PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 0, Fasc. 0 (0000), pp. 000–000

STRONG STATIONARY DUALITY FOR MÖBIUS MONOTONE MARKOV CHAINS: EXAMPLES*

BY

PAWEŁ LOREK AND RYSZARD SZEKLI

Abstract. We construct strong stationary dual chains for non-symmetric random walks on square lattice, for random walks on hypercube and for some Ising models on the circle. The strong stationary dual chains are all sharp and have the same state space as original chains. We use Möbius monotonicity of these chains with respect to some natural orderings of the corresponding state spaces. This method provides an alternative way to study mixing times for studied models.

2000 AMS Mathematics Subject Classification: Primary: 60J10; Secondary: 06A06; 60G40

Key words and phrases: Markov chains; stochastic monotonicity; eigenvalues; Möbius monotonicity; strong stationary duality; strong stationary times; separation distance; mixing time; Ising model; hypercube

1. INTRODUCTION

Consider an ergodic Markov chain $\mathbf{X} = (X_n)_{n\geqslant 0}$ on a discrete (finite or countable) state space $\mathbb{E}$ with transition matrix $\mathbf{P}$ and initial distribution ν. One way of studying the speed of convergence of $\mathbf{X}$ to its stationary distribution π is to find (and bound its tail) so-called *Strong Stationary Time* (SST), i.e. such a stopping

* Work of both authors supported by NCN Research Grant DEC-2011/01/B/ST1/01305

time T (T implicitly depends on ν) that it is independent from X_T, and X_T has distribution π. SST's were introduced by Aldous and Diaconis [2, 3], who also gave examples of SST and their applications. Many examples can also be found in Diaconis [4]. First examples of SST's were created by *ad hoc* methods. A general approach was invented by Diaconis and Fill [5] who introduced *dual processes*. They showed that for $\mathbf{X}$ there always exists so-called *Strong Stationary Dual* (SSD) absorbing chain $\mathbf{X}^*$, such that its time to absorption T^* is equal, in distribution, to a SST T for $\mathbf{X}$. Their proof is an existence type argument which does not show how to construct a dual chain in general. They showed one tractable case [5, Theorem 4.6], where the state space is linearly ordered. Under the condition of stochastic monotonicity (related to the linear order) of the corresponding time-reversed chain (and some assumptions on the initial distribution) they gave a recipe of how to construct a dual chain on the same state space. A special, and important, case is a stochastically monotone birth-and-death chain for which the dual chain is an absorbing birth-and-death chain.

Strong stationary dual chains have a variety of applications. Diaconis and Fill [6] gave an extension of this theory to countable state spaces. Fill [12] gave a stochastic proof of a well-known theorem (usually attributed to Keilson), which states that the first passage time from 0 to M of a stochastically monotone birth-and-death process on $\{0, \ldots, M\}$ is equal, in distribution, to a sum of geometric random variables related to the spectral values of $\mathbf{X}$. Similar results for continuous time birth-and-death processes were obtained by Diaconis and Miclo [7]. Diaconis and Saloff-Coste [8] studied cut-off phenomena for birth-and-death chains by using SSD theory. Different dualities in Markov chains are utilized in a variety of contexts, see, e.g., Huillet and Martinez [16].

All the mentioned examples above (although very interesting) somehow rely on Theorem 4.6 of [5] which involves linearly ordered states space. That is why

most of the known examples are related to birth-and-death chains. The main underlying assumption is (classical) stochastic monotonicity of the time-reversed chain. Although this monotonicity is defined also for partially ordered state spaces, it is not sufficient for an analogous construction of a SSD chain as in Diaconis and Fill [5]. Lorek and Szekli [20] gave a recipe of how to construct dual chains on partially ordered state spaces with a special feature that the duals have the same state space as original chains. The assumption of the classical stochastic monotonicity was replaced by the assumption of Möbius monotonicity. This extension (to partially ordered state spaces) opens a new way of finding SSD chains defined for not linearly ordered state spaces. The purpose of this paper is to get a new SSD insight to some classical examples of finite state Markov chains. In section 2 we recall needed definitions and facts about Möbius monotone chains. In section 3 we present strong stationary duals for non-symmetric random walk on a square lattice, for a random walk on the hypercube, and some Ising models on the circle. For the latter one we give duals for specific cases, and conjecture the general case.

In section 4 we give proofs of the main results. We believe that the presented method should be applicable for many other examples and can be used to find bounds on the speed of convergence to stationarity, and to study cut-off phenomena.

2. MÖBIUS MONOTONICITY AND DUALITY

In this section we recall needed results on SSD and Möbius monotone chains. For a more complete material on duality see Diaconis and Fill [5], and for results on Möbius monotone chains, see Lorek and Szekli [20].

2.1. Strong Stationary Duality. For an ergodic Markov chain $\mathbf{X} = (X_n)_{n\geqslant 0}$ with the transition matrix $\mathbf{P}$ and initial distribution ν, we are interested in bounding a distance between $\nu\mathbf{P}^k$ (a distribution of a chain at step k) and its stationary

distribution π. Often used distance is the total variation distance $d_{TV}(\nu\mathbf{P}^k, \pi) = \max_{A\subset\mathbb{E}} |\nu\mathbf{P}^k(A) - \pi(A)|$. Another useful distance is the *separation distance* s defined as follows: $s(\nu\mathbf{P}^k, \pi) = \max_{\mathbf{e}\in\mathbb{E}}(1 - \nu\mathbf{P}^k(\mathbf{e})/\pi(\mathbf{e}))$. For random times T which are SST, Aldous and Diaconis [3] show that $d_{TV}(\nu\mathbf{P}^k, \pi) \leqslant s(\nu\mathbf{P}^k, \pi) \leqslant P(T > n)$.

Let $\mathbf{X}^*$ be a Markov chain with transition matrix $\mathbf{P}^*$, initial distribution ν^* and a state space $\mathbb{E}^*$, with an absorbing state $\mathbf{e}_a^*$. Let $\Lambda \equiv \Lambda(\mathbf{e}^*, \mathbf{e}), \mathbf{e}^* \in \mathbb{E}^*, \mathbf{e} \in \mathbb{E}$ be a stochastic kernel (called *a link*), such that $\Lambda(\mathbf{e}_a^*, \cdot) = \pi$, for $\mathbf{e}_a^* \in \mathbb{E}^*$. $\mathbf{X}^*$ is a Strong Stationary Dual (SSD) chain for $\mathbf{X}$ if

$$\nu = \nu^*\Lambda \quad \text{and} \quad \Lambda\mathbf{P} = \mathbf{P}^*\Lambda. \tag{2.1}$$

Diaconis and Fill [5] proved that the absorption time T^* of $\mathbf{X}^*$ is a SST for $\mathbf{X}$. Thus, the problem of finding SST for $\mathbf{X}$ translates into the problem of studying the absorption time of $\mathbf{X}^*$.

DEFINITION 2.1. Strong Stationary Dual chain $\mathbf{X}^*$ is called *sharp* if $s(\nu\mathbf{P}^n, \pi) = P(T^* > n)$.

REMARK 2.1. The relation (2.1) implies that for finite $\mathbb{E}$ and $\mathbb{E}^*$, $\mathbf{P}$ and $\mathbf{P}^*$ have the same set of eigenvalues.

It turns out, that in some examples we can easily identify the eigenvalues of $\mathbf{P}^*$, and thus, by the above remark, we will also obtain the eigenvalues of $\mathbf{P}$ which are usually not easy to obtain directly.

2.2. Duality for Möbius monotone chains. In this section we recall how to construct a SSD chain for finite partially ordered state spaces. We shall consider a finite state space $\mathbb{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_M\}$ with a partial ordering $\preceq$. From the very beginning we shall choose an enumeration of $\mathbb{E}$ such that $\mathbf{e}_i \preceq \mathbf{e}_j$ implies $i < j$ (which is always possible). We call such an enumeration consistent with $\preceq$. With

this enumeration the partial ordering can by represented by an upper-triangular, 0-1 valued matrix $\mathbf{C}$. The inversion $\mathbf{C}^{-1}$ represents (in the incidence algebra) the so called *Möbius function*, usually denoted by μ, see Rota [23]. The Möbius function allows for the following calculus: it is possible to recover f from the relation $\bar{F}(\mathbf{e}) = \sum_{\mathbf{e}:\mathbf{e}\succeq\mathbf{e}_i} f(\mathbf{e})$, namely $f(\mathbf{e}_i) = \sum_{\mathbf{e}:\mathbf{e}\succeq\mathbf{e}_i} \mu(\mathbf{e}_i,\mathbf{e})\bar{F}(\mathbf{e})$.

DEFINITION 2.2. Let $\mathbf{P}$ be a transition matrix with enumaration of states consistent with $\mathbf{C}$. We say that $\mathbf{P}$ (or alternatively, $\mathbf{X}$) is $^{\downarrow}$-Möbius monotone ($^{\uparrow}$-Möbius monotone) if $\mathbf{C}^{-1}\mathbf{P}\mathbf{C} \geqslant 0$ ($(\mathbf{C}^T)^{-1}\mathbf{P}\mathbf{C}^T \geqslant 0$) (each entry is nonnegative).

We say that $\mathbf{f} : \mathbb{E} \to \mathbb{R}^M$ is $^{\downarrow}$-Möbius monotone ($^{\uparrow}$-Möbius monotone) if $\mathbf{f}(\mathbf{C}^T)^{-1} \geqslant 0$ ($\mathbf{f}\mathbf{C}^{-1} \geqslant 0$). In terms of the transition probabilities, we have

$$^{\downarrow}\text{-Möbius monotonicity:} \quad \forall(\mathbf{e}_i,\mathbf{e}_j \in \mathbb{E}) \quad \sum_{\mathbf{e}:\mathbf{e}\succeq\mathbf{e}_i} \mu(\mathbf{e}_i,\mathbf{e})\, \mathbf{P}(\mathbf{e},\{\mathbf{e}_j\}^{\downarrow}) \geqslant 0,$$

$$^{\uparrow}\text{-Möbius monotonicity:} \quad \forall(\mathbf{e}_i,\mathbf{e}_j \in \mathbb{E}) \quad \sum_{\mathbf{e}:\mathbf{e}\preceq\mathbf{e}_j} \mathbf{P}(\mathbf{e},\{\mathbf{e}_i\}^{\uparrow})\mu(\mathbf{e},\mathbf{e}_j) \geqslant 0,$$

where $\{\mathbf{e}_j\}^{\downarrow} = \{\mathbf{e} : \mathbf{e} \preceq \mathbf{e}_j\}$, $\{\mathbf{e}_j\}^{\uparrow} = \{\mathbf{e} : \mathbf{e} \succeq \mathbf{e}_j\}$, and $\mathbf{P}(\mathbf{e},A) = \sum_{\mathbf{e}'\in A} \mathbf{P}(\mathbf{e},\mathbf{e}')$.

We recall the SSD result of Lorek and Szekli [20] ($\overleftarrow{\mathbf{X}}$ denotes the time-reversed process).

THEOREM 2.1 (Lorek and Szekli [20]). Let $\mathbf{X}$ be an ergodic Markov chain on a finite state space $\mathbb{E} = \{\mathbf{e}_1,\ldots,\mathbf{e}_M\}$, which is partially ordered with $\preceq$, and has a unique maximal state $\mathbf{e}_M$. For the stationary distribution π and an initial distribution ν we assume that

(i) $g(\mathbf{e}) = \frac{\nu(\mathbf{e})}{\pi(\mathbf{e})}$ is $^{\downarrow}$-Möbius monotone,

(ii) $\overleftarrow{\mathbf{X}}$ is $^{\downarrow}$-Möbius monotone.

Then there exists a Strong Stationary Dual chain $\mathbf{X}^*$ on $\mathbb{E}^* = \mathbb{E}$ with link being a truncated stationary distribution $\Lambda(\mathbf{e}_j,\mathbf{e}_i) = \mathbb{I}(\mathbf{e}_i \preceq \mathbf{e}_j)\frac{\pi(\mathbf{e}_i)}{H(\mathbf{e}_j)}$, where $H(\mathbf{e}_j) = \sum_{\mathbf{e}:\mathbf{e}\preceq\mathbf{e}_j} \pi(\mathbf{e})$. The initial distribution and transitions of $\mathbf{X}^*$ are given, respectively,

by

$$\begin{aligned} \nu^*(\mathbf{e}_i) &= H(\mathbf{e}_i) \sum_{\mathbf{e}:\mathbf{e}\succeq\mathbf{e}_i} \mu(\mathbf{e}_i,\mathbf{e}) g(\mathbf{e}), \\ \mathbf{P}^*(\mathbf{e}_i,\mathbf{e}_j) &= \frac{H(\mathbf{e}_j)}{H(\mathbf{e}_i)} \sum_{\mathbf{e}:\mathbf{e}\succeq\mathbf{e}_j} \mu(\mathbf{e}_j,\mathbf{e}) \overleftarrow{\mathbf{P}}(\mathbf{e},\{\mathbf{e}_i\}^{\downarrow}). \end{aligned} \tag{2.2}$$

REMARK 2.2. Following Remark 2.39 of Diaconis and Fill [5] and the terminology used there, the Strong Stationary Dual $\mathbf{X}^*$ in Theorem 2.1 is sharp, and the corresponding strong stationary time is the time to stationarity, i.e., $s(\nu\mathbf{P}^n,\pi) = P(T>n)$. The reason for this is that $\Lambda(\mathbf{e}^*,\mathbf{e}_M)=0$ for all, $\mathbf{e}^*\neq\mathbf{e}_M\in\mathbb{E}^*$.

REMARK 2.3. Theorem 2.1 is stated for $^{\downarrow}$-Möbius monotonicity, but it can be similarly stated for $^{\uparrow}$-Möbius monotonicity (see Corollary 3.1 in [20]). The other formulation is potentially useful, because a chain can be, e.g., $^{\downarrow}$-Möbius monotone but not $^{\uparrow}$-Möbius monotone.

REMARK 2.4. The assumption on the initial distribution is not very restrictive, for example if $\mathbf{e}_1$ is a unique minimal state and $\nu=\delta_{\mathbf{e}_1}(\cdot)$, then the assumption is fulfilled, and also $\nu^*=\delta_{\mathbf{e}_1}(\cdot)$. For simplicity of presentation, in all subsequent examples the initial distribution will be the single atom at the minimal element (this assumption may be relaxed).

In order to find and use the above constructed SSD chains one has to find *an approperiate* ordering (w.r.t which the chain is Möbius monotone). It is worth mentioning, that for linearly ordered state space $^{\uparrow}$-Möbius monotonicity is equivalent to the usual stochastic monotonicity, in general partially ordered spaces this is not the case. It turns out that for partially ordered spaces some *natural* orderings work. A non-symmetric random walk on the unit cube is an example presented in [20]. In the next section we shall give new examples.

3. MÖBIUS MONOTONE MARKOV CHAINS: EXAMPLES

3.1. Random walk on weighted directed graph. Consider a random walk on a directed weighted graph $G = (V, E)$ with vertices $V = \{v_1, v_2, \ldots, v_n\}$, edges $E = \{(i, j) : \text{edge from } v_i \text{ to } v_j\}$ and with a weighting function $w : E \to [0, \infty)$. Denote by $w_{i,j}$ the nonnegative weight of the directed edge from node v_i to v_j. If there is no edge between these nodes, i.e., $(i, j) \notin E$, then $w_{i,j} = 0$. We allow $w_{i,i}$ be nonzero.

Let $\mathcal{N}(i) = \{j : (i, j) \in E\}$ be a set of neighbours of node v_i. Random walk may be viewed as a process of sequential vertex visiting. We assume that weights are normalized, i.e., for all $i \in \{1, \ldots, n\}$ we have $w_{i,i} + \sum_{r\in\mathcal{N}(i)} w_{i,r} = 1$. The probability of a single step from node i to j is then given by $P(i, j) = w_{i,j}$.

In this section we consider the following example: Let $V = \{0, 1, \ldots, N\}^2$ with edges

$$((x_1, y_1), (x_2, y_2)) \in E \iff |x_1 - x_2| + |y_1 - y_2| = 1 \tag{3.1}$$

for $x_1, x_2, y_1, y_2 \in \{0, \ldots, N\}$. Thus, for each node there are at most four edges in four directions: *up, down, left, right* plus a possible self-loop. The weighting function depends only on the direction in the following way: for $((x_1, y_1), (x_2, y_2)) \in E$ and nonnegative parameters $\lambda_1, \lambda_2, \mu_1, \mu_2$ such that $\lambda_1 + \lambda_2 + \mu_1 + \mu_2 \leqslant 1$

(3.2)

$$w_{((x_1,y_1),(x_2,y_2))} = \begin{cases} \lambda_1 & \text{if} \quad x_2 = x_1 + 1, y_2 = y_1, \\ \mu_1 & \text{if} \quad x_2 = x_1 - 1, y_2 = y_1, \\ \lambda_2 & \text{if} \quad x_2 = x_1, y_2 = y_1 + 1, \\ \mu_2 & \text{if} \quad x_2 = x_1, y_2 = y_1 - 1, \\ 1 - \displaystyle\sum_{(x,y)\in\mathcal{N}((x_1,y_1))} w_{((x_1,y_1),(x,y))} & \text{if} \quad x_2 = x_1, y_2 = y_1. \end{cases}$$

We associate weights directly with one step probabilities:

$$\mathbf{P}((x_1, y_1), (x_2, y_2)) = w_{((x_1,y_1),(x_2,y_2))}.$$

Roughly speaking, we consider a random walk on square lattice $\{0, \ldots, N\}^2$, at each step we can move (if feasible): *right* with probability λ_1, *left* with probability μ_1, *up* with probability λ_2 and *down* with probability μ_2. With remnant probability we stay at a given vertex. For convenience, we let $\rho_1 := \lambda_1/\mu_1$, and $\rho_2 := \lambda_2/\mu_2$. Denote the transition matrix of a corresponding Markov chain $\mathbf{X}$ by $\mathbf{P}$. The chain is time-reversible (i.e. $\overleftarrow{\mathbf{P}} = \mathbf{P}$) and has (time-reversibility equations can be easily checked) the stationary distribution on V

$$\pi((x, y)) = C^{-1}\rho_1^x \rho_2^y$$

for $(x, y) \in V = \{0, \ldots, N\}^2$, where the normalizing constant C for $\rho_1 \neq 1$ and $\rho_2 \neq 1$ is given by

$$C = \frac{1 - \rho_1^{N+1}}{1 - \rho_1} \cdot \frac{1 - \rho_2^{N+1}}{1 - \rho_2},$$

and C for other cases can be obtained by obvious modifications.

We shall use the coordinate-wise partial ordering $(x_1, y_1) \preceq (x_2, y_2) \iff x_1 \leqslant x_2$ and $y_1 \leqslant y_2$. Then we have unique minimal element $\mathbf{e}_1 = (0, 0)$ and the maximal one $\mathbf{e}_M = (N, N)$, where $M = (N + 1)^2$. It turns out that $\mathbf{X}$ is Möbius monotone for any set of parameters $\lambda_1, \mu_1, \lambda_2, \mu_2 > 0$, such that $\lambda_1 + \lambda_2 + \mu_1 + \mu_2 \leqslant 1$, and applying Theorem 2.1 we have:

THEOREM 3.1. Let $\mathbf{X}$ be a random walk on directed weighted graph with $G = (V, E)$, with $V = \{0, \ldots, N\}^2$, and E given in (3.1), weights given in (3.2) and with positive parameters $\lambda_1 \neq \mu_1,\ \lambda_2 \neq \mu_2$, such that $\lambda_1 + \lambda_2 + \mu_1 + \mu_2 \leqslant 1$. Assume, that $\mathbf{X}$ starts at $\mathbf{e}_1 = (0, 0)$. Then there exists sharp SSD chain $\mathbf{X}^*$ which is an absorbing Markov chain (with $\mathbf{e}_M = (N, N)$ being the single absorbing state) on the state space $\mathbb{E}^* = \mathbb{E} = \{0, \ldots, N\}^2$, starting at $\mathbf{e}_1 = (0, 0)$, with the follow-

ing transition probabilities (for $x, x', y, y' \in \{0, \ldots, N\}$)

(3.3)

$$\mathbf{P}^*((x,y),(x',y')) =$$

$$\begin{cases} \frac{1-\rho_1^{x+2}}{1-\rho_1^{x+1}} \cdot \mu_1 & \text{if } x' = x+1,\ y' = y \\ \frac{1-\rho_2^{y+2}}{1-\rho_2^{y+1}} \cdot \mu_2 & \text{if } y' = y+1,\ x' = x \\ \frac{1-\rho_2^{y}}{1-\rho_2^{y+1}} \cdot \lambda_2 & \text{if } x' = x,\ y' = y-1,\ y \neq N \\ \frac{1-\rho_1^{x}}{1-\rho_1^{x+1}} \cdot \lambda_1 & \text{if } y' = y,\ x' = x-1,\ x \neq N \\ 1-(\lambda_1+\lambda_2+\mu_1+\mu_2) & \text{if } x' = x, y' = y,\ (x,y) \in \{0,\ldots,N-1\}^2 \\ 1-(\lambda_2+\mu_2) & \text{if } x' = x = N, y' = y,\ \ y \in \{0,\ldots,N-1\} \\ 1-(\lambda_1+\mu_1) & \text{if } x' = x, y' = y = N,\ x \in \{0,\ldots,N-1\} \\ 1 & \text{if } x' = x = y = y = N \end{cases}$$

Thus, the SSD chain $\mathbf{X}^*$ is again a chain on $\mathbb{E}$, with feasible moves in the same directions as $\mathbf{X}$ *except* for movements on the upper borders of this square lattice. Once the chain hits the border $(\cdot, N)$ (or $(N, \cdot)$), then it can only move *left* or *right* (*up* or *down*) until it hits the absorbing state (N, N). Note that probability of changing i-th coordinate, $i = 1, 2$, is independent of the value of $(3-i)$-th coordinate. The chain $\mathbf{X}^*$, for a suitable selection of the parameters, can have a *drift* towards the absorbing state. Note that the case $\rho_1 = 1$, and/or $\rho_2 = 1$ can be obtained by obvious modifications in computing $H(x, y)$ (see the proof in section 4.1).

One can study the time to absorption T^* in the following way: it is the time of hitting a border $(\cdot, N)$ or $(N, \cdot)$ plus the time for the one dimensional birth-and-death chain with birth probability λ_1 and death probability μ_1 (or λ_2 and μ_2 respectively) to reach the state N (worst cases scenarios can be used).

3.2. Random change of single coordinate on a cube. Let us consider a discrete time Markov chain $\mathbf{X}$ with state space $\mathbb{E} = \{0, \ldots, k\}^n$, which evolves in the following way: it stays with probability $1/2$ or (with probability $1/2$) for one coordinate chosen uniformly, it changes uniformly its value to any other different value. In terms of the transition probabilities, for $\mathbf{e} = (\mathbf{e}(1), \ldots, \mathbf{e}(n)) \in \mathbb{E}$, $\mathbf{e}(i) \in \{0, \ldots, k\}$, we set

(3.4)
$$\mathbf{P}(\mathbf{e}, \mathbf{e}') = \begin{cases} \frac{1}{2} & \text{if } \mathbf{e} = \mathbf{e}' \\ \frac{1}{2nk} & \text{if for some } i \quad \mathbf{e}(i) \neq \mathbf{e}'(i) \text{ and } \mathbf{e}(j) = \mathbf{e}'(j),\ j \neq i \\ 0 & \text{otherwise} \end{cases}$$

Since $\mathbf{P}$ is symmetric, the corresponding stationary distribution is uniform, i.e.,

$$\pi(\mathbf{e}) = \frac{1}{(k+1)^n},\ \mathbf{e} \in \mathbb{E}.$$

The motivation for this example comes from DNA sequence alignment. Given n sequences of length $k+1$ the task is to find points of references in each one such that, starting *reading* sequence i from it's reference point $r(i)$ we have the biggest agreement in all sequences. Since the state space is huge (of size $(k+1)^n$), often Monte Carlo methods are used. One constructs a chain such that its stationary distribution assigns higher mass to states with high agreements. The chain given in (3.4) is a simplified version of such a chain.

The chain $\mathbf{X}$ can be seen as an extension of the standard lazy random walk on the unit cube (obtained for $k = 1$). Using the coordinate-wise ordering $\preceq$ on $\mathbb{E}$, it turns out that $\mathbf{X}$ (which is reversible) is Möbius monotone. For this ordering, the state $\mathbf{e}_1 = (0, \ldots, 0)$ is the minimal state and $\mathbf{e}_M = (k, \ldots, k)$ is the maximal state (with $M = (k+1)^n$), where we use an enumeration of $\mathbb{E}$ consistent with $\preceq$. Applying Theorem 2.1 we obtain.

THEOREM 3.2. Consider the chain $\mathbf{X}$ described above, on state space $\mathbb{E} = \{0, \ldots, k\}^n$, with transition probabilities given in (3.4). Assume that $\mathbf{X}$ starts at $\mathbf{e}_1$. Then, there exists sharp SSD chain $\mathbf{X}^*$ on the state space $\mathbb{E}^* = \mathbb{E}$, with the state $\mathbf{e}_M$ being the absorbing one, starting with probability 1 at $\mathbf{e}_1$, and having transition probabilities, for all $A \subseteq \{1, \ldots, n\}$, $j \notin A$

$$\begin{aligned} \mathbf{P}^*(\mathbf{e}_A^{(k)}, \mathbf{e}_{A\cup\{j\}}^{(k)}) &= \frac{(k+1)}{2nk}, \\ \mathbf{P}^*(\mathbf{e}_A^{(k)}, \mathbf{e}_A^{(k)}) &= \frac{n(k-1)+|A|(k+1)}{2nk}, \end{aligned}$$

where $\mathbf{e}_A^{(k)} = (\mathbf{e}(1), \ldots, \mathbf{e}(n))$ with $\mathbf{e}(i) = k$ if $i \in A$ and $\mathbf{e}(i) = 0$ if $i \notin A$, and all other transitions have probability 0.

Note that SSD chain $\mathbf{X}^*$ jumps, with probability 1, only to greater or equal states in the ordering $\preceq$, thus its eigenvalues are the entries on the diagonal of the matrix $\mathbf{P}^*$ written using an enumeration of the states consistent with this ordering. The states which can be traversed by $\mathbf{X}^*$ are of the form $\mathbf{e}_A^{(k)}$, which means that $\mathbf{X}^*$ can be identified with a random walk on the unit cube $\{0, k\}^n$. Again, by Remark 2.1, the eigenvalues of $\mathbf{P}$ are the same as diagonal entries of $\mathbf{P}^*$, i.e.,

$$\frac{n(k-1)+i(k+1)}{2nk}, \quad i = 0, 1, \ldots, n.$$

Similarly as in the Ising model example, we can consider the time to absorption of one dimensional projection $Z_t^* := S(X_t^*)$, where $S(\mathbf{e}) = \sum_{i=1}^n \mathbf{1}\{\mathbf{e}(i) = k\}$. If $Z_0^* = 0$, then the time to absorption T^* of Z_t^* is the same as for X_t^*, and is distributed as the sum of independent variables $\sum_{i=0}^{n-1} Y_i$, where Y_i has geometric distribution with the success parameter $p_i = \frac{(n-i)(k+1)}{2nk}$. For the expected absorption time we have

$$ET^* = \sum_{i=0}^{n-1} \frac{1}{p_i} = \sum_{i=0}^{n-1} \frac{1}{n-i} \frac{2nk}{k+1} = \frac{2nk}{k+1} \sum_{i=1}^{n} \frac{1}{i} \leqslant \frac{2k}{k+1}(n+1)\log n$$

For the variance of T^* we have

$$VarT^* = \sum_{i=0}^{n-1} \frac{1-p_i}{p_i^2} = \frac{2nk}{(k+1)^2} \sum_{i=0}^{n-1} \frac{nk-n+ki+i}{(n-i)^2}$$

$$= \frac{2nk}{(k+1)^2}\left[nk\sum_{i=0}^{n-1}\frac{1}{(n-i)^2} + k\sum_{i=0}^{n-1}\frac{i}{(n-i)^2} - \sum_{i=1}^{n}\frac{1}{i}\right] \overset{(*)}{\leqslant} \left(\frac{2nk}{k+1}\right)^2\frac{\pi^2}{6},$$

where in $^{(*)}$ we used the following ineqalities

$$\sum_{i=0}^{n-1}\frac{1}{(n-i)^2} \leqslant \frac{\pi^2}{6}, \quad \sum_{i=0}^{n-1}\frac{i}{(n-i)^2} \leqslant n\frac{\pi^2}{6}.$$

By Remark 2.2 and from Chebyshev's inequality, we have that after $m = \frac{2k}{k+1}(n+1)\log n + c\frac{2k}{k+1}\frac{\pi}{\sqrt{6}}n, c \geqslant 0$ steps we have

$$s(\nu\mathbf{P}^m, \pi) = P(T > m) \leqslant P(T - ET \leqslant c\sqrt{Var}) \leqslant P(|T - ET| \leqslant c\sqrt{Var}) \leqslant \frac{1}{c^2}.$$

3.3. Ising model on a circle. Let $G = (V, E)$ be a finite graph. Elements of state space $\mathbb{E} = \{-1, 1\}^V$ are called *configurations*, and for $\mathbf{e} \in \mathbb{E}$ the value $\mathbf{e}(v)$ is called the *spin* at vertex v. For a given configuration $\mathbf{e}$ its *energy* is defined as

$$\mathcal{H}(\mathbf{e}) \quad = -\sum_{\{x,y\}\in E}\mathbf{e}(x)\cdot\mathbf{e}(y),$$

where the sum is over all edges of the graph. For $\beta \geqslant 0$, the **Ising model** on the graph G with parameter β is the probability measure on $\mathbb{E}$ given by

$$\pi(\mathbf{e}) = \frac{e^{-\beta\mathcal{H}(\mathbf{e})}}{Z_\beta}, \tag{3.5}$$

where $Z_\beta = \sum_{\mathbf{e}\in\mathbb{E}} e^{-\beta\mathcal{H}(\mathbf{e})}$ is a normalizing constant. The parameter β has a physical interpretation as the inverse of the temperature of the configuration. Note, that for $\beta = 0$ (equivalent to infinite temperature), every spin configuration is equally likely, i.e., it is the same as setting spin at each vertex to -1 or +1 with probability 1/2 independently. In general, β represents the influence of energy $\mathcal{H}$ on π.

This model has focused a lot of attention in the context of speed of convergence to equilibrium of particle systems. Propp and Wilson [22] introduced *Coupling From The Past* algorithm and used it to show how to draw exact sample from (3.5) in the case of square lattice. Recently Ding and Peres [10] showed that

for Ising models on each graph it takes at least $(1/4 + o(1))n \log n$ steps for the Glauber dynamics to mix, where n is the corresponding number of vertices. In Ding and Peres [11] a simple proof for the bound $n \log n/2$ was presented.

We shall consider the Ising model on a circle. We will present the dual chain for a case of 2 and 3 vertices and a conjecture on a general number of vertices. We set $V = \{0, \dots, N-1\}$ and $E = \{(i, (i+1) \mod N) : i = 0, \dots, N-1\}$. The stationary distribution (3.5) in this case can be rewritten as $\pi(\mathbf{e}) = \frac{1}{Z_\beta} exp\left(\beta \sum_{i=0}^{N-1} \mathbf{e}(i)\mathbf{e}(i+1)\right)$, where we always mean vertex number modulo N. The following is a classical *Gibbs sampler* for this model, a Markov chain with stationary distribution (3.5):

- Given a configuration $\mathbf{e}$ at step n, i.e., $X_n = \mathbf{e}$, choose a vertex $v \in V$ with probability $1/N$.

- Take U_{n+1}, a random variable with the uniform distribution $U(0,1)$, independent of $U_i, i \leqslant n$. Update the spin at vertex v in the following way

$$X_{n+1}(i) = \begin{cases} +1 & \text{if } U_{n+1} < \dfrac{e^{2\beta(k_+(v,\mathbf{e}) - k_-(v,\mathbf{e}))}}{e^{2\beta(k_+(v,\mathbf{e}) - k_-(v,\mathbf{e}))} + 1}, \\ -1 & \text{otherwise,} \end{cases}$$

where $k_+(v, \mathbf{e})$ is the number of neighbours of vertex v, in configuration $\mathbf{e}$, with spin values +1, and $k_-(v, \mathbf{e})$ is the number of neighbours of vertex v, in configuration $\mathbf{e}$, with spin values -1.

The chain $\mathbf{X}$ constructed in this way is reversible. Moreover, $\mathbf{X}$ can be viewed as a random walk on $N-$dimensional cube, where the probability of changing coordinate (corresponding to some vertex v) depends on the values of the neighbouring (w.r.t. underlying graph G) coordinates.

It turns out that if we consider the coordinate-wise ordering, i.e., $\mathbf{e} \preceq \mathbf{e}'$ if $\mathbf{e}(v) \leqslant \mathbf{e}'(v)$ for every vertex $v \in V$. Let $M := 2^{|V|} = 2^N$. Denote by $\mathbf{e}_1$ the state with all spins equal to -1 (minimal state), and by $\mathbf{e}_M$ the state with all spins equal to $+1$ (maximal state). We identify $\mathbb{E} = \{-1, 1\}^V$ with the enumerated set $\{\mathbf{e}_1, \dots, \mathbf{e}_M\}$, where the enumeration is consistent with $\preceq$. For specific cases $N =$

2 and $N = 3$ we can directly calculate the dual chain in matrix form from equation (2.1) with link given in Theorem 2.1, namely

$$\mathbf{P}^* = \Lambda \mathbf{P} \Lambda^{-1}.$$

If the resulting $\mathbf{P}^*$ is not a stochastic matrix, then it means that X is not $^{\downarrow}$-Möbius monotone (which will not be the case).

3.3.1. *2 vertices case.* Let us order the states in the following way: $\mathbf{e}_1 = (-1,-1)$, $\mathbf{e}_2 = (+1,-1)$, $\mathbf{e}_3 = (-1,+1)$, $\mathbf{e}_4 = (+1,+1)$. Using this enumeration matrices $\mathbf{P}$ and $\mathbf{C}$ are as follows:

$$\mathbf{P} = \begin{pmatrix} \frac{1}{e^{-2\beta}+1} & \frac{1}{2}\frac{e^{-2\beta}}{e^{-2\beta}+1} & \frac{1}{2}\frac{e^{-2\beta}}{e^{-2\beta}+1} & 0 \\ \frac{1}{2}\frac{1}{e^{-2\beta}+1} & \frac{1}{e^{2\beta}+1} & 0 & \frac{1}{2}\frac{e^{2\beta}}{e^{2\beta}+1} \\ \frac{1}{2}\frac{1}{e^{-2\beta}+1} & 0 & \frac{1}{e^{2\beta}+1} & \frac{1}{2}\frac{e^{2\beta}}{e^{2\beta}+1} \\ 0 & \frac{1}{2}\frac{1}{e^{2\beta}+1} & \frac{1}{2}\frac{1}{e^{2\beta}+1} & \frac{e^{2\beta}}{e^{2\beta}+1} \end{pmatrix}, \mathbf{C} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

From Theorem 2.1 we obtain the following dual chain:

$$\mathbf{P}^* = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2}\frac{e^{2\beta}-1}{e^{2\beta}+1} & \frac{1}{e^{2\beta}+1} \\ 0 & \frac{1}{2}\frac{e^{2\beta}-1}{e^{2\beta}+1} & \frac{1}{2} & \frac{1}{e^{2\beta}+1} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

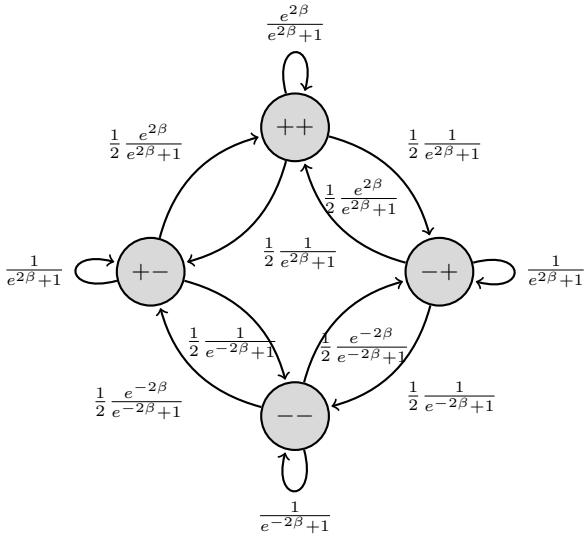
The chain started at $(-1,-1)$ goes with equal probability in first step to $(+1,-1)$ or $(-1,+1)$. Then it either is absorbed in next step with probability $\frac{1}{e^{2\beta}+1}$ or not, with remnant probability $\frac{1}{2} + \frac{1}{2}\frac{e^{2\beta}-1}{e^{2\beta}+1} = \frac{e^{2\beta}}{e^{2\beta}+1}$. Time to absorption in this case is of the following form: $T^* = X_0 + X_1, \quad X_0 \sim Geo(1)$ (i.e., $X_0 \equiv 1$), $X_1 \sim Geo\left(\frac{1}{e^{2\beta}+1}\right)$, i.e.,

$$P(T^* = k) = \left(\frac{e^{2\beta}}{e^{2\beta}+1}\right)^{k-2} \left(\frac{1}{e^{2\beta}+1}\right), \quad k = 2, 3, \ldots$$

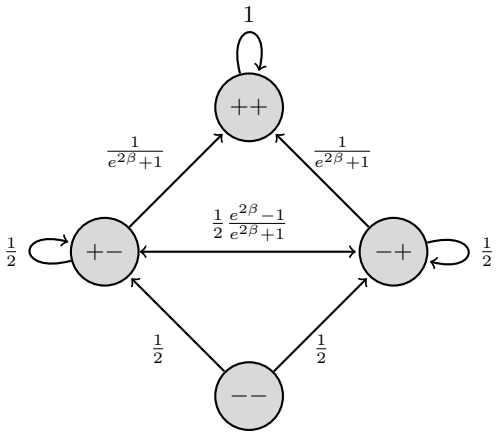
In particular, $ET^* = 2 + e^{2\beta}$ and after $k = 1 + \frac{1}{2\beta - \ln(2\beta+1)} \ln(\varepsilon)$ steps we have

$$s(\delta_{\mathbf{e}_1}\mathbf{P}^k, \pi) = P(T^* > k) = \varepsilon.$$

We can depict the chain and its dual on a graph:



original chain $\mathbf{P}$



dual chain $\mathbf{P}^*$

3.3.2. *3 vertices case.* Consider a simple case of 3 vertices all being each other's neighbour, say labeled v_1, v_2, v_3. Let us enumerate the states in the following way: $\mathbf{e}_1 = (-1,-1,-1)$, $\mathbf{e}_2 = (+1,-1,-1)$, $\mathbf{e}_3 = (-1,+1,-1)$, $\mathbf{e}_4 = (-1,-1,+1)$, $\mathbf{e}_5 = (+1,+1,-1)$, $\mathbf{e}_6 = (+1,-1,+1)$, $\mathbf{e}_7 = (-1,+1,+1)$, $\mathbf{e}_8 = (+1,+1,+1)$. To shorten notation, denote: $p = \frac{e^{4\beta}}{1+e^{4\beta}}$ and $q = \frac{e^{-4\beta}}{1+e^{-4\beta}}$. The tran-

sition matrix is following:

$$\mathbf{P} = \begin{pmatrix}
1-q & \frac{1}{3}q & \frac{1}{3}q & \frac{1}{3}q & 0 & 0 & 0 & 0 \\
\frac{1}{3}-\frac{1}{3}q & \frac{1}{3}+\frac{1}{3}q & 0 & 0 & \frac{1}{6} & \frac{1}{6} & 0 & 0 \\
\frac{1}{3}-\frac{1}{3}q & 0 & \frac{1}{3}+\frac{1}{3}q & 0 & \frac{1}{6} & 0 & \frac{1}{6} & 0 \\
\frac{1}{3}-\frac{1}{3}q & 0 & 0 & \frac{1}{3}+\frac{1}{3}q & 0 & \frac{1}{6} & \frac{1}{6} & 0 \\
0 & \frac{1}{6} & \frac{1}{6} & 0 & \frac{2}{3}-\frac{1}{3}p & 0 & 0 & \frac{1}{3}p \\
0 & \frac{1}{6} & 0 & \frac{1}{6} & 0 & \frac{2}{3}-\frac{1}{3}p & 0 & \frac{1}{3}p \\
0 & 0 & \frac{1}{6} & \frac{1}{6} & 0 & 0 & \frac{2}{3}-\frac{1}{3}p & \frac{1}{3}p \\
0 & 0 & 0 & 0 & \frac{1}{3}-\frac{1}{3}p & \frac{1}{3}-\frac{1}{3}p & \frac{1}{3}-\frac{1}{3}p & p
\end{pmatrix}$$

The stationary distribution in this case is following:

$$\begin{aligned}
\pi((-1,-1,-1)) &= \pi((+1,+1,+1)) = \tfrac{1}{Z} \\
\pi(\mathbf{e}) &= \tfrac{e^{-4\beta}}{Z} \text{ for } \mathbf{e} \notin \{(-1,-1,-1),(+1,+1,+1)\},
\end{aligned}$$

where Z is a normalizing constant equal to $2 + 6e^{-4\beta}$. In this case, function H is as follows:

$$\begin{aligned}
&H((+1,-1,-1)) = H((-1,+1,-1)) = H((-1,-1,+1)) = \tfrac{1+e^{-4\beta}}{2(1+3e^{-4\beta})} \\
&H((+1,+1,-1)) = H((+1,-1,+1)) = H((-1,+1,+1)) = \tfrac{1}{2} \\
&H((-1,-1,-1)) = \tfrac{1}{2(1+3e^{-4\beta})}, H((+1,+1,+1)) = 1
\end{aligned}$$

We obtain the following dual chain:

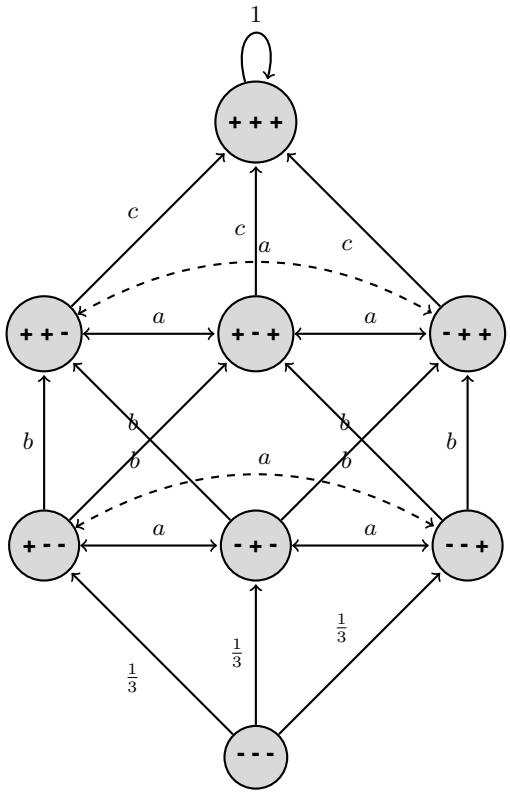
$$\mathbf{P}^* = \begin{pmatrix}
0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 \\
0 & \frac{1}{3} & \frac{1}{6}\frac{1-e^{-4\beta}}{1+e^{-4\beta}} & \frac{1}{6}\frac{1-e^{-4\beta}}{1+e^{-4\beta}} & \frac{1}{6}\frac{1+3e^{-4\beta}}{1+e^{-4\beta}} & \frac{1}{6}\frac{1+3e^{-4\beta}}{1+e^{-4\beta}} & 0 & 0 \\
0 & \frac{1}{6}\frac{1-e^{-4\beta}}{1+e^{-4\beta}} & \frac{1}{3} & \frac{1}{6}\frac{1-e^{-4\beta}}{1+e^{-4\beta}} & \frac{1}{6}\frac{1+3e^{-4\beta}}{1+e^{-4\beta}} & 0 & \frac{1}{6}\frac{1+3e^{-4\beta}}{1+e^{-4\beta}} & 0 \\
0 & \frac{1}{6}\frac{1-e^{-4\beta}}{1+e^{-4\beta}} & \frac{1}{6}\frac{1-e^{-4\beta}}{1+e^{-4\beta}} & \frac{1}{3} & 0 & \frac{1}{6}\frac{1+3e^{-4\beta}}{1+e^{-4\beta}} & \frac{1}{6}\frac{1+3e^{-4\beta}}{1+e^{-4\beta}} & 0 \\
0 & 0 & 0 & 0 & \frac{2}{3} & \frac{1}{6}\frac{1-e^{-4\beta}}{1+e^{-4\beta}} & \frac{1}{6}\frac{1-e^{-4\beta}}{1+e^{-4\beta}} & \frac{2}{3(1+e^{4\beta})} \\
0 & 0 & 0 & 0 & \frac{1}{6}\frac{1-e^{-4\beta}}{1+e^{-4\beta}} & \frac{2}{3} & \frac{1}{6}\frac{1-e^{-4\beta}}{1+e^{-4\beta}} & \frac{2}{3(1+e^{4\beta})} \\
0 & 0 & 0 & 0 & \frac{1}{6}\frac{1-e^{-4\beta}}{1+e^{-4\beta}} & \frac{1}{6}\frac{1-e^{-4\beta}}{1+e^{-4\beta}} & \frac{2}{3} & \frac{2}{3(1+e^{4\beta})} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}$$

Note that, when the chain is on level 1, i.e. in states $(+1,-1,-1),(-1,+1,-1)$ or $(-1,-1,+1)$ then it either gets to some state on level 2 with probability $\frac{2}{6}\frac{1+3e^{-4\beta}}{1+e^{-4\beta}}$ or stays somewhere on level 1 with remnant probability. Similarly, when dual is on level 2 it can get to absorbing state $(+1,+1,+1)$ wit probability $\frac{2}{3(1+e^{4\beta})}$ or stays at the level with remnant probability. Thus, time to absorption is of form:

$$T^* = X_0 + X_1 + X_2, \text{ where}$$

$$X_0 \sim Geo(1), X_1 \sim Geo\left(\frac{1+3e^{-4\beta}}{3(1+e^{-4\beta})}\right), X_2 \sim Geo\left(\frac{2}{3(1+e^{4\beta})}\right).$$

The dual for this case can be depicted as follows:



where $a = \frac{1}{6}(1-e^{-4\beta})/(1+e^{-4\beta}), b = \frac{1}{6}(1+3e^{-4\beta})/(1+e^{-4\beta})$ and $c = 2/(3+3e^{4\beta})$. Self-loops (strictly positive at each vertex except (-1,-1,-1)) are not depicted.

3.3.3. *General N vertices case.* Here we will present conjecture for case of any number of vertices N. First, we conjecture that the chain for this Ising model is always $^{\downarrow}$-Möbius monotone for arbitrary graph G with respect to coordinate-wise ordering. Second, for G being a circle (what is under consideration throughout a section), the structure of a dual is "nice" in a sense we are about to describe. The details for this case will appear in a subsequent paper.

Recall, we consider $G = (V, E)$, where $V = \{0, \ldots, N-1\}$ and $E = \{(i, (i+1) \mod N) : i = 0, \ldots, N-1\}$. We identify a vertex v with a number v and keep general-graph notation. For example $(v, w) \in E$ actually means, that $w = v \pm 1$. For notational convenience, let $\mathbf{e}_{v \leftrightarrow w}$ denotes a configuration $\mathbf{e}$ with swapped spins at v and w. Recall also, that within coordinate-wise ordering, the state with all spins equal to -1 (denoted by $\mathbf{e}_1$) is a minimal state, and state with all spins equal to +1 (denoted by $\mathbf{e}_M$, where $M = 2^N$) is a maximal state.

CONJECTURE 1. Consider the Gibbs sampler $\mathbf{X}$ for the Ising model on arbitrary graph $G = (V, E)$. Assume that $\mathbf{X}$ starts with the configuration $\mathbf{e}_1$. Then, there exists sharp SSD chain $\mathbf{X}^* = (X_n^*)_{n \geqslant 0}$ on the state space $\mathbb{E}^* = \mathbb{E}$, with the state $\mathbf{e}_M$ being the absorbing one, starting with probability 1 at $\mathbf{e}_1$, and having transition probabilities for $\mathbf{e}, \mathbf{e}' \in \{\mathbf{e}_1, \ldots, \mathbf{e}_M\}$

(3.6)

$$\mathbf{P}^*(\mathbf{e}, \mathbf{e}') =$$

$$\begin{cases} 0 & \text{if} \quad \mathbf{e} \succ \mathbf{e}' \\ \frac{1}{N} S(\mathbf{e}) & \text{if} \quad \mathbf{e} = \mathbf{e}' \\ \frac{H(\mathbf{e}')}{H(\mathbf{e})} \frac{1}{N} \left(1 - \frac{e^{2\beta(k_+(v,\mathbf{e}) - k_-(v,\mathbf{e}))}}{e^{2\beta(k_+(v,\mathbf{e}) - k_-(v,\mathbf{e}))} + 1}\right) & \text{if} \quad \mathbf{e}' = \mathbf{e} + s_v,\ \mathbf{e}(v) = -1 \\ \frac{H(\mathbf{e}_{v \leftrightarrow w})}{H(\mathbf{e})} \frac{1}{N} \left(\frac{e^{2\beta(k_+(v,\mathbf{e}) - k_-(v,\mathbf{e}) + 2)}}{e^{2\beta(k_+(v,\mathbf{e}) - k_-(v,\mathbf{e}) + 2)} + 1} - \frac{e^{2\beta(k_+(v,\mathbf{e}) - k_-(v,\mathbf{e}))}}{e^{2\beta(k_+(v,\mathbf{e}) - k_-(v,\mathbf{e}))} + 1}\right) & \text{if} \quad \mathbf{e}' = \mathbf{e}_{v \leftrightarrow w}, (v, w) \in E \\ & \qquad\quad \mathbf{e}(v) = -1, \cdot \mathbf{e}(w) = -1 \end{cases}$$

where $s_v = (0, \dots, 0, 2, 0, \dots, 0)$ (2 on the coordinate corresponding to vertex v), $S(\mathbf{e}) = \sum_{v \in V} \mathbf{1}\{\mathbf{e}(v) = 1\}$ (called *a level*), and $H(\mathbf{e}) = \sum_{\mathbf{e}' \preceq \mathbf{e}} \pi(\mathbf{e})$.

The conjectured dual chain, being at some state $\mathbf{e}_1$ (thus on level $S(\mathbf{e}_1)$)

- can stay at the same state
- cannot go to any state with lower level
- can go to higher levels only to states of form $\mathbf{e} + s_j$ (i.e., only one lever higher, to a comparable state)
- can move "across" the level only by swapping spins at some vertices v and w which are neigbours, i.e. $(v, w) \in E$

The special cases $N = 2$ and $N = 3$ were particularly "nice". For $N = 2$ the transitions from $\mathbf{e}$ to $\mathbf{e}' = \mathbf{e}_{v \leftrightarrow w}$ are all zero. For $N = 3$ the transitions to another state on the same level or to higher level, do not depend on particular state, but only on the current level. This is mainly due to fact that "*neigbours of* v" actually means "*all other vertices of* v". That is implied, that time to absorption in these case was distributed as sum of geometric random variables. For $N \geqslant 4$, the probabilities depend on a particular state. However, the dual has a nice block-matrix form

$$\mathbf{P}^* = \begin{pmatrix} \mathbf{P}^*_{0,0} & \mathbf{P}^*_{0,1} & 0 & \dots & & & & 0 \\ 0 & \mathbf{P}^*_{1,1} & \mathbf{P}^*_{1,2} & 0 & \dots & & & 0 \\ 0 & & \ddots & & 0 & \dots & & 0 \\ 0 & \dots & & \mathbf{P}^*_{i,i} & \mathbf{P}^*_{i,i+1} & 0 & \dots & 0 \\ 0 & & & & \ddots & 0 & \dots & 0 \\ 0 & \dots & & & & \mathbf{P}^*_{N-1,N-1} & \mathbf{P}^*_{N-1,N} \\ 0 & & & & \dots & 0 & 1 \end{pmatrix},$$

where $\mathbf{P}^*(i, i)$ is a square matrix corresponding to all states with level i of size $\binom{N}{i} \times \binom{N}{i}$, and $\mathbf{P}^*(i, i+1)$ is a matrix of size $\binom{N}{i} \times \binom{N}{i+1}$, $i = 0, \dots, N-1$.

REMARK 3.1. *The time to absorption for such block-matrices can be further studied again via some dualities. Roughly speaking, for given absorbing chain* $\mathbf{P}^*$ *there is a method for finding another "dual" absorbing chain* $\hat{\mathbf{P}}$ *which has the same time to absorption. The method is due to Fill and Lyzinski [14].*

3.4. Further examples and developements.

3.4.1. *Simple symmetric random walk on a circle.* Let $\mathbb{Z}_d$ be the set of integers modulo d, regarded as d labelled $1, 2, \ldots, d$ points arranged anti-clockwise around a circle with 1 at the bottom, say. Suppose we start at 1, that is $\nu = \delta_1$ and with probability $1/3$ random walk $\mathbf{X}$ moves one step in either direction along the circle or remains. The stationary distribution for this Markov chain is the uniform distribution π_u on $\mathbb{E} = \mathbb{Z}_d$. Diaconis and Fill [5] showed for $d = 2^a$, a SST T, such that

$$d_{TV}(\nu \mathbf{P}^n, \pi_u) \leqslant P(T > n) \leqslant \frac{3}{16} d^2/n,$$

and pointed out some extensions of this model. The point of this example was that a SST was found by identifying sets of states increasing in size and times at which the process is uniform on each set. This transformed the original problem of studying convergence to stationarity into a different problem of analysing first passage times. It might be interesting to know that there is an alternative way of finding a SSD chain for this walk using Möbius monotonicity. It is possible if we order linearly the state space by a zig-zag ordering on the circle. For example, for $d = 2^3$, the ordering is $1 < 2 < 8 < 3 < 7 < 4 < 6 < 5$, with 5 being the maximal state, or, for $d = 2^4$, the ordering is $1 < 2 < 16 < 3 < 15 < 4 < 14 < 5 < 13 < 6 < 12 < 7 < 11 < 8 < 10 < 9$, with 9 being the maximal state. It turns out that with this ordering $\mathbf{X}$ is Möbius monotone. The corresponding dual is a Markov chain with 1 step transitions only on one side of the circle, that is for example in the case $d = 16$ moving as birth- death chain on $1, 16, 15, 14, 13, 12, 11, 10, 9$, with the maximal state 9 being the absorbing state. That means that the original problem of studying convergence to stationarity for a symmetric random walk on a circle can be transformed into a different problem of analysing passage time for a birth and death process on linearly ordered segment from the minimal to maximal state, which is more standard one. Möbius monotonicity here is a delicate property,

because changing the probability of remaining in a state from $1/3$ to a smaller value destroys this monotonicity, however changing the probability to remain from $1/3$ to a bigger value does not influence the property of Möbius monotonicity. Also making the walk not symmetric destroys this kind of monotonicity.

4. PROOFS

4.1. Proof of Theorem 3.1. We start with a detailed expression for the transition probabilities of $\mathbf{X}$

$$\mathbf{P}((x,y),(x',y')) = \begin{cases} \lambda_1 & \text{if } x' = x+1 \leqslant N, y' = y \\ \lambda_2 & \text{if } x' = x, y' = y+1 \leqslant N \\ \mu_1 & \text{if } x' = x-1 \geqslant 0, y' = y \\ \mu_2 & \text{if } y' = y-1 \geqslant 0, x' = x \\ 1-(\lambda_1+\lambda_2+\mu_1+\mu_2) & \text{if } x' = x > 0, y' = y > 0 \\ 1-(\lambda_1+\lambda_2+\mu_1) & \text{if } x' = x > 0, y' = y = 0 \\ 1-(\lambda_1+\lambda_2+\mu_2) & \text{if } x' = x = 0, y' = y > 0 \\ 1-(\mu_1+\mu_2) & \text{if } x' = x = y = y' = N \\ 1-(\mu_1+\mu_2+\lambda_1) & \text{if } x' = x > 0, y' = y = N \\ 1-(\mu_1+\mu_2+\lambda_2) & \text{if } x' = x = N, y' = y > 0 \\ 1-(\lambda_1+\lambda_2) & \text{if } x' = x = y = y' = 0 \\ 1-(\lambda_1+\mu_2) & \text{if } x' = x = N, y = y' = 0 \\ 1-(\lambda_2+\mu_1) & \text{if } x' = x = 0, y = y' = N \end{cases}$$

In a standard way we can check that $\mathbf{X}$ is reversible and the stationary distribution is given by

$$\pi((x,y)) = C^{-1}\rho_1^x\rho_2^y$$

where C is the normalizing constant, and $\rho_i = \lambda_i/\mu_i,\ i = 1, 2$. For the coordinate-wise ordering

$$(x, y) \preceq (x', y') \iff x \leqslant x' \text{ and } y \leqslant y',$$

with the minimal state $\mathbf{e}_1 = (0, 0)$, and the maximal state $\mathbf{e}_M = (N, N)$, $(M = (N + 1)^2)$ directly from Proposition 5 in [23], we find the corresponding Möbius function:

$$\mu((x, y), (x', y')) = \begin{cases} 1 & \text{if} \quad x' = x,\ y' = y \\ -1 & \text{if} \quad x' = x + 1,\ y' = y \\ -1 & \text{if} \quad x' = x,\ y' = y + 1 \\ 1 & \text{if} \quad x' = x + 1,\ y' = y + 1 \\ 0 & \text{otherwise} \end{cases}$$

For

$$H(x, y) = C^{-1} \sum_{x' \leqslant x} \rho_1^{x'} \sum_{y' \leqslant y} \rho_2^{y'} = C^{-1}(1 - \rho_1)^{-1}(1 - \rho_2)^{-1} \left(1 - \rho_1^{x+1}\right) (1 - \rho_2^{y+1}),$$

we shall compute

(4.1)

$$\mathbf{P}^*((x, y), (x_2, y_2)) = \frac{H(x_2, y_2)}{H(x, y)} \sum_{(x', y') \succeq (x_2, y_2)} \mu((x_2, y_2), (x', y')) \overleftarrow{\mathbf{P}}((x', y'), \{(x, y)\}^{\downarrow}).$$

Set

$$S := \sum_{(x', y') \succeq (x_2, y_2)} \mu((x_2, y_2), (x', y')) \overleftarrow{\mathbf{P}}((x', y'), \{(x, y)\}^{\downarrow}).$$

Note that in order to prove that $\overleftarrow{\mathbf{X}}$ is $^{\downarrow}$-Möbius monotone it is enough to show that $S \geqslant 0$. Since $\mathbf{X}$ is reversible, we take $\mathbf{P}$ instead of $\overleftarrow{\mathbf{P}}$ in the above formula. We shall consider all possible transitions, case by case.

- (inside lattice, up x direction)

$x_2 = x + 1,\ y_2 = y$

$$S = \sum_{(x',y') \succeq (x+1,y)} \mu((x+1,y),(x',y'))\mathbf{P}((x',y'),\{(x,y)\}^{\downarrow})$$

where μ will be non-zero only in the following cases

$$\mu((x+1,y),(x+1,y)) = 1, \quad \mu((x+1,y),(x+1,y+1)) = -1,$$

$$\mu((x+1,y),(x+2,y)) = -1, \quad \mu((x+1,y),(x+2,y+1)) = 1.$$

Combining these cases with the values of $\mathbf{P}((x',y'),\{(x,y)\}^{\downarrow})$ we get

$$S = \mu((x+1,y),(x+1,y))\mathbf{P}((x+1,y),\{(x,y)\}^{\downarrow}) - 1\cdot 0 - 1\cdot 0 + 1\cdot 0 = \mu_1,$$

- (inside lattice, up y direction)

$x_2 = x,\ y_2 = y+1$

in a similar way as ubove, we get

$$S = \mu((x,y+1),(x,y+1))\mathbf{P}((x,y+1),\{(x,y)\}^{\downarrow}) - 1\cdot 0 - 1\cdot 0 + 1\cdot 0 = \mu_2,$$

- (inside lattice, down x direction)

$x_2 = x-1 \geqslant 0,\ y_2 = y$

using the formula for S we have

$$\begin{aligned} S &= \mu((x-1,y),(x-1,y))\mathbf{P}((x-1,y),\{(x,y)\}^{\downarrow}) \\ &\quad + \mu((x-1,y),(x,y))\mathbf{P}((x,y),\{(x,y)\}^{\downarrow}) \\ &\quad + \mu((x-1,y),(x-1,y+1))\mathbf{P}((x-1,y+1),\{(x,y)\}^{\downarrow}) \\ &\quad + \mu((x-1,y),(x,y+1))\mathbf{P}((x,y+1),\{(x,y)\}^{\downarrow}) \\ &= 1\cdot(1-\lambda_2) - 1\cdot(1-\lambda_2-\lambda_1) - 1\cdot\mu_2 + 1\cdot\mu_2 = \lambda_1, \end{aligned}$$

- (inside lattice, down y direction)

$x_2 = x,\ y_2 = y - 1 \geqslant 0$

$$\begin{aligned} S &= \mu((x,y-1),(x,y-1))\mathbf{P}((x,y-1),\{(x,y)\}^{\downarrow}) \\ &\quad + \mu((x,y-1),(x,y))\mathbf{P}((x,y),\{(x,y)\}^{\downarrow}) \\ &\quad + \mu((x,y-1),(x+1,y-1))\mathbf{P}((x+1,y-1),\{(x,y)\}^{\downarrow}) \\ &\quad + \mu((x,y-1),(x+1,y))\mathbf{P}((x+1,y),\{(x,y)\}^{\downarrow}) \\ &= 1\cdot(1-\lambda_1) - 1\cdot(1-\lambda_2-\lambda_1) - 1\cdot\mu_1 + 1\cdot\mu_1 = \lambda_2, \end{aligned}$$

- (inside lattice, down on both axes)

$x_2 = x - 1 \geqslant 0,\ y_2 = y - 1 \geqslant 0$

$$\begin{aligned} S &= \mu((x-1,y-1),(x-1,y-1))\mathbf{P}((x-1,y-1),\{(x,y)\}^{\downarrow}) \\ &\quad + \mu((x-1,y-1),(x-1,y))\mathbf{P}((x-1,y),\{(x,y)\}^{\downarrow}) \\ &\quad + \mu((x-1,y-1),(x,y-1))\mathbf{P}((x,y-1),\{(x,y)\}^{\downarrow}) \\ &\quad + \mu((x-1,y-1),(x,y))\mathbf{P}((x,y),\{(x,y)\}^{\downarrow}) \\ &= 1\cdot 1 - 1\cdot(1-\lambda_2) - (1-\lambda_1) + 1 - (\lambda_1+\lambda_2) = 0. \end{aligned}$$

In a similar way it is possible to check that inside the lattice the only one remaining movement with positive probability is the feedback movement

- (feedback inside lattice) $x_2 = x > 0,\ y_2 = y > 0$

$$\mathbf{P}^*((x,y),(x,y)) = 1 - \lambda_1 - \lambda_2 - \mu_1 - \mu_2 = \mathbf{P}((x,y),(x,y)),$$

- (upper border, up x direction)

$x_2 = x + 1 \leqslant N,\ y_2 = y = N$

$$S = \mu((x+1,N),(x+1,N))\mathbf{P}((x+1,N),\{(x,N)\}^{\downarrow}) = \mu_1,$$

- (upper border, down y direction)

$x_2 = x < N, y = N, y_2 = N - 1$

$$\begin{aligned} S &= \mu((x,N-1),(x,N-1))\mathbf{P}((x,N-1),\{(x,N)\}^{\downarrow}) \\ &\quad + \mu((x,N-1),(x,N))\mathbf{P}((x,y),\{(x,N)\}^{\downarrow}) \\ &\quad + \mu((x,N-1),(x+1,N-1))\mathbf{P}((x+1,N-1),\{(x,N)\}^{\downarrow}) \\ &\quad + \mu((x,N-1),(x+1,N))\mathbf{P}((x+1,N),\{(x,N)\}^{\downarrow}) \\ &= 1\cdot(1-\lambda_1) - 1\cdot(1-\lambda_1) - 1\cdot\mu_1 + 1\cdot\mu_1 = 0, \end{aligned}$$

- (upper-right corner, down y direction)

$x_2 = x = N, y = N, y_2 = N - 1$

$$\begin{aligned} S &= \mu((N,N-1),(N,N-1))\mathbf{P}((N,N-1),\{(N,N)\}^{\downarrow}) \\ &\quad + \mu((N,N-1),(N,N))\mathbf{P}((N,N),\{(N,N)\}^{\downarrow}) = 1 - 1 = 0 \end{aligned}$$

- (upper border, down x direction)

$N - 1 > x_2 = x - 1 \geqslant 0, y_2 = y = N$

$$\begin{aligned} S &= \mu((x-1,N),(x-1,N))\mathbf{P}((x-1,N),\{(x,N)\}^{\downarrow}) \\ &\quad + \mu((x-1,N),(x,N))\mathbf{P}((x,N),\{(x,N)\}^{\downarrow}) = 1\cdot 1 - 1\cdot(1-\lambda_1) = \lambda_1, \end{aligned}$$

- (lower border, up x direction)

$N - 1 > x_2 = x + 1 \leqslant N, y_2 = y = 0$

$$S = \mu((x+1,0),(x+1,0)\mathbf{P}((x+1,0),\{(x,0)\}^{\downarrow}) = \mu_1,$$

- (lower border, down x direction)

$x_2 = x - 1 \geqslant 0, y_2 = y = 0$

$$\begin{aligned} S &= \mu((x-1,0),(x-1,0))\mathbf{P}((x-1,0),\{(x,0)\}^{\downarrow}) \\ &\quad + \mu((x-1,0),(x,0))\mathbf{P}((x,0),\{(x,0)\}^{\downarrow}) \\ &\quad + \mu((x-1,0),(x-1,1))\mathbf{P}((x-1,1),\{(x,0)\}^{\downarrow}) \\ &\quad + \mu((x-1,0),(x,1))\mathbf{P}((x,1),\{(x,0)\}^{\downarrow}) \\ &= 1\cdot(1-\lambda_2) - 1\cdot(1-\lambda_1-\lambda_2) - 1\cdot\mu_2 + \mu_2 = \lambda_1, \end{aligned}$$

- (lower border, up y direction)

$x_2 = x \geqslant 0, y_2 = 1, y = 0$

$$S = \mu((x,1),(x,1))\mathbf{P}((x,1),\{(x,0)\}^{\downarrow}) = \mu_2.$$

In a similar way we get

- (right border, up y direction)

$x_2 = x = N, y_2 = y+1 \leqslant N, \quad S = \mu_2,$

- (right border, down y direction)

$x_2 = x = N, N-1 > y_2 = y-1 \geqslant 0, \quad S = \lambda_2,$

- (right border, down x direction)

$x_2 = N-1, x = N, y_2 = y, \quad S = 0,$

- (left border, up y direction)

$x_2 = x = 0, y_2 = y+1 \leqslant N, \quad S = \mu_2,$

- (left border, up x direction)

$x = 0, x_2 = 1, y_2 = y, \quad S = \mu_1,$

- (left border, down y direction)

$x_2 = x = 0, N-1 > y_2 = y-1 \geqslant 0, \quad S = \lambda_2,$

- (absorbing state)

$x_2 = x = N, y_2 = y = N, \quad S = 1.$

- (feedback movements)

for all $(x,y) \in \{0,\dots,N-1\}^2$, $S = 1-(\lambda_1+\lambda_2+\mu_1+\mu_2)$,

for $x = N$, and $y \in \{0,\dots,N-1\}$, $S = 1-(\lambda_2+\mu_2)$,

for $y = N$, and $x \in \{0,\dots,N-1\}$, $S = 1-(\lambda_1+\mu_1)$.

Now using (4.1), and using values of $H(x,y)$ we obtain $\mathbf{P}^*$ given in (3.3).

4.2. Proof of Theorem 3.2. Consider the coordinate-wise ordering

$$\mathbf{e} = (\mathbf{e}(1),\dots,\mathbf{e}(n)) \preceq (\mathbf{e}'(1),\dots,\mathbf{e}'(n)) = \mathbf{e}' \text{ iff } \mathbf{e}(i) \leqslant \mathbf{e}'(i), i = 1,\dots,n.$$

Again, for this ordering with minimal element $\mathbf{e}_1 = (0,\dots,0)$ and maximal element $\mathbf{e}_M = (k,\dots,k)$ (with $M = (k+1)^n$) , directly from Proposition 5 in Rota [23], we find the corresponding Möbius function

$$\mu((\mathbf{e}(1),\dots,\mathbf{e}(n)),(\mathbf{e}(1)+d_1,\dots,\mathbf{e}(n)+d_n)) = \begin{cases} (-1)^{\sum_{i=1}^n d_i} & d_i \in \{0,1\},\ \mathbf{e}(i)+d_i \leqslant k, \\ & \qquad i = 1,\dots,n \\ 0 & \text{otherwise.} \end{cases}$$

For $H(\mathbf{e}) = \sum_{\mathbf{e}'\preceq\mathbf{e}} \pi(\mathbf{e}') = |\{\mathbf{e}' : \mathbf{e}' \leqslant \mathbf{e}\}| \cdot 1/(k+1)^n$, we shall compute directly transitions of the dual chain (2.2) from Theorem 2.1. Note, that in order to prove that $\overleftarrow{\mathbf{X}}$ is $^{\downarrow}$-Möbius monotone, it is enough to show that all summands in (2.2) are non-negative. We take $\mathbf{P}$ instead of $\overleftarrow{\mathbf{P}}$ since this chain is reversible.
For convenience, we shall consider states of the following form

$$\mathbf{e}_A^{(k)} = (\mathbf{e}_A^{(k)}(1),\dots,\mathbf{e}_A^{(k)}(n)),\ \ A \subseteq \{1,\dots,n\},$$

with $\mathbf{e}_A^{(k)}(i) = k$ if $i \in A$ and 0 otherwise. Note, that there are $(k+1)^{|A|}$ states smaller or equal (w.r.t. $\preceq$) to $\mathbf{e}_A^{(k)}$, and we have

$$\frac{H(\mathbf{e}_{A\cup\{j\}}^{(k)})}{H(\mathbf{e}_A^{(k)})} = \frac{(k+1)^{|A\cup\{j\}|}}{(k+1)^{|A|}} = k+1 \text{ for } j \notin A. \tag{4.2}$$

Let us calculate transitions of the dual chain from state $\mathbf{e}_A^{(k)}$. We shall use $s_i = (0,\dots,0,1,0,\dots,0)$ with 1 at position i. For the probability of staying at this state we get

$$\mathbf{P}^*(\mathbf{e}_A^{(k)},\mathbf{e}_A^{(k)}) = 1\cdot \sum_{\mathbf{e}\succeq \mathbf{e}_A^{(k)}} \mu(\mathbf{e}_A^{(k)},\mathbf{e})\mathbf{P}(\mathbf{e},\{\mathbf{e}_A^{(k)}\}^{\downarrow})$$

$$= \mu(\mathbf{e}_A^{(k)},\mathbf{e}_A^{(k)})\mathbf{P}(\mathbf{e}_A^{(k)},\{\mathbf{e}_A^{(k)}\}^{\downarrow}) + \sum_{i\in A^c} \mu(\mathbf{e}_A^{(k)},\mathbf{e}_A^{(k)}+s_i)\mathbf{P}(\mathbf{e}_A^{(k)}+s_i,\{\mathbf{e}_A^{(k)}\}^{\downarrow})$$

$$= 1\cdot\left(\frac{1}{2}+\sum_{i\in A} k\cdot\frac{1}{2nk}\right) - \sum_{i\in A^c}\frac{1}{2nk} = \frac{1}{2}+\frac{k|A|}{2nk} - \frac{n-|A|}{2nk} = \frac{n(k-1)+|A|(k+1)}{2nk},$$

since $\mathbf{P}(\mathbf{e}_A^{(k)}+s_i,\{\mathbf{e}_A^{(k)}\}^{\downarrow}) = \mathbf{P}(\mathbf{e}_A^{(k)}+s_i,\mathbf{e}_A^{(k)})$.

Now, for the probability of transition from $\mathbf{e}_A^{(k)}$ to $\mathbf{e}_{A\cup\{j\}}^{(k)}$, $j\notin A$ we obtain

$$\mathbf{P}^*(\mathbf{e}_A^{(k)},\mathbf{e}_{A\cup\{j\}}^{(k)}) = \frac{H(\mathbf{e}_{A\cup\{j\}}^{(k)})}{H(\mathbf{e}_A^{(k)})} \sum_{\mathbf{e}\succeq\mathbf{e}_{A\cup\{j\}}^{(k)}} \mu(\mathbf{e}_{A\cup\{j\}}^{(k)},\mathbf{e})\mathbf{P}(\mathbf{e},\{\mathbf{e}_A^{(k)}\}^{\downarrow}).$$

The only state $\mathbf{e}$ for which $\mathbf{P}(\mathbf{e},\{\mathbf{e}_A^{(k)}\}^{\downarrow})>0$ is $\mathbf{e}=\mathbf{e}_{A\cup\{j\}}^{(k)}$, thus (using (4.2)) we have

$$\mathbf{P}^*(\mathbf{e}_A^{(k)},\mathbf{e}_{A\cup\{j\}}^{(k)}) = (k+1)\mu(\mathbf{e}_{A\cup\{j\}}^{(k)},\mathbf{e}_{A\cup\{j\}}^{(k)})\mathbf{P}(\mathbf{e}_{A\cup\{j\}}^{(k)},\{\mathbf{e}_A^{(k)}\}^{\downarrow}) = \frac{k+1}{2nk}.$$

This completes our argument since all other transitions have probability 0, which is clear from the following summation

$$\mathbf{P}^*(\mathbf{e}_A^{(k)},\mathbf{e}_A^{(k)}) + \sum_{j\in A^c}\mathbf{P}^*(\mathbf{e}_A^{(k)},\mathbf{e}_{A\cup\{j\}}^{(k)}) = \frac{n(k-1)+|A|(k+1)}{2nk} + (n-|A|)\cdot\frac{k+1}{2nk}$$

$$= \frac{n(k-1)+n(k+1)+|A|(k+1)-|A|(k+1)}{2nk} = 1.$$

Note that the dual chain starts at the minimal state which is also of the form $\mathbf{e}_A^{(k)}$, namely with $A=\emptyset$.

ACKNOWLEDGMENTS

Authors thank an anonymous referee for his/her comments that allowed us to improve the paper. All authors kindly acknowledge partial support by the project RARE-318984, a Marie Curie IRSES Fellowship within the 7th European Community Framework Programme.

REFERENCES

[1] Aldous, D.J. Finite-Time Implications of Relaxation Times for Stochastically Monotone Processes. *Probab. Th. Rel. Fields*, **77**, 137-145, (1988)

[2] Aldous, D.J., Diaconis, P. Shuffling cards and stopping times. *American Mathematical Monthly*, **93**, 333–348, (1986)

[3] Aldous, D.J., Diaconis, P. Strong uniform times and finite random walks. *Advances in Applied Mathematics*, **8**, 69–97, (1987)

[4] Diaconis, P. *Group Representations in Probability and Statistics*. IMS, Hayward, CA. (1988)

[5] Diaconis, P., Fill, J.A. Strong stationary times via a new form of duality. *The Annals of Probability*, **18**, 1483-1522, (1990)

[6] Diaconis, P., Fill, J.A. Examples for the theory of strong stationary duality with countable state spaces. *Probability in the Engineering and Informational Sciences*, **4**, 157–180, (1990)

[7] Diaconis, P. and Miclo, L. On times to quasi-stationarity for birth and death processes. *Journal of Theoretical Probability* **22**, 558-586, (2009)

[8] Diaconis, P. and Saloff-Coste, L. Separation cut-offs for birth and death chains. *The Annals of Applied Probability* **16**(4), 2098–2122, (2006)

[9] Ding, J. and Lubetzky, E. and Peres, Y. The mixing time evolution of Glauber dynamics for the mean-field Ising model *Communications in Mathematical Physics*, **289**, 725-764, (2009)

[10] Ding, J. and Peres, Y. Mixing time for the Ising model: A uniform lower bound for all graphs. *Ann. Inst. H. Poincaré Probab. Statist.*, **47**(4), 1020–1028, (2011)

[11] Ding, J. and Peres, Y. Mixing time for the Ising model: A uniform lower bound for all graphs. arXiv:0909.5162v2 [math.PR] (2013)

[12] Fill, J. A. The passage time distribution for a birth-and-death chain:Strong stationary duality gives a first stochastic proof. *Journal of Theoretical Probability* **22**, 543-557, (2009a)

[13] Fill, J. A. On hitting times and fastest strong stationary times for skip-free and more general chains. *Journal of Theoretical Probability* **22**(3), 587-600, (2009)

[14] Fill, J. and Lyzinski, V. Hitting times and interlacing eigenvalues: a stochastic approach using intertwinings *Journal of Theoretical Probability* **27**(3), 954–981, (2014)

[15] Hayes, T. P. and Sinclair, A. A general bound for mixing of single-site dynamics on graphs. *The Annals of Applied Probability* **17**(3), 931–952, (2007)

[16] Huillet, T. and Martinez, S. On Möbius Duality and Coarse-Graining. *Journal of Theoretical Probability*, doi:10.1007/s10959-014-0569-5, (2014)

[17] Levin, D. and Peres, Y. and Wilmer, E. Markov chains and mixing times. *American Mathematical Soc.* (2009)

[18] Levin, D. A. and Luczak, M. and Peres, Y. Glauber dynamics for the mean-field Ising model: cut-off, critical power law, and metastability *Probability Theory and Related Fields* **146**, 223–265, (2010)

[19] Lorek, P. , Szekli, R. Computable bounds on the spectral gap for unreliable Jackson networks. *Advances in Applied Probability*, **47**(2), (2015) arXiv:1101.0332 [math.PR]

[20] Lorek, P., Szekli, R. Strong stationary duality for Möbius monotone Markov chains. *Queueing Systems*, **71**, 79–95, (2012)

[21] Lubetzky, E. and Sly, A. Cutoff for the Ising model on the lattice. *Inventiones mathematicae***191**, 719-755, (2013)

[22] Propp, J. G. and Wilson, D. B. Exact sampling with coupled Markov chains and applications to statistical mechanics *Random Structures & Algorithms*, **9**(1), 223–252, (1996)

[23] Rota, G. C. On the Foundations of Combinatorial Theory I. Theory of Möbius Functions. *Z. Wahrseheinlichkeitstheorie* **2**, 340–368, (1964)

University of Wrocław

E-mail: Pawel.Lorek@math.uni.wroc.pl

University of Wrocław

E-mail: Ryszard.Szekli@math.uni.wroc.pl

Received on 20.11.2014;
revised version on xx.xx.xxxx