which corresponds to a given permutation $\tau \in \mathfrak{S}_n$ and $D \in \mathcal{C}_\mathcal{B}$ is the unique chamber of $\mathcal{B}$ which contains C, then the acyclic orientation of $\mathcal{G}$ corresponding to D is the one which orients an edge $\{a, b\}$ of $\mathcal{G}$ as $b \to a$ if and only if $\tau^{-1}(a) < \tau^{-1}(b)$. It follows that the induced chain on $\mathrm{AO}(\mathcal{G})$ proceeds from a given acyclic orientation by selecting the node i of $\mathcal{G}$ with probability w_i and reorienting all edges of $\mathcal{G}$ incident to this node towards itself, to reach a new acyclic orientation of $\mathcal{G}$, leaving the orientations of all other edges of $\mathcal{G}$ unchanged. Equivalently, if K^* is the transition matrix of the induced chain on $\mathrm{AO}(\mathcal{G})$, then $K^*(\mathfrak{o}, \mathfrak{o}')$ is given by the right-hand side of (4.2) for $\mathfrak{o}, \mathfrak{o}' \in \mathrm{AO}(\mathcal{G})$.

Given a subset T of the node set of $\mathcal{G}$, we denote by $\mathcal{G} \smallsetminus T$ the graph obtained from $\mathcal{G}$ by removing all nodes in T and all incident to them edges (in other words, $\mathcal{G} \smallsetminus T$ is the induced subgraph of $\mathcal{G}$ on the node set $[n] \smallsetminus T$). Thus T is dominating in $\mathcal{G}$ if and only if the graph $\mathcal{G} \smallsetminus T$ has no edges. The following statements can be added to the conclusions of Proposition 5.1.

Proposition 5.3. *Let $\mathcal{G}$ be a simple graph on the node set $[n]$ and let K^* be the transition matrix of the Markov chain on $\mathrm{AO}(\mathcal{G})$ which is induced from the Tsetlin library with weights $w_1, w_2, \ldots, w_n$.*

(i) *The matrix K^* is diagonalizable with characteristic polynomial given by*

(5.3)
$$\det(xI - K^*) = \prod_{S \subseteq [n]} (x - \lambda_S)^{m_S},$$

where

(5.4)
$$\lambda_S = \sum_{i \in S} w_i$$

and

(5.5)
$$m_S = \sum_{S \subseteq T \subseteq [n]} (-1)^{|T \smallsetminus S|} \, \#\mathrm{AO}(\mathcal{G} \smallsetminus T)$$

for $S \subseteq [n]$, where the number of acyclic orientations of the graph with empty node set is equal to one, by convention.

(ii) *K^* has a unique stationary distribution $\bar{\pi}$ if and only if there is no edge $\{i, j\}$ of $\mathcal{G}$ such that $w_i = w_j = 0$. Moroever, we have*

$$\bar{\pi}(\mathfrak{o}) = \sum_{\tau \in \mathcal{E}(\mathfrak{o})} \frac{w_{\tau(1)} w_{\tau(2)} \cdots w_{\tau(n)}}{(1 - w_{\tau(1)})(1 - w_{\tau(1)} - w_{\tau(2)}) \cdots (1 - w_{\tau(1)} - \cdots - w_{\tau(n-1)})}$$

for $\mathfrak{o} \in \mathrm{AO}(\mathcal{G})$, if $w_1, w_2, \ldots, w_n$ are all positive. In particular, (5.2) holds for every $\mathfrak{o} \in \mathrm{AO}(\mathcal{G})$ if $w_1 = \cdots = w_n = 1/n$.

(iii) *Assuming that $\bar{\pi}$ exists, the conclusions of Corollary 4.1 (iii) hold if π is replaced there by $\bar{\pi}$ and $K_{\mathfrak{o}}^l$ is replaced by the distribution $(K_{\mathfrak{o}}^*)^l$ of the induced chain started from the orientation $\mathfrak{o} \in \mathrm{AO}(\mathcal{G})$ after l steps.*

Proof. We first recall from [41, Section 2.3] the following description of the intersection poset $\mathcal{L}_\mathcal{B}$ of the graphical arrangement $\mathcal{B} \subseteq \mathcal{A}$ corresponding to $\mathcal{G}$. A set partition π of $[n]$ is said to be $\mathcal{G}$-*connected* if the induced subgraph of $\mathcal{G}$ on each block of π is connected. We denote by $\mathcal{L}_\mathcal{G}$ the set of $\mathcal{G}$-connected partitions of $\mathcal{G}$, ordered by refinement. This poset, known as the "bond lattice", or "lattice of contractions" of $\mathcal{G}$, is isomorphic to $\mathcal{L}_\mathcal{B}$, where the isomorphism is induced by the one between the lattice of partitions of $[n]$ and $\mathcal{L}_\mathcal{A}$, discussed in Section 2.2.

Given $W \in \mathcal{L}_\mathcal{B}$, we write λ_σ for the eigenvalue λ_W of K^* which appears in Theorem 3.3, where $\sigma \in \mathcal{L}_\mathcal{G}$ is the $\mathcal{G}$-connected partition corresponding to W. The definition of the measure w on $\mathcal{F}_\mathcal{A}$ of Example 2.3 and the definition of λ_W in (2.2) imply that

$$\tag{5.6} \lambda_\sigma \;=\; \sum_{\{i\} \in \sigma} w_i,$$

where the sum runs over all singleton blocks $\{i\}$ of $\sigma \in \mathcal{L}_\mathcal{G}$. Theorem 3.3 (i) gives

$$\det(xI - K^*) \;=\; \prod_{\sigma \in \mathcal{L}_\mathcal{G}} (x - \lambda_\sigma)^{m_\sigma^*},$$

where $m_\sigma^* = |\mu_\mathcal{G}(\hat{0}, \sigma)|$ and $\mu_\mathcal{G}$ is the Möbius function of $\mathcal{L}_\mathcal{G}$. The previous two equations imply that (5.3) holds if we define

$$m_S \;=\; \sum_{\sigma \in \mathcal{L}_\mathcal{G}:\ \mathrm{sing}(\sigma) = S} |\mu_\mathcal{G}(\hat{0}, \sigma)|,$$

where $\mathrm{sing}(\sigma) = \{i \in [n] : \{i\} \in \sigma\}$ denotes the set of singleton blocks of σ. To complete the proof of part (i), it remains to prove (5.5). By inclusion-exclusion we can write

$$\tag{5.7} m_S \;=\; \sum_{S \subseteq T \subseteq [n]} (-1)^{|T \smallsetminus S|}\, n_T,$$

where

$$n_T \;=\; \sum_{\sigma \in \mathcal{L}_\mathcal{G}:\ T \subseteq \mathrm{sing}(\sigma)} |\mu_\mathcal{G}(\hat{0}, \sigma)|.$$

Clearly, writing $\sigma = \{B_1, B_2, \ldots, B_k\}$, the closed interval $[\hat{0}, \sigma]$ in $\mathcal{L}_\mathcal{G}$ is isomorphic to the direct product of the lattices $\mathcal{L}_{\mathcal{G}_i}$ for $1 \leq i \leq k$, where $\mathcal{G}_i$ is the induced subgraph of $\mathcal{G}$ on the node set B_i. It follows easily from this observation and the multiplicativity of the Möbius function [38, Proposition 3.8.2] that

$$\tag{5.8} n_T \;=\; \sum_{\sigma \in \mathcal{L}_{\mathcal{G} \smallsetminus T}} |\mu_{\mathcal{G} \smallsetminus T}(\hat{0}, \sigma)|.$$

By Zaslavsky's formula [41, Theorem 2.5] [47], the right-hand side of (5.8) is equal to the number of chambers of the graphical arrangement corresponding to $\mathcal{G} \smallsetminus T$ and hence to the number of acyclic orientations of $\mathcal{G} \smallsetminus T$. Thus (5.5) follows from (5.7) and (5.8).

Part (ii) follows from Theorem 3.3 (iii), equation (2.9) and Proposition 5.1.

Part (iii) follows from Theorem 3.3 (iv), since a product of faces of $\mathcal{A}$ corresponding to ordered partitions of the form $(i, [n] \smallsetminus \{i\})$ is not contained in any of the hyperplanes of $\mathcal{B}$ if

and only if the corresponding set of nodes v_i is dominating in $\mathcal{G}$ and since $\lambda_H = 1 - w_i - w_j$ holds for the hyperplane H of $\mathcal{B}$ corresponding to the edge $\{i, j\}$ of $\mathcal{G}$. $\qquad\square$

Part (ii) of Proposition 5.3 and the following statement determine the stationary distribution of the Markov chain on the set $\mathrm{O}(\mathcal{G})$ of all orientations of $\mathcal{G}$, discussed in part C of Section 4. We note that if $\mathcal{G}$ is a forest, then every orientation of $\mathcal{G}$ is acyclic and hence the two Markov chains on $\mathrm{O}(\mathcal{G})$ and $\mathrm{AO}(\mathcal{G})$ coincide.

Proposition 5.4. *Consider the chain on the set* $\mathrm{O}(\mathcal{G})$ *of all orientations and the chain on the set* $\mathrm{AO}(\mathcal{G})$ *of acyclic orientations of* $\mathcal{G}$, *with weights* $w_1, w_2, \ldots, w_n$. *Assuming there is no edge* $\{i, j\}$ *of* $\mathcal{G}$ *such that* $w_i = w_j = 0$, *their respective stationary distributions* Π *and* $\bar{\pi}$ *are related by*

$$(5.9) \qquad \Pi(\mathfrak{o}) \;=\; \begin{cases} \bar{\pi}(\mathfrak{o}), & \text{if } \mathfrak{o} \text{ is acyclic} \\ 0, & \text{otherwise} \end{cases}$$

for $\mathfrak{o} \in \mathrm{O}(\mathcal{G})$.

Proof. We denote by K the transition matrix of the chain on $\mathrm{O}(\mathcal{G})$ and recall that $\Pi(\mathfrak{o}) = \lim_{l \to \infty} K^l(\mathfrak{o}_*, \mathfrak{o})$ for $\mathfrak{o} \in \mathrm{O}(\mathcal{G})$, where the limit is independent of the starting orientation $\mathfrak{o}_* \in \mathrm{O}(\mathcal{G})$. Choosing $\mathfrak{o}_* \in \mathrm{AO}(\mathcal{G})$, all orientations in the chain stay in $\mathrm{AO}(\mathcal{G})$ and the limit becomes equal to the right-hand side of (5.9). $\qquad\square$

Example 5.5. Fix an integer $1 \le k \le n$ and let $\mathcal{G}$ be the graph with edges $\{i, j\}$ for $1 \le i < j \le k$. The set $\mathrm{AO}(\mathcal{G})$ can be identified with the subgroup $\mathfrak{S}_k$ of permutations in $\mathfrak{S}_n$ which fix the set $\{k+1, \ldots, n\}$ pointwise and the induced chain is the process which records the relative ordering of $\{1, 2, \ldots, k\}$, when $\tau \in \mathfrak{S}_n$ evolves as in the Markov chain of Example 2.3. The eigenvalues and stationary distribution $\bar{\pi}$ of the transition matrix K^* can be easily deduced from those of the transition matrix K of the parent chain, since in this case K^* differs by a multiple of the identity matrix from the restriction of K on $\mathfrak{S}_k$. For instance, K^* has eigenvalues (2.8), one for each $\tau \in \mathfrak{S}_k$. By Proposition 5.3 (iii), the bound from (4.5) applies and gives

$$\| (K_\tau^*)^l - \bar{\pi} \|_{\mathrm{TV}} \;\le\; \sum_{1 \le i < j \le k} (1 - w_i - w_j)^l.$$

In particular, if $w_1 = \cdots = w_k = 1/k$, so that $\bar{\pi}$ is uniform, then

$$(5.10) \qquad \| (K_\tau^*)^l - \bar{\pi} \|_{\mathrm{TV}} \;\le\; \binom{k}{2} \left(1 - \frac{2}{n} \right)^l,$$

which is an improvement over (2.11). Thus the left-hand side of (5.10) is bounded above by $e^{-2c}/2$ if $l \ge n(\log k + c)$.

The bound (5.10) is quite sharp across the whole range of k. For instance, if $k = 2$ it shows that l must grow as cn, with c approaching infinity. This is correct since if c stays bounded, then there is a nonzero chance that neither 1 nor 2 has been moved, and thus that they have stayed in their original relative order. At the other extreme, we have

already commented in our discussion of (2.11) that the bound is sharp if $k = n$. Similar remarks hold for other values of k. □

The following example gives a concrete case in which the bound of (3.8) is better than that of (3.9).

Example 5.6 (Birth and Extinction). Let $\mathcal{G}$ be the graph with edges $\{i, n\}$ for $1 \leq i \leq m$, where $m \leq n - 1$ is a positive integer, and choose weights $w_1 = \cdots = w_n = 1/n$. Every orientation of $\mathcal{G}$ is acyclic and hence the set $\mathrm{AO}(\mathcal{G})$ can be identified with $\{-, +\}^m$, as described in part C of Section 4. The chain proceeds, at each stage, from a sign vector $x \in \{-, +\}^m$ by picking a coordinate i uniformly at random and switching this coordinate to $+$, if $1 \leq i \leq m$, leaving x unchanged, if $m + 1 \leq i < n$, and switching all coordinates of x to $-$, if $i = n$, to reach a new sign vector. Such processes are studied in mathematical genetics with many variations.

From our current point of view, we may think of this chain as the process which records the subset of $[m]$ consisting of those integers which precede n in the current permutation τ, when $\tau \in \mathfrak{S}_n$ evolves as in the Markov chain of Example 2.3 with uniform weights (random to top model). This is because a number $i \in [m]$ precedes n in some (equivalently, every) linear extension of the orientation $\mathfrak{o}$ if and only if the edge $\{i, n\}$ is directed as $n \to i$ in $\mathfrak{o}$.

Proposition 5.7. *Consider the Markov chain on* $\mathrm{AO}(\mathcal{G})$, *which is induced from the Tsetlin library with uniform weights, as a chain on the set* $\{-, +\}^m$ *and let* K^* *be its transition matrix.*

(i) *The matrix* K^* *is diagonalizable with eigenvalues*

$$\begin{cases} 1, & \text{with multiplicity one} \\ \frac{n-j-1}{n}, & \text{with multiplicity } \binom{m}{j}, \text{ for } 1 \leq j \leq m. \end{cases}$$

(ii) *The stationary distribution of* K^* *is given by*

$$(5.11) \qquad \bar{\pi}(x) \;=\; \frac{1}{(m+1)\binom{m}{k}}$$

for $x \in \{-, +\}^m$, *where* k *is the number of coordinates of* x *equal to* $+$.

(iii) *Assume* $m = n - 1$ *and let* $(K_x^*)^l$ *be the distribution of the chain started from* x *after* l *steps. We have*

$$(5.12) \qquad \|(K_x^*)^l - \bar{\pi}\|_{\mathrm{TV}} \;\leq\; \left(1 - \frac{1}{n}\right)^l \;\leq\; e^{-c}$$

for $l \geq cn$ *and* $c > 0$. *Moreover this bound is sharp, in the sense that there exists* $0 < \theta < 1$ *such that* $\|(K_x^*)^n - \bar{\pi}\|_{\mathrm{TV}} \geq \theta$ *for all large* n.

Proof. Part (i) follows from Corollary 4.1 (i). Alternatively, it follows from the proof of Proposition 5.3 (i) and, in particular, equation (5.6), since all values of the Möbius function $\mu_{\mathcal{G}}$ in this case have absolute value 1. Part (ii) follows from Proposition 5.3 (ii) and equation (5.2), which applies in our situation, since the number of linear extensions

of the poset on $[n]$ defined by any orientation of $\mathcal{G}$ with k edges pointing away from n is equal to $\binom{n}{m+1} k!(m-k)!(n-m-1)!$.

Assuming that $m = n-1$, (5.12) follows from the bound given in (4.4), since a dominating set in $\mathcal{G}$ is formed as soon as node n is picked and the chance that this has not happen in the first l steps is equal to $(1 - 1/n)^l$. Finally, suppose that the starting sign vector x has all its coordinates equal to $-$ and let A be the set of all $y \in \{-, +\}^{n-1}$ having at least $(n-1)/2$ coordinates equal to $+$. An elementary calculation shows that after n steps in the chain, the expected number of $+$ coordinates is equal to

$$\frac{n-1}{2}\left(1 - \left(1 - \frac{2}{n}\right)^n\right) \sim \frac{n-1}{2}\left(1 - \frac{1}{e^2}\right).$$

It follows that $(K_x^*)^n(A) \to 0$ as $n \to \infty$, while clearly $\bar{\pi}(A) \geq 1/2$. Since the total variation distance $\|(K_x^*)^n - \bar{\pi}\|_{\mathrm{TV}}$ is bounded below by $|(K_x^*)^n(A) - \bar{\pi}(A)|$, we conclude that given any $0 < \theta < 1/2$ we have $\|(K_x^*)^n - \bar{\pi}\|_{\mathrm{TV}} \geq \theta$ for n large enough. This completes the proof of part (iii). A similar argument works for all $2 \leq m \leq n-1$. $\qquad\square$

Example 5.8 (Descent Set). Let $\mathcal{G}$ be the path with edges $\{i, i+1\}$ for $1 \leq i \leq n-1$ and choose weights $w_1 = \cdots = w_n = 1/n$. Once again, every orientation of $\mathcal{G}$ is acyclic and hence the set $\mathrm{AO}(\mathcal{G})$ can be identified with the set of sign vectors $\{-, +\}^{n-1}$. We leave it to the reader to give a description of the evolution of this chain on the set $\{-, +\}^{n-1}$ similar to that of Example 5.6.

We find it more convenient to identify $\mathrm{AO}(\mathcal{G})$ with the set of subsets of $[n-1]$, where an orientation $\mathfrak{o}$ of $\mathcal{G}$ is identified with the set of indices $i \in [n-1]$ for which the edge $\{i, i+1\}$ is directed as $i \to i+1$ in $\mathfrak{o}$. We denote by $\mathcal{B}$ the graphical arrangement associated to $\mathcal{G}$, as usual, and recall that there is a directed edge $i \to i+1$ in $\mathfrak{o}$ if and only if $x_i < x_{i+1}$ holds in the chamber D of $\mathcal{B}$ corresponding to $\mathfrak{o}$. In turn, this happens if and only if $i+1$ precedes i in any of the permutations τ which correspond to chambers of $\mathcal{A}$ contained in D or, equivalently, if and only if i belongs to the descent set

$$\mathrm{Des}(\tau^{-1}) = \{i \in [n-1] : \tau^{-1}(i) > \tau^{-1}(i+1)\}$$

of the inverse permutation τ^{-1}. Therefore, our chain on the set of subsets of $[n-1]$ is the process which records the descent set $\mathrm{Des}(\tau^{-1})$, when $\tau \in \mathfrak{S}_n$ evolves as in the Markov chain of Example 2.3 with uniform weights. We recall that a *composition* of n is an ordered sequence of positive integers (called parts) which sum to n.

Proposition 5.9. *Consider the Markov chain on* $\mathrm{AO}(\mathcal{G})$, *which is induced from the Tsetlin library with uniform weights, as a chain on the set of subsets of* $[n-1]$ *and let* K^* *be its transition matrix.*

(i) *The matrix* K^* *is diagonalizable with eigenvalues* j/n *for* $j \in \{0, 1, \ldots, n-2\} \cup \{n\}$, *where the multiplicity of* j/n *is equal to the number of compositions of* n *having exactly* j *parts equal to 1.*

(ii) *The stationary distribution of* K^* *is given by*

$$(5.13) \qquad\qquad \bar{\pi}(S) = \frac{1}{n!} \#\{\tau \in \mathfrak{S}_n : \mathrm{Des}(\tau) = S\}$$

for every $S \subseteq [n-1]$.
(iii) *We have*

$$\|(K_S^*)^l - \bar{\pi}\|_{\text{TV}} \ \leq \ (n-1)\left(1 - \frac{2}{n}\right)^l,$$

where $(K_S^*)^l$ is the distribution of the chain started from S after l steps.

Proof. Part (i) follows once again from Corollary 4.1 (i), or from the proof of Proposition 5.3 (i). For part (ii) it suffices to note that given $S \subseteq [n-1]$ with corresponding orientation $\mathfrak{o} \in \text{AO}(\mathcal{G})$, the set of linear extensions of the partial order on $[n]$ defined by $\mathfrak{o}$ is in bijection with the set of elements of $\mathfrak{S}_n$ with descent set equal to S, as already discussed before the statement of the proposition. Then (5.13) follows from (5.2) and Proposition 5.3 (ii). Part (iii) is a consequence of (4.6). $\qquad\square$

Part (iii) of this proposition implies that for $c > 0$, the distance $\|(K_S^*)^l - \bar{\pi}\|_{\text{TV}}$ is bounded above by e^{-c} if $l \geq \frac{n}{2}(\log n + c)$. This can be shown to be sharp, in the sense of Proposition 5.7 (iii), by an argument similar to the one in the proof of this proposition. $\qquad\square$

Example 5.10 (Cyclic Descent Set). For notational convenience, in this example we replace the node set $[n]$ by the abelian group $\mathbb{Z}_n$ of integers modulo n. We let $\mathcal{G}$ be the cycle with edges $\{i, i+1\}$ for $1 \leq i \leq n$ and choose weights $w_1 = \cdots = w_n = 1/n$. Since there are exactly two orientations of $\mathcal{G}$ which have a directed cycle, the number of acyclic orientations of $\mathcal{G}$ is equal to $2^n - 2$. We may identify $\text{AO}(\mathcal{G})$ with the set of proper subsets of $[n]$, where an acyclic orientation $\mathfrak{o}$ of $\mathcal{G}$ corresponds to the set of indices $i \in [n]$ for which the edge $\{i, i+1\}$ is directed as $i \to i+1$ in $\mathfrak{o}$. Arguing as in Example 5.8, we see that this chain on the set of proper subsets of $[n]$ is the process which records the cyclic descent set

$$\text{cDes}(\tau^{-1}) \ = \ \{i \in [n] : \tau^{-1}(i) > \tau^{-1}(i+1)\}$$

when $\tau \in \mathfrak{S}_n$ evolves as in the Markov chain of Example 2.3 with uniform weights. Cyclic descents of permutations were introduced by Cellini [18] and further studied by Fulman; see [23] and references therein.

The bond lattice $\mathcal{L}_\mathcal{G}$ is isomorphic to the set of subsets of $[n]$, other than those of cardinality $n-1$, partially ordered by inclusion. Since this lattice has a well known Möbius function, one can deduce easily from (5.6) the following description of the eigenvalues of the transition matrix K^* of this chain. They are the numbers j/n for $j \in \{0, 1, \ldots, n\}$ and for $j \geq 1$, the multiplicity of j/n is equal to the number of set partitions of $\mathbb{Z}_n$ into blocks of the form $\{a, a+1, \ldots, b\}$ having exactly j singleton blocks. The multiplicity of zero is two less than the number of such partitions of $\mathbb{Z}_n$ having no singleton block. Arguing as in Example 5.8, we find that the stationary distribution of K^* is given by

$$(5.14) \qquad\qquad \bar{\pi}(S) \ = \ \frac{1}{n!} \, \#\{\tau \in \mathfrak{S}_n : \text{cDes}(\tau) = S\}$$

for proper subsets $S \subseteq [n]$ and that

$$(5.15) \qquad\qquad \|(K_S^*)^l - \bar{\pi}\|_{\text{TV}} \ \leq \ n\left(1 - \frac{2}{n}\right)^l,$$

where $(K_S^*)^l$ is the distribution of the chain started from S after l steps. As in Example 5.8, it follows that $\|(K_S^*)^l - \bar{\pi}\|_{\mathrm{TV}}$ is bounded above by e^{-c} if $l \geq \frac{n}{2}(\log n + c)$. $\square$

B. Inverse a-shuffling. Let $a \geq 2$ be an integer and let w be the probability measure on $\mathcal{F}_{\mathcal{A}}$ of Example 2.4, so that the hyperplane walk associated to w is the Markov chain of inverse a-shuffles on $\mathfrak{S}_n$. Using similar reasoning to the one in the case of the Tsetlin library, one can describe the induced chain of Proposition 5.1 on the set $\mathrm{AO}(\mathcal{G})$ as follows: The chain proceeds from a given acyclic orientation of $\mathcal{G}$ by selecting uniformly at random a weak ordered partition $B = (B_1, B_2, \ldots, B_a)$ of $[n]$ with a blocks. Then the orientation of any edge of $\mathcal{G}$ whose endpoints belong to the same block of B is left unchanged and any other edge $\{u, v\}$ of $\mathcal{G}$ is reoriented as $u \to v$, if $i < j$ holds for the unique indices i and j with $v \in B_i$ and $u \in B_j$, to reach a new acyclic orientation of $\mathcal{G}$. We will refer to the induced chain as the chain of inverse a-shuffles on $\mathrm{AO}(\mathcal{G})$. Its transition matrix K^* satisfies

$$K^*(\mathfrak{o}, \mathfrak{o}') \;=\; \frac{\nu^*(\mathfrak{o}, \mathfrak{o}')}{a^n}$$

for $\mathfrak{o}, \mathfrak{o}' \in \mathrm{AO}(\mathcal{G})$, where $\nu^*(\mathfrak{o}, \mathfrak{o}')$ is the number of weak ordered partitions of $[n]$ with a blocks, the action of which on $\mathfrak{o}$, just described, results in $\mathfrak{o}'$.

We denote by $\chi_{\mathcal{G}}$ the chromatic polynomial [41, Section 2.3] of $\mathcal{G}$. Thus for every positive integer q, $\chi_{\mathcal{G}}(q)$ is equal to the number of colorings $\kappa : [n] \to [q]$ of the nodes of $\mathcal{G}$ with q colors satisfying $\kappa(u) \neq \kappa(v)$ for every edge $\{u, v\}$ of $\mathcal{G}$. Part (i) of the following corollary provides an interpretation to the coefficients of $\chi_{\mathcal{G}}$ which strengthens a theorem of Stanley [37] [41, Corollary 2.3], stating that the sum of the unsigned coefficients of $\chi_{\mathcal{G}}$ is equal to the number of acyclic orientations of $\mathcal{G}$. There are other interpretations to these coefficients; see, for instance, [41, Theorem 4.12] and [25, 44].

Proposition 5.11. *Let $\mathcal{G}$ be a simple graph on the node set $[n]$ and let K^* be the transition matrix of the Markov chain of inverse a-shuffles on $\mathrm{AO}(\mathcal{G})$.*

(i) *The matrix K^* is diagonalizable with characteristic polynomial given by*

$$\det(xI - K^*) \;=\; \prod_{i=0}^{n-1} (x - \frac{1}{a^i})^{p_i},$$

where

$$\chi_{\mathcal{G}}(q) \;=\; \sum_{i=0}^{n-1} (-1)^i \, p_i q^{n-i}$$

is the chromatic polynomial of $\mathcal{G}$.

(ii) *The stationary distribution $\bar{\pi}$ of K^* is given by (5.2).*

(iii) *We have*

(5.16) $$\|(K_{\mathfrak{o}}^*)^l - \bar{\pi}\|_{\mathrm{TV}} \;\leq\; m \left(\frac{1}{a}\right)^l,$$

where m is the number of edges of $\mathcal{G}$ and $(K_{\mathfrak{o}}^*)^l$ is the distribution of the chain started from the acyclic orientation $\mathfrak{o}$, after l steps.

Proof. Let $\mathcal{B}$ denote the graphical arrangement corresponding to $\mathcal{G}$, as usual, and $\mu_\mathcal{B}$ denote the Möbius function of the intersection poset $\mathcal{L}_\mathcal{B}$. It follows from Theorem 3.3 (i) and (2.12) that the distinct eigenvalues of K^* are $1, 1/a, 1/a^2, \ldots, 1/a^{n-1}$ and that the multiplicity p_i of the eigenvalue $1/a^i$ satisfies

$$p_i \;=\; \sum_{W \in \mathcal{L}_\mathcal{B}:\, \mathrm{codim}(W,V)=i} (-1)^i \, \mu_\mathcal{B}(V,W).$$

Equivalently, $(-1)^i p_i$ is equal to the coefficient of q^{n-i} in the characteristic polynomial [31, Section 2.3] [41, Section 1.3] of $\mathcal{B}$, which is known to equal $\chi_\mathcal{G}(q)$ [41, Theorem 2.7]. This proves part (i). Part (ii) follows from Proposition 5.1, since the chain of inverse a-shuffles converges to the uniform distribution on $\mathfrak{S}_n$. Part (iii) follows from (3.9), since $\lambda_H = 1/a$ for every $H \in \mathcal{B}$. $\qquad\square$

For the applications discussed in the following example, we think of $\mathfrak{S}_n$ as the set of linear orderings of a deck of n cards, labeled by the elements of $[n]$. Since inverse a-shuffling, followed by passing to the inverse of the current permutation, gives the same distribution as ordinary a-shuffling, there is a straightforward translation of our results into the language of a-shuffles.

Example 5.12. (i) Suppose that $\mathcal{G}$ is the star of Example 5.6, say with $n-1$ edges. The chain of inverse a-shuffles on $\mathrm{AO}(\mathcal{G})$ is the process which records the set $S \subseteq [n-1]$ of the labels of cards which precede card n in the current linear ordering, in the chain of inverse a-shuffles on $\mathfrak{S}_n$ (this is stronger than just recording the current position of card n; see [3] for a summary of results on that Markov chain).

Since $\mathcal{G}$ is a tree, its chromatic polynomial is given by $\chi_\mathcal{G}(q) = q(q-1)^{n-1}$ and hence, by Proposition 5.11 (i), the matrix K^* has eigenvalues $1/a^i$ with multiplicity $\binom{n-1}{i}$, for $0 \leq i \leq n-1$. The stationary distribution $\bar{\pi}(S)$ is given by the right-hand side of (5.11), where $m = n - 1$ and k is the number of elements of S. Moreover, (5.16) gives

$$(5.17) \qquad\qquad \|(K^*)^l - \bar{\pi}\|_{\mathrm{TV}} \;\leq\; (n-1)\left(\frac{1}{a}\right)^l$$

and hence the total variation distance on the left is bounded above by a^{-c} if $l \geq \log_a n + c$, for $c > 0$. This shows a speedup over the $(3/2)\log_a n + c$, required for the parent chain of inverse a-shuffles on $\mathfrak{S}_n$ to reach stationarity, and is essentially sharp by the results of [3, Section 2].

(ii) Suppose that $\mathcal{G}$ is the path of Example 5.8. The induced chain records the descent set $\mathrm{Des}(\tau^{-1})$ of the inverse of the current permutation τ in the chain of inverse a-shuffles on $\mathfrak{S}_n$. The stationary distribution is given by (5.13). Since the path $\mathcal{G}$ is also a tree, the description of the eigenvalues for the star example and (5.17) continue to hold. The result on the rate of convergence in this case was obtained earlier in [17, Section 3], where it is also shown that $(1/2)\log_a n + c$ steps are necessary and sufficient for $\mathrm{Des}(\tau^{-1})$ to reach stationarity if n is large.

(iii) Suppose that $\mathcal{G}$ is the cycle of Example 5.10. The induced chain now records the cyclic descent set $\mathrm{cDes}(\tau^{-1})$ of the inverse of the current permutation τ in the chain of

inverse a-shuffles on $\mathfrak{S}_n$. It follows, in particular, that this process is a Markov chain on the set of proper subsets of $[n]$. As in the previous case, we find that the stationary distribution is given by (5.14) and that

$$\|(K^*)^l - \bar{\pi}\|_{\mathrm{TV}} \ \leq \ n \left(\frac{1}{a}\right)^l,$$

so that $\|(K^*)^l - \bar{\pi}\|_{\mathrm{TV}}$ is bounded above by a^{-c} if $l \geq \log_a n + c$, for $c > 0$. We leave further details to the interested reader. $\qquad\square$

Remark 5.13. The difference between ordinary and inverse a-shuffles is easy to appreciate by considering the graph with a single edge $\{1, 2\}$. Then the induced process records the relative order of cards labeled 1 and 2. After fewer than $\log_a n$ ordinary a-shuffles, there is a good chance (close to 1) that they are still in their original order. However, their relative order is close to random after a growing number of inverse a-shuffles. $\qquad\square$

6. Semigroup walks

The theory of hyperplane walks was generalized to random walks on semigroups by Brown [11, 12]. This section shows how our main results can be extended in this direction. Some familiarity with the ideas of [11, 12] will be assumed. The algebraic aspects of Brown's theory of semigroup walks have been further studied in [1, 42, 43], with probabilistic developments in [19]. These references contain examples to which the following theory may be applied.

The *face semigroup* of a hyperplane arrangement $\mathcal{A}$ is defined as the set $\mathcal{F}_\mathcal{A}$ of faces of $\mathcal{A}$, endowed with the product operation discussed in Section 2. The set $\mathcal{C}_\mathcal{A}$ of chambers is a left ideal of $\mathcal{F}_\mathcal{A}$, meaning that it is a nonempty subset which is closed under left multiplication by elements of $\mathcal{F}_\mathcal{A}$ (of course, $\mathcal{C}_\mathcal{A}$ is a right ideal as well). Given a finite semigroup $\mathcal{S}$, a left ideal $\mathcal{C}$ of $\mathcal{S}$ and a probability measure w on $\mathcal{S}$, one can define a Markov chain on the state space $\mathcal{C}$ with transition matrix K given by

$$(6.1) \qquad\qquad K(c, c') \ = \ \sum_{x \in \mathcal{S}:\, xc = c'} w(x)$$

for $c, c' \in \mathcal{C}$. We refer to this chain as the *semigroup walk* on $\mathcal{C}$ associated to w; it coincides with the hyperplane walk on $\mathcal{A}$ associated to w, if $\mathcal{S} = \mathcal{F}_\mathcal{A}$ and $\mathcal{C} = \mathcal{C}_\mathcal{A}$. The semigroup $\mathcal{S}$ is said to be a *band* if $x^2 = x$ for every $x \in \mathcal{S}$. To such a semigroup, one can associate a join semilattice L and a surjective map supp : $\mathcal{S} \to L$, satisfying

$$(6.2) \qquad\qquad \mathrm{supp}(x) \leq_L \mathrm{supp}(y) \ \Leftrightarrow \ y = yxy$$

for $x, y \in \mathcal{S}$; see [12, Section A.2] for further details. The support map has the additional property that

$$(6.3) \qquad\qquad \mathrm{supp}(xy) \ = \ \mathrm{supp}(x) \vee \mathrm{supp}(y)$$

for $x, y \in \mathcal{S}$, where $u \vee v$ denotes the least upper bound (join) of u and v in L. In the special case of a face semigroup $\mathcal{F}_\mathcal{A}$, the support of a face $F \in \mathcal{F}_\mathcal{A}$ is the linear span of

F and the semilattice L is the dual of the intersection poset $\mathcal{L}_\mathcal{A}$, defined in Section 2. A band $\mathcal{S}$ is called *left-regular* if $xyx = xy$ for all $x, y \in \mathcal{S}$.

Assume that $\mathcal{S}$ is a finite band. Then L is a finite join semilattice and hence it has a maximum element, denoted $\hat{1}$. It follows from (6.3) that the set $\mathcal{C}_\mathcal{S} = \{c \in \mathcal{S} : \operatorname{supp}(c) = \hat{1}\}$ is a left ideal of $\mathcal{S}$. The elements of $\mathcal{C}_\mathcal{S}$ are called *chambers*. Given $x \in \mathcal{S}$, the subsemigroup $x\mathcal{S}$ is a finite band whose number of chambers depends only on the support $u = \operatorname{supp}(x)$ of x in L; see [12, Section B.3]. We denote this number by $n(u)$. The following theorem was proved for left-regular bands in [11] and generalized to all bands in [12].

Theorem 6.1. *Let $\mathcal{S}$ be a finite band with corresponding semilattice L and set of chambers $\mathcal{C}_\mathcal{S}$. Let w be a probability measure on $\mathcal{S}$ and let K be the transition matrix of the semigroup walk on $\mathcal{C}_\mathcal{S}$ associated to w.*

   (i) *The characteristic polynomial of K is given by*

$$(6.4) \qquad \det(xI - K) \;=\; \prod_{u \in L} (x - \lambda_u)^{m_u},$$

      where

$$(6.5) \qquad \lambda_u \;=\; \sum_{x \in \mathcal{S}:\ \operatorname{supp}(x) \leq_L u} w(x)$$

      is an eigenvalue,

$$(6.6) \qquad m_u \;=\; \sum_{u \leq_L v} \mu_L(u, v)\, n(v),$$

      μ_L is the Möbius function of L and $n(v)$ is the number of chambers of $x\mathcal{S}$ for any $x \in \mathcal{S}$ with $\operatorname{supp}(x) = v$.

  (ii) *The matrix K is diagonalizable.*

 (iii) *If the set $\{x \in \mathcal{S} : w(x) > 0\}$ generates $\mathcal{S}$, then K has a unique stationary distribution π and*

$$(6.7) \qquad \|K_c^l - \pi\|_{\mathrm{TV}} \;\leq\; P\{x_1 x_2 \cdots x_l \notin \mathcal{C}_\mathcal{S}\} \;\leq\; \sum_u \lambda_u^l,$$

      where K_c^l is the distribution of the chain started from $c \in \mathcal{C}_\mathcal{S}$ after l steps, $(x_1, x_2, \dots)$ consists of independent and identically distributed picks from the measure w on $\mathcal{S}$ and u runs through the set of elements of L covered by $\hat{1}$.

The results of Section 3 can be extended to this setting as follows. Let $\phi : \mathcal{S}_\mathcal{A} \to \mathcal{S}_\mathcal{B}$ be an epimorphism of finite semigroups, meaning that ϕ is a surjective map which satisfies $\phi(xy) = \phi(x)\phi(y)$ for all $x, y \in \mathcal{S}_\mathcal{A}$. Given a left ideal $\mathcal{C}$ of $\mathcal{S}_\mathcal{A}$ and a probability measure w on $\mathcal{S}_\mathcal{A}$, the semigroup walk (6.1) on $\mathcal{C}$ associated to w induces a stochastic process on the state space $\phi(\mathcal{C})$, in the sense of Section 2.3. Since ϕ is surjective, the image $\phi(\mathcal{C})$ is a left ideal of $\mathcal{S}_\mathcal{B}$. This setup generalizes that of the map $f : \mathcal{F}_\mathcal{A} \to \mathcal{F}_\mathcal{B}$ of face semigroups of Section 3, where $\mathcal{B}$ is a subarrangement of a hyperplane arrangement $\mathcal{A}$ and $f(F)$ is the unique face of $\mathcal{B}$ which contains F, for $F \in \mathcal{F}_\mathcal{A}$. The following proposition generalizes Corollary 3.2.

Proposition 6.2. *Let $\phi : \mathcal{S}_\mathcal{A} \to \mathcal{S}_\mathcal{B}$ be an epimorphism of semigroups, $\mathcal{C} \subseteq \mathcal{S}_\mathcal{A}$ be a left ideal and w be a probability measure on $\mathcal{S}_\mathcal{A}$. For every starting distribution on $\mathcal{C}$, the stochastic process on $\phi(\mathcal{C})$ which is induced from the semigroup walk on $\mathcal{C}$ associated to w by the map ϕ is Markov. Moreover, this induced chain is itself a semigroup walk on $\phi(\mathcal{C})$, with associated probability measure w^* on $\mathcal{S}_\mathcal{B}$ defined by*

$$(6.8) \qquad\qquad w^*(z) \;=\; \sum_{x \in \mathcal{S}_\mathcal{A} \,:\, \phi(x) = z} w(x).$$

Proof. This follows by computations similar to those in the proofs of Proposition 3.1 and Corollary 3.2. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Suppose now that $\phi : \mathcal{S}_\mathcal{A} \to \mathcal{S}_\mathcal{B}$ is an epimorphism of finite bands. The definition of the support semilattice in [12, Section A.2] and (6.2) imply that ϕ induces an order preserving, surjective map $\phi_* : L_\mathcal{A} \to L_\mathcal{B}$ of the associated semilattices which makes the diagram

$$
\begin{array}{ccc}
\mathcal{S}_\mathcal{A} & \xrightarrow{\ \ \phi\ \ } & \mathcal{S}_\mathcal{B} \\[4pt]
{\scriptstyle \mathrm{supp}_\mathcal{A}}\big\downarrow & & \big\downarrow{\scriptstyle \mathrm{supp}_\mathcal{B}} \\[4pt]
L_\mathcal{A} & \xrightarrow{\ \ \phi_*\ \ } & L_\mathcal{B}
\end{array}
$$

commute, where $\mathrm{supp}_\mathcal{A}$ and $\mathrm{supp}_\mathcal{B}$ are the support maps of $\mathcal{S}_\mathcal{A}$ and $\mathcal{S}_\mathcal{B}$, respectively. We denote by $\mathcal{C}_\mathcal{A}$ and $\mathcal{C}_\mathcal{B}$ the set of chambers of $\mathcal{S}_\mathcal{A}$ and $\mathcal{S}_\mathcal{B}$, respectively. Since $\mathcal{C}_\mathcal{A}$ is a left ideal of $\mathcal{S}_\mathcal{A}$, the image $\phi(\mathcal{C}_\mathcal{A})$ is a left ideal of $\mathcal{S}_\mathcal{B}$.

Lemma 6.3. *Let $\mathcal{S}_\mathcal{A}$ and $\mathcal{S}_\mathcal{B}$ be finite bands with sets of chambers $\mathcal{C}_\mathcal{A}$ and $\mathcal{C}_\mathcal{B}$, respectively, and let $\phi : \mathcal{S}_\mathcal{A} \to \mathcal{S}_\mathcal{B}$ be an epimorphism of semigroups.*

   (i) *We have $\phi(\mathcal{C}_\mathcal{A}) \subseteq \mathcal{C}_\mathcal{B}$.*
   (ii) *If $\mathcal{S}_\mathcal{A}$ is left-regular, then $\phi(\mathcal{C}_\mathcal{A}) = \mathcal{C}_\mathcal{B}$.*

Proof. To prove (i), suppose that $c \in \mathcal{C}_\mathcal{A}$. Then we have $\mathrm{supp}_\mathcal{A}(x) \le \mathrm{supp}_\mathcal{A}(c)$ in $L_\mathcal{A}$ for every $x \in \mathcal{S}_\mathcal{A}$. By (6.2), this means that $c = cxc$ holds in $\mathcal{S}_\mathcal{A}$ for every $x \in \mathcal{S}_\mathcal{A}$. Since ϕ is an epimorphism of semigroups, it follows that $\phi(c) = \phi(c)z\phi(c)$ for every $z \in \mathcal{S}_\mathcal{B}$. By reversing the first part of the argument, we conclude that $\phi(c) \in \mathcal{C}_\mathcal{B}$.

Assume now that $\mathcal{S}_\mathcal{A}$ is left-regular. It was shown in [11, Sections 2.2 and B.3] that the relation $\preceq_\mathcal{A}$, defined by letting $x \preceq_\mathcal{A} y \Leftrightarrow xy = y$ for $x, y \in \mathcal{S}_\mathcal{A}$, is a partial order on $\mathcal{S}_\mathcal{A}$ and that the chambers of $\mathcal{S}_\mathcal{A}$ are precisely the maximal elements of $\preceq_\mathcal{A}$. Similar remarks hold for the band $\mathcal{S}_\mathcal{B}$, which is also left-regular as a homomorphic image of $\mathcal{S}_\mathcal{A}$. To prove (ii), suppose that $d \in \mathcal{C}_\mathcal{B}$ and let $x \in \mathcal{S}_\mathcal{A}$ be such that $\phi(x) = d$. Then there exists $c \in \mathcal{C}_\mathcal{A}$ such that $x \preceq_\mathcal{A} c$. Clearly, the map $\phi : \mathcal{S}_\mathcal{A} \to \mathcal{S}_\mathcal{B}$ is order preserving and hence $d \preceq_\mathcal{B} \phi(c)$. Since $d \in \mathcal{C}_\mathcal{B}$ is maximal in $\preceq_\mathcal{B}$, we must have $d = \phi(c)$. This shows that $d \in \phi(\mathcal{C}_\mathcal{A})$ and hence that $\mathcal{C}_\mathcal{B} \subseteq \phi(\mathcal{C}_\mathcal{A})$. In view of part (i), it follows that $\phi(\mathcal{C}_\mathcal{A}) = \mathcal{C}_\mathcal{B}$. $\qquad\qquad\square$

Lemma 6.3 implies that if $\phi : \mathcal{S}_\mathcal{A} \to \mathcal{S}_\mathcal{B}$ is an epimorphism of finite left-regular bands and $\mathcal{C} = \mathcal{C}_\mathcal{A}$, then the induced Markov chain of Proposition 6.2 is a semigroup walk on the state space $\mathcal{C}_\mathcal{B}$ of chambers of $\mathcal{S}_\mathcal{B}$. Thus all conclusions of Theorem 6.1 apply to the

induced chain. We leave it to the reader to formulate the exact analogue of Theorem 3.3 in this situation and end with a remark on the rest of the material of Section 3.

Remark 6.4. The proofs of parts (i) and (iv) of Theorem 2.1, given in Section 3, extend easily in the setup of Theorem 6.1. For part (i), for instance, one should replace (3.13) by the equality

$$(6.9) \qquad \#\{c \in \mathcal{C}_\mathcal{S} : xc = c\} = \sum_{u \in L:\ \mathrm{supp}(x) \leq_L u} m_u$$

for $x \in \mathcal{S}$, where the m_u are defined by (6.6). A computation similar to that given in the proof of Theorem 2.1 (i) in Section 3, equation (6.3) and a slight variant of Lemma 3.4 then imply that the m_u are necessarily nonnegative integers and that (6.4) holds. To check the validity of (6.9), we observe that the set $\{c \in \mathcal{C}_\mathcal{S} : xc = c\}$ is equal to the set of chambers of $x\mathcal{S}$ (see, for instance, [12, Example A.13]). Thus the left-hand side of (6.9) equals $n(v)$, where $v = \mathrm{supp}(x) \in L$, and hence (6.9) is equivalent to

$$n(v) = \sum_{v \leq_L u} m_u$$

for $v \in L$. This is in turn equivalent to (6.6) by Möbius inversion on L. We leave the details of the coupling proof of (6.7) to the interested reader. $\square$

Acknowledgements

References

[1] M. Aguiar, S. Mahajan, Coxeter Groups and Hopf Algebras, Amer. Math. Society, Providence, RI, 2006.

[2] D. Aldous, Random walks on groups and rapidly mixing Markov chains, in: Seminar on Probability XVII, Springer Lecture Notes in Mathematics **986**, Springer, Berlin, 1983, pp. 243–297.

[3] S. Assaf, P. Diaconis, K. Soundararajan, A rule of thumb for riffle schuffling, `arXiv:0908.3462`, Ann. Appl. Probab. (to appear).

[4] C.A. Athanasiadis, Spectra of some interesting combinatorial matrices related to oriented spanning trees on a directed graph, J. Algebraic Combin. **5** (1996) 5–11.

[5] C.A. Athanasiadis, Characteristic polynomials of subspace arrangements and finite fields, Adv. Math. **122** (1996) 193–233.

[6] A.D. Barbour, L. Holst, S. Janson, Poisson Approximation, Oxford University Press, New York, 1992.

[7] J. Barrera, T. Huillet, C. Paroissin, Limiting search cost for move-to-front rule with random request probabilities, Oper. Res. Lett. **34** (2006) 557–563.

[8] D. Bayer, P. Diaconis, Trailing the dovetail shuffle to its lair, Ann. Appl. Probab. **2** (1992) 294–313.

[9] 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.

[10] A. Björner, M. Las Vergnas, B. Sturmfels, N. White, G.M. Ziegler, Oriented Matroids, Cambridge University Press, Cambridge, 1999.

[11] K.S. Brown, Semigroups, rings, and Markov chains, J. Theoret. Probab. **13** (2000) 871–938.

[12] K.S. Brown, Semigroup and ring theoretical methods in probability, in: Representations of finite dimensional algebras in Lie theory and geometry, Fields Inst. Commun. **40**, Amer. Math. Society, Providence, RI, 2004, pp. 3–26.

[13] K.S. Brown, P. Diaconis, Random walks and hyperplane arrangements, Ann. Probab. **26** (1998) 1813–1854.

[14] P. Diaconis, Group Representations in Probability and Statistics, IMS, Hayward, CA, 1988.

[15] P. Diaconis, Mathematical developements from the analysis of riffle shuffling, in: Groups, Combinatorics & Geometry (Durham, 2001), World. Sci. Publ, River Edge, NJ, 2003, pp. 73–97.

[16] P. Diaconis, J.A. Fill, J. Pitman, Analysis of top to random shuffles, Combin. Probab. Comput. **1** (1992) 135–155.

[17] P. Diaconis, J. Fulman, Carries, shuffling and symmetric functions, Adv. in Appl. Math. **43** (2009) 176–196.

[18] P. Cellini, A general commutative descent algebra, J. Algebra **175** (1995) 990–1014.

[19] P. Diaconis, B. Steinberg, Colored shuffles and random walks on semigroups, preprint, Stanford University, 2009.

[20] R.P. Dobrow, J.A. Fill, The move-to-front rule for self-organizing lists with Markov dependent requests, in: Discrete Probability and Algorithms, IMA Vol. Math. Appl. **72**, 1995, pp. 57–80.

[21] J.A. Fill, Limits and rates of convergence for the distribution of search cost under the move-to-front rule, Theoret. Comput. Sci. **164** (1996) 185–206.

[22] D. Freedman, From association to causation via regression, Adv. in Appl. Math. **18** (1997) 59–110.

[23] J. Fulman, Affine shuffles, shuffles with cuts, the Whitehouse module and patience sorting, J. Algebra **231** (2000) 614–639.

[24] S.B. Gillispie, M.D. Perlman, The size distribution for Markov equivalence classes of acyclic digraph models, Artificial Intelligence **141** (2002) 137–155.

[25] P. Hanlon, A Hodge decomposition interpretation for the coefficients of the chromatic polynomial, Proc. Amer. Math. Soc. **136** (2008) 3741–3749.

[26] J.G. Kemeny, J.L. Snell, Finite Markov chains, D. Van Nostrand Co., Inc., Princeton NJ/Toronto/London/New York, 1960.

[27] V.F. Kolchin, B.A. Sevast'yanov, V.P. Chistyakov, Random Allocations, John Wiley & Sons, New York-Toronto, Ont.-London, 1978.

[28] S.L. Lauritzen, Graphical models, Oxford University Press, New York, 1996.

[29] D. Levin, Y. Peres, E. Wilmer, Markov chains and mixing times, Amer. Math. Society, Providence, RI, 2009.

[30] B. Mann, How many times should you shuffle a deck of cards?, UMAP J. **15** (1994), 303–332; reprinted in: J.L. Snell, (Ed.), Topics in Contemporary Probability and its Applications, CRC Press, Boca Raton, FL, 1995, pp. 261–289.

[31] P. Orlik, H. Terao, Arrangements of Hyperplanes, Springer-Verlag, New York, 1992.

[32] R.M. Phatarfod, On the matrix occurring in a linear search problem, J. Appl. Probab. **28** (1991) 336–346.

[33] A. Postnikov, R.P. Stanley, Deformations of Coxeter hyperplane arrangements, J. Combin. Theory Series A **91** (2000) 544–597.

[34] V. Retakh, R. Wilson, Algebras associated to acyclic directed graphs, Adv. in Appl. Math. **42** (2009) 42–59.

[35] L.C.G. Rogers, J.W. Pitman, Markov functions, Ann. Probab. **9** (1981) 573–582.

[36] L. Saloff-Coste, J. Zúñiga, Convergence of some time inhomogeneous Markov chains via spectral techniques, Stochastic Process. Appl. **117** (2007) 961–979.

[37] R.P. Stanley, Acyclic orientations of graphs, Discrete Math. **5** (1973) 171–178.

[38] R.P. Stanley, Enumerative Combinatorics, vol. 1, Wadsworth & Brooks/Cole, Pacific Grove, CA, 1986; second printing, Cambridge University Press, Cambridge, 1997.

[39] R.P. Stanley, Hyperplane arrangements, interval orders and trees, Proc. Nat. Acad. Sci. **93** (1996) 2620–2625.

[40] R.P. Stanley, Enumerative Combinatorics, vol. 2, Cambridge University Press, Cambridge, 1999.

[41] R.P. Stanley, An Introduction to Hyperplane Arrangements, in: E. Miller, V. Reiner, B. Sturmfels, (Eds.), Geometric Combinatorics, Amer. Math. Society, Providence, RI, 2007, pp. 389–496.

[42] B. Steinberg, Möbius functions and semigroup representation theory, J. Combin. Theory Series A **113** (2006) 866–881.

[43] B. Steinberg, Möbius functions and semigroup representation theory II: Character formulas and multiplicities, Adv. Math. **217** (2008) 1521–1557.

[44] E. Steingrímsson, The coloring ideal and coloring complex of a graph, J. Algebraic Combin. **14** (2001) 73–84.

[45] J. Tits, Buildings of spherical type and finite BN-pairs, Lecture Notes in Mathematics **386**, Springer-Verlag, Berlin, 1974.

[46] J.-C. Uyemura Reyes, Random walk, semi-direct products and card shuffling, Ph.D thesis, Stanford University, 2002.

[47] T. Zaslavsky, Facing up to arrangements: face-count formulas for partitions of space by hyperplanes, Mem. Amer. Math. Soc. **154** (1975).

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ATHENS, PANEPISTIMIOUPOLIS, ATHENS 15784, GREECE

E-mail address: `caath@math.uoa.gr`

DEPARTMENT OF MATHEMATICS AND STATISTICS, STANFORD UNIVERSITY, STANFORD, CA 94305