

STRONG CONVERGENCE OF INFINITE COLOR BALANCED URNS UNDER UNIFORM ERGODICITY

ANTAR BANDYOPADHYAY, SVANTE JANSON, AND DEBLEENA THACKER

ABSTRACT. We consider the generalization of the Pólya urn scheme with possibly infinite many colors as introduced in [37, 4, 5, 6]. For countable many colors, we prove almost sure convergence of the urn configuration under *uniform ergodicity* assumption on the associated Markov chain. The proof uses a stochastic coupling of the sequence of chosen colors with a *branching Markov chain* on a weighted *random recursive tree* as described in [6, 31, 26]. Using this coupling we estimate the covariance between any two selected colors. In particular, we reprove the limit theorem for the classical urn models with finitely many colors.

1. INTRODUCTION

Pólya urn schemes and its various generalizations have been a key element of study for random processes with reinforcements. Starting from the seminal work of Pólya [35], various types of urn schemes with finitely many colors have been widely studied in literature, see [33] for an extensive survey of the known classical results and some of the modern works can be found in [24, 25, 3, 17, 8, 9, 12, 11, 10].

The Pólya urn models with colors indexed by a general *Polish space* was first introduced by Blackwell and MacQueen [7]. They showed that the so called *Ferguson distribution* [16] on the set of probabilities on a Polish space can be obtained as a limit of Pólya-type urn model. Other seminal work on urn models with possibly infinitely many colors was done by Hoppe [22, 23] in the context of population genetics. He introduced a new type of urn scheme [22], where at each step with some positive probability a new color can be introduced. Several authors studied this slightly different model in the context of *Griffiths-Engen-McCloskey (GEM) Model*, *Poisson-Dirichlet Distribution* and *Ewens Sampling Formula* [13, 20, 21, 38]. It was also observed that the *Hoppe's urn scheme* has deep relation with certain combinatorial stochastic processes, known as, *Chinese Restaurant Process* [1, 34]. It is to be noted here that one of the key ingredient in studying these urn models is the fact that the sequence of colors are *exchangeable*.

A somewhat different generalization for *balanced* urn schemes with infinitely many colors was introduced in [37] and subsequently in the papers [4, 5, 6]. These models are significantly different than the two other models discussed above. Typically for these class of models the observed color sequence need not be exchangeable and hence new method of analysis was introduced in [4, 5, 6]. These works have

Date: April 12, 2019; revised November 24, 2019.

2010 Mathematics Subject Classification. Primary: 60F05, 60F10; Secondary: 60G50.

Key words and phrases. almost sure convergence, branching Markov chain, infinite color urn, random recursive tree, reinforcement processes, uniform ergodicity, urn models.

since then generated a lot of interest and such models are now receiving considerable attention [26, 27, 31]. In this paper, we will consider the infinite color balanced urn model as introduced in [4, 5, 6], where the color set is countably infinite.

1.1. Model. In this work we will consider the same generalization of the Pólya urn scheme with infinite many colors as defined in [4, 5, 6]. However, we will focus on the special case where the set of colors, is *countably infinite*, which will be denoted by S . We follow similar framework and notations as in [4, 5, 6]. For the sake of completeness, we provide here a brief description of the model.

Let R be a $S \times S$ (infinite) matrix with non-negative entries, representing the *replacement scheme*. We will assume that R is *balanced*, that is, each row sum is equal and finite. In that case, it is customary to take R a *stochastic matrix* (see [6] for details).

We will denote by $U_n := (U_{n,v})_{v \in S} \in [0, \infty)^S$, the random configuration of the urn at time $n \geq 0$. We will view it as an infinite vector (with non-negative entries) which is in $\ell_1 \equiv \ell_1(S)$, and thus can also be viewed as a (random) finite measure on S . Intuitively, we will define U_n , such that, if Z_n is the randomly chosen color at the $(n+1)$ -th draw, then the conditional distribution of Z_n given the “*past*”, will satisfy, for all $z \in S$,

$$\mathbb{P}\left(Z_n = z \mid U_n, U_{n-1}, \dots, U_0\right) \propto U_n(z).$$

Formally, starting with a non-random $U_0 \in \ell_1$, we define $(U_n)_{n \geq 0} \subseteq \ell_1$, recursively as follows

$$(1.1) \quad U_{n+1} = U_n + R_{Z_n},$$

where R_z denotes the z -th row of the matrix R , and

$$(1.2) \quad \mathbb{P}\left(Z_n = z \mid U_n, U_{n-1}, \dots, U_0\right) = \frac{U_{n,z}}{n+t},$$

where U_0 is a ℓ_1 -vector with total mass denoted by $0 < t < \infty$, that is, $\sum_{v \in S} U_{0,v} = t \in (0, \infty)$.

Observe that, one can now associate with such an urn model a Markov chain $(X_n)_{n \geq 0}$ on the countable state space S , with transition matrix R and initial distribution U_0/t . Conversely, given any Markov chain $(X_n)_{n \geq 0}$, on the countable state space S , with transition matrix R and a vector $U_0 \in \ell_1$, one can associate a balanced urn model $(U_n)_{n \geq 0}$, satisfying equations (1.1) and (1.2). We will call such a Markov chain $(X_n)_{n \geq 0}$, as a Markov chain associated with the urn model $(U_n)_{n \geq 0}$.

It has been observed in [5, 31] that the asymptotic properties of the urn model so defined are determined by the asymptotic properties of the associated Markov chain. In fact, in [5, 31], the authors have shown that the urn sequence $(U_n)_{n \geq 0}$ has same law as that of a *branching Markov chain* with transition matrix R , initial distribution U_0/t and defined on a *random recursive tree*. In Section 2, we provide the details of this representation.

1.2. Main Result. In this paper, we consider the case when R is irreducible, aperiodic and positive recurrent. From classical theory (see Section XV.7 of [15] for the details), it is well know that, in that case, the chain has unique stationary

distribution, say, π , satisfying the equation

$$(1.3) \quad \pi R = \pi.$$

Moreover, such a chain is *ergodic*, that is, for any $u, v \in S$,

$$\lim_{n \rightarrow \infty} R^n(u, v) = \pi_v,$$

where R^n is the n -step transition matrix, which is nothing but the n -fold composition of R with itself. Note that as S is countable, R^n is just the n -fold multiplication of R .

In this work we will further assume that the chain is *uniformly ergodic*. For the sake of completeness, we provide the definition here. (One often uses a version with summation over v in (1.4); we need only the version below.)

Definition 1.1. *A Markov chain with transition matrix R on a countable state space S is called uniformly ergodic, if there exists positive constants, $0 < \rho < 1$ and $C > 0$, such that, for any time $n \geq 1$ and for any states $u, v \in S$,*

$$(1.4) \quad |R^n(u, v) - \pi_v| \leq C\rho^n.$$

We note here that if S is finite then an irreducible and aperiodic chain is necessarily uniformly ergodic (see Theorem 4.9 of [29]). However, when S is infinite (even countable) there are ergodic chains which are not uniformly ergodic (see e.g. [19]).

Our main result is as follows:

Theorem 1.2. *Consider an urn model $(U_n)_{n \geq 0}$ as defined by the equations (1.1) and (1.2), with colors indexed by a countably infinite set S , a balanced replacement matrix R , and an initial configuration U_0 . We assume that R is a stochastic matrix which is irreducible, aperiodic, positive recurrent with stationary distribution π , and uniformly ergodic. That is, it satisfies (1.4).*

(i) *Then as $n \rightarrow \infty$,*

$$(1.5) \quad \frac{U_n}{n+t} \longrightarrow \pi \text{ a.s.},$$

where the convergence is coordinate-wise and also in ℓ_1 .

(ii) *For any $v \in S$, let $N_{n,v} := \sum_{k=0}^n \mathbb{1}_{\{Z_k=v\}}$, denote the number of times the color v is chosen up to time n . Then, as $n \rightarrow \infty$*

$$(1.6) \quad \frac{N_{n,v}}{n+1} \longrightarrow \pi_v \text{ a.s.},$$

where the convergence is coordinate-wise and also in ℓ_1 .

1.3. Background and Motivation. It is known (see, for example, Theorems 3.3(a) and 3.4(a) of [5]) that under our setup, as $n \rightarrow \infty$,

$$(1.7) \quad \frac{U_n}{n+t} \xrightarrow{p} \pi,$$

and also for any $v \in S$,

$$(1.8) \quad \mathbb{P}(Z_n = v) = \frac{\mathbb{E}[U_{n,v}]}{n+t} \longrightarrow \pi_v.$$

Recall, Z_n denotes the randomly chosen color at the $(n+1)$ -th draw, from the urn, when its (random) configuration is U_n . Our result strengthens this result to strong convergence. However, we would like to point out that the results in [5] (Theorems

3.3(a) and 3.4(a)), only needs assumption of ergodicity for the associated Markov chain, while our main result in this work needs more stronger assumption of uniform ergodicity of the associated Markov chain. As discussed above the two assumptions are identical when S is finite. It is worthwhile to note here that for S finite our result is essentially the classical result for Freedman-Pólya-Eggenberger type urn models [18, 3, 24, 2]. The classical results mainly use three types of techniques, namely, the martingale techniques [18, 8, 9, 12]; stochastic approximations [28] and embedding into continuous time pure birth processes [2, 24, 25, 3]. Typically, the analysis of a finite color urn is heavily dependent on the *Perron-Frobenius theory* [36] of matrices with positive entries and *Jordan Decomposition* of finite dimensional matrices [2, 18, 24, 25, 3, 8, 12]. Unfortunately, such techniques are unavailable when S is infinite, even when countable. Our method bypasses the use of such techniques and instead uses the newer approach developed in [5, 31]. Our extra assumption (uniform ergodicity) is needed only when S infinite. Thus the result stated above re-proves the classical result for the finite color urn model using the new technique. The result essentially completes the work developed in [5, 31] for the case when S is countable. We would like to note here that similar results for a null recurrent case (when the chain is a random walk) has been derived in [31, 26].

1.4. Discussion on the assumption of uniform ergodicity. As discussed above, when S is finite the assumption of uniform ergodicity is equivalent to the assumption of ergodicity of the associate Markov chain [29]. In particular, it holds for irreducible and aperiodic chain. However, when S is infinite it is indeed a much stronger assumption. Necessary and sufficient condition under which a chain is uniformly ergodic can be found in [32]. In particular, an irreducible and aperiodic chain on a countable state space is uniformly ergodic, if and only if, the so called *Doebelin's condition* is satisfied (see Section 16.2 of [32]). This condition is satisfied by many Markov chains on countable infinite state space, but it is indeed restrictive. We need this assumption in the proof we provide in Section 3. However, we do feel that this condition is not necessary in general. We, in fact, make the following conjecture:

Conjecture 1.3. *Consider an urn model $(U_n)_{n \geq 0}$ as defined by the equations (1.1) and (1.2), with colors indexed by a countably infinite set S , a balanced replacement matrix R , and an initial configuration U_0 . Assume that R is a stochastic matrix which is irreducible, aperiodic, positive recurrent with stationary distribution π . Then, the convergence in (1.5) and (1.6) holds a.s. and also in ℓ_1 .*

1.5. Outline. In the following section we provide some details about the representation of a balanced urn in terms of a branching Markov chain on a weighted random recursive tree, which is our main tool to prove Theorem 1.2. Section 3 provides the proof of Theorem 1.2. In Section 4, we discuss a non-trivial application of our main result.

2. COUPLING OF BRANCHING MARKOV CHAINS AND URN MODELS

It is known from [37, 5, 31, 26] that the law for the entire sequence of randomly selected colors $(Z_n)_{n \geq 0}$ can be represented in terms of a *branching Markov chain* on a *random recursive trees*. For the sake of completeness, we will briefly discuss this representation here. We will later use this representation to prove the main result of the paper.

2.1. Weighted random recursive tree. Random recursive trees (RRT) are well studied in literature, see Ch.6 of [14]. The weighted version for RRT has been introduced and defined in [26]. For $n \geq -1$, let \mathcal{T}_n be the random recursive tree on $n+2$ vertices, with o as the root, and the other vertices labeled as $\{w_0, w_1, \dots, w_n\}$, where the increasing subscripts of the vertices indicate the order in which they are attached. The root is given some initial weight $t > 0$. Every other node has weight 1. Initially, we start with \mathcal{T}_{-1} which consists only of the root, denoted by o . Now we construct recursively the sequence of trees $(\mathcal{T}_n)_{n \geq -1}$, where the parent of the incoming node in \mathcal{T}_n is chosen proportional to its weight, that is, the parent is the root o , with probability $t/(n+t+1)$, and any other vertex with probability $1/(n+t+1)$. Define the infinite random recursive tree as

$$(2.1) \quad \mathcal{T} := \bigcup_{n \geq -1} \mathcal{T}_n.$$

2.2. Branching Markov chain on RRT. The definition for *branching Markov chain on the random recursive tree*, which we abbreviate as BMC on RRT, as discussed in the context of this paper is available in details in [5, 31]. To facilitate convenient reading of this paper, we discuss briefly the BMC on RRT as available in [5].

Recall that the set of colors are indexed by a set S . Let $\Delta \notin S$ be a symbol. We say a stochastic process $(W_n)_{n \geq -1}$ with state space $S \cup \Delta$ is a *branching Markov chain on \mathcal{T}* , starting at the root o and at a position $W_{-1} = \Delta$, if for any $n \geq 0$ and for any $z \in S$

$$(2.2) \quad \mathbb{P}(W_n = z \mid W_{n-1}, W_{n-2}, \dots, W_{-1}; \mathcal{T}_n) = \begin{cases} U_0(z)/t, & \text{if } \overleftarrow{w}_n = o, \\ R(W_j, z) & \text{if } \overleftarrow{w}_n = w_j, \end{cases}$$

where \overleftarrow{w}_n is the parent of the vertex w_n in \mathcal{T}_n . Note that we here denote the vertices of \mathcal{T}_n as $\{o, w_0, w_1, \dots, w_n\}$.

2.3. Representation Theorems. The coupling of $(Z_n)_{n \geq 0}$ and $(W_n)_{n \geq 0}$ is available in details in [5, 31]. Here, we follow same notations as in [5]. The following representation is available in the Theorem 2.1 in [5].

$$(2.3) \quad (Z_n)_{n \geq 0} \stackrel{d}{=} (W_n)_{n \geq 0}.$$

3. PROOF OF THE MAIN RESULTS

Recall that \mathcal{T}_n denotes the weighted RRT with $n+2$ vertices; for convenience we use \mathcal{T}_n also to denote its vertex set $\{o, w_0, w_1, \dots, w_n\}$. Let $\mathcal{T}'_n := \mathcal{T}_n \setminus \{o\} = \{w_0, w_1, \dots, w_n\}$, the set of $n+1$ vertices excluding the root. Note that the RRT \mathcal{T}_n is random, but the vertex set is non-random.

For $u, w \in \mathcal{T}_n$, let $d(u, w)$ denote the graph distance between u and w . In particular, $d(o, u)$ is the depth of u , which we also denote by $d(u)$.

We begin by proving the following lemma, where ρ is as in Definition 1.1.

Lemma 3.1. *Let $\mathcal{L}(u, w)$ denote the least common ancestor for the vertices u, w in the random recursive tree (RRT). Given the RRT \mathcal{T}_n , we have for some suitable constant $C > 0$,*

$$(3.1) \quad \text{Cov}(\mathbb{1}_{\{W_u=v\}}, \mathbb{1}_{\{W_w=v\}} \mid \mathcal{T}_n) \leq C \rho^{\max(d(u, \mathcal{L}(u, w)), d(w, \mathcal{L}(u, w)))} \leq C \rho^{d(u, w)/2}.$$

Proof. Let us denote by \mathbb{P}_n the conditional probability given the RRT \mathcal{T}_n . By definition,

$$\text{Cov}(\mathbb{1}_{\{W_u=v\}}, \mathbb{1}_{\{W_w=v\}} \mid \mathcal{T}_n) = \mathbb{P}_n(W_u=v, W_w=v) - \mathbb{P}_n(W_u=v)\mathbb{P}_n(W_w=v).$$

With $\mathcal{L}(u, w)$ denoting the least common ancestor between u and w , it is easy to see that

$$\mathbb{P}_n(W_u=v, W_w=v) = \sum_{s \in S} \mathbb{P}_n(W_{\mathcal{L}(u,w)}=s) R^{d(u, \mathcal{L}(u,w))}(s, v) R^{d(w, \mathcal{L}(u,w))}(s, v).$$

Thus,

$$\begin{aligned} & \text{Cov}(\mathbb{1}_{\{W_u=v\}}, \mathbb{1}_{\{W_w=v\}} \mid \mathcal{T}_n) \\ &= \sum_{s \in S} \mathbb{P}_n(W_{\mathcal{L}(u,w)}=s) R^{d(u, \mathcal{L}(u,w))}(s, v) R^{d(w, \mathcal{L}(u,w))}(s, v) \\ & \quad - \sum_{s, s' \in S} \mathbb{P}_n(W_{\mathcal{L}(u,w)}=s) \mathbb{P}_n(W_{\mathcal{L}(u,w)}=s') R^{d(u, \mathcal{L}(u,w))}(s, v) R^{d(w, \mathcal{L}(u,w))}(s', v) \\ &= \sum_{s \in S} \mathbb{P}_n(W_{\mathcal{L}(u,w)}=s) R^{d(u, \mathcal{L}(u,w))}(s, v) \left[R^{d(w, \mathcal{L}(u,w))}(s, v) \right. \\ & \quad \left. - \sum_{s' \in S} \mathbb{P}_n(W_{\mathcal{L}(u,w)}=s') R^{d(w, \mathcal{L}(u,w))}(s', v) \right] \\ &= \sum_{s \in S} \mathbb{P}_n(W_{\mathcal{L}(u,w)}=s) R^{d(u, \mathcal{L}(u,w))}(s, v) \left[(R^{d(w, \mathcal{L}(u,w))}(s, v) - \pi_v) \right. \\ & \quad \left. - \sum_{s' \in S} \mathbb{P}_n(W_{\mathcal{L}(u,w)}=s') (R^{d(w, \mathcal{L}(u,w))}(s', v) - \pi_v) \right]. \end{aligned}$$

The last equality is obtained by adding and subtracting π_v inside the final square bracket. Recall that we have assumed uniform ergodicity for the Markov chain, so for both s, s' we have $|R^{d(w, \mathcal{L}(u,w))}(s, v) - \pi_v| < C\rho^{d(w, \mathcal{L}(u,w))}$, which implies that

$$(3.2) \quad |\text{Cov}(\mathbb{1}_{\{W_u=v\}}, \mathbb{1}_{\{W_w=v\}} \mid \mathcal{T}_n)| \leq 2C\rho^{d(w, \mathcal{L}(u,w))}.$$

The first inequality in (3.1) follows by symmetry. The second inequality is obvious as $0 < \rho < 1$. \square

Lemma 3.2. *Fix r with $0 < r < 1$ and define*

$$(3.3) \quad A_n = A_n(r) := \mathbb{E} \sum_{u \in \mathcal{T}'_n} r^{d(u)},$$

$$(3.4) \quad B_n = B_n(r) := \mathbb{E} \sum_{u, w \in \mathcal{T}'_n} r^{d(u,w)}.$$

Then, for some constant C (possibly depending on r and t) and all $n \geq 1$,

$$(3.5) \quad A_n \leq Cn^r,$$

$$(3.6) \quad B_n \leq \begin{cases} Cn^{2r}, & \frac{1}{2} < r < 1, \\ Cn \log(n+1), & r = \frac{1}{2}, \\ Cn, & 0 < r < \frac{1}{2}. \end{cases}$$

Much more precise asymptotic formulas can be derived by the same method, but we do not need them.

Proof. Recall that w_n is the $(n+1)$ -th coming vertex, and assume that w_n is attached to $w \in \mathcal{T}_{n-1}$. Then, for all $u \in \mathcal{T}_{n-1}$,

$$(3.7) \quad d(u, w_n) = d(u, w) + 1.$$

Hence,

$$(3.8) \quad \begin{aligned} A_n &= A_{n-1} + \mathbb{E}r^{d(w_n)} \\ &= A_{n-1} + \frac{1}{n+t} \mathbb{E} \left(\sum_{u \in \mathcal{T}'_{n-1}} r^{d(u)+1} \right) + \frac{t}{n+t} r \\ &= \left(1 + \frac{r}{n+t} \right) A_{n-1} + \frac{tr}{n+t}. \end{aligned}$$

Consequently, by induction and using $A_0 = r$,

$$(3.9) \quad A_n = r \sum_{k=0}^n \left(\frac{t}{k+t} \prod_{j=k+1}^n \left(1 + \frac{r}{j+t} \right) \right) = rt \sum_{k=0}^n \frac{\Gamma(n+1+t+r)}{\Gamma(n+1+t)} \frac{\Gamma(k+t)}{\Gamma(k+1+t+r)}.$$

By standard asymptotics for the Gamma function (following from Stirling's formula), this yields

$$(3.10) \quad A_n \leq rt \sum_{k=0}^n C \frac{(n+1)^r}{(k+1)^{r+1}} \leq C(n+1)^r,$$

showing (3.5).

For (3.6) we argue similarly. We have,

$$\begin{aligned} B_n &= B_{n-1} + 2\mathbb{E} \left(\sum_{u \in \mathcal{T}'_{n-1}} r^{d(u, w_n)} \right) + 1 \\ &= B_{n-1} + \frac{2}{n+t} \mathbb{E} \left(\sum_{u, w \in \mathcal{T}'_{n-1}} r^{d(u, w)+1} \right) + \frac{2t}{n+t} \mathbb{E} \left(\sum_{u \in \mathcal{T}'_{n-1}} r^{d(u, o)+1} \right) + 1 \\ &= \left(1 + \frac{2r}{n+t} \right) B_{n-1} + \frac{2rt}{n+t} A_{n-1} + 1 \end{aligned}$$

and, with $A_{-1} := 0$, using $B_0 = 1$,

$$(3.11) \quad B_n = \sum_{k=0}^n \left(\left(1 + \frac{2rt}{k+t} A_{k-1} \right) \prod_{j=k+1}^n \left(1 + \frac{2r}{j+t} \right) \right).$$

We use the crude estimate $A_{k-1} \leq k$ and estimate the product in (3.11) using Gamma functions as in (3.9)–(3.10) (with r replaced by $2r$); this yields

$$(3.12) \quad B_n \leq C \sum_{k=0}^n \prod_{j=k+1}^n \left(1 + \frac{2r}{j+t} \right) \leq C \sum_{k=0}^n \frac{(n+1)^{2r}}{(k+1)^{2r}}.$$

This implies (3.6) by a simple summation. \square

3.1. Proof of the Theorem 1.2.

Proof. We observe that the basic recursion (1.1) can also be written as

$$(3.13) \quad U_{n+1} = U_n + \chi_{n+1}R$$

where $\chi_{n+1} = (\chi_{n+1,v})_{v \in S}$ is such that $\chi_{n+1,Z_n} = 1$ and $\chi_{n+1,u} = 0$ if $u \neq Z_n$. In other words,

$$(3.14) \quad U_{n+1} = U_n + R_{Z_n}$$

where R_{Z_n} is the Z_n -th row of the matrix R . Hence,

$$(3.15) \quad U_{n+1} = U_0 + \sum_{k=1}^{n+1} \chi_k R,$$

$$(3.16) \quad \frac{U_{n+1} - U_0}{n+t} = \frac{1}{n+t} \sum_{k=1}^{n+1} \chi_k R.$$

To prove (1.5), it is thus by (3.16), and since $\frac{n+1}{n+t} \rightarrow 1$ as $n \rightarrow \infty$, enough to show that

$$(3.17) \quad \frac{1}{n+1} \sum_{k=1}^{n+1} \chi_k R \rightarrow \pi \quad \text{in } \ell^1(S), \text{ a.s.}$$

Since R is balanced, the mapping $x \mapsto xR$ is a bounded map $\ell^1(S) \rightarrow \ell^1(S)$, and since furthermore $\pi R = \pi$, to prove (3.17) it is enough to show that as $n \rightarrow \infty$,

$$(3.18) \quad \frac{1}{n+1} \sum_{k=1}^{n+1} \chi_k \rightarrow \pi \quad \text{in } \ell^1(S), \text{ a.s.}$$

Both sides of (3.18) can be regarded as probability distributions on S , and therefore, the convergence in ℓ^1 is equivalent to convergence of every coordinate, i.e., to

$$(3.19) \quad \frac{1}{n+1} \sum_{k=1}^{n+1} \chi_{k,v} \rightarrow \pi_v \quad \text{a.s. for every } v \in S.$$

Moreover, (1.6) is just another way to write (3.19). Hence, to show the theorem, it suffices to show (3.19).

Recall that $\chi_{k,v} = \mathbb{1}_{\{Z_{k-1}=v\}}$. From Theorem 3.3(a) of [5] (which easily is extended to general t), it follows that

$$(3.20) \quad \frac{1}{n+1} \mathbb{E} \left[\sum_{k=1}^{n+1} \chi_{k,v} \right] = \frac{1}{n+1} \sum_{k=0}^n \mathbb{P}(Z_k = v) \rightarrow \pi_v \text{ as } n \rightarrow \infty.$$

Note that $|\mathbb{1}_{\{Z_k=v\}} - \mathbb{E}\mathbb{1}_{\{Z_k=v\}}| \leq 1$. Therefore, from the Strong Law of Large Numbers for correlated random variables [30, Theorem 1], it follows that, if we prove

$$(3.21) \quad \sum_{n \geq 0} \frac{1}{n+1} \text{Var} \left(\frac{1}{n+1} \sum_{k=0}^n \mathbb{1}_{\{Z_k=v\}} \right) < \infty,$$

then, as $n \rightarrow \infty$,

$$(3.22) \quad \frac{1}{n+1} \sum_{k=1}^{n+1} \chi_{k,v} = \frac{1}{n+1} \sum_{k=0}^n \mathbb{1}_{\{Z_k=v\}} \rightarrow \pi_v \text{ a.s.,}$$

which will complete the proof. In other words, if we define

$$(3.23) \quad J_{n,v} := \text{Var} \left(\sum_{k=0}^n \mathbb{1}_{\{Z_k=v\}} \right),$$

then, it suffices to show that

$$(3.24) \quad \sum_{n=1}^{\infty} \frac{1}{n^3} J_{n,v} < \infty.$$

Now, recalling (2.3), (3.23) can be expanded as

$$(3.25) \quad J_{n,v} = \text{Var} \left(\sum_{k=0}^n \mathbb{1}_{\{W_k=v\}} \right) = \sum_{u,w \in \mathcal{T}'_n} \text{Cov} \left(\mathbb{1}_{\{W_u=v\}}, \mathbb{1}_{\{W_w=v\}} \right).$$

We use the conditional covariance formula to get

$$(3.26) \quad \begin{aligned} \text{Cov} \left(\mathbb{1}_{\{W_u=v\}}, \mathbb{1}_{\{W_w=v\}} \right) &= \mathbb{E} \left[\text{Cov} \left(\mathbb{1}_{\{W_u=v\}}, \mathbb{1}_{\{W_w=v\}} \mid \mathcal{T}_n \right) \right] \\ &+ \text{Cov} \left(\mathbb{E} \left[\mathbb{1}_{\{W_u=v\}} \mid \mathcal{T}_n \right], \mathbb{E} \left[\mathbb{1}_{\{W_w=v\}} \mid \mathcal{T}_n \right] \right). \end{aligned}$$

Now, using Lemma 3.1, we obtain

$$\text{Cov} \left(\mathbb{1}_{\{W_u=v\}}, \mathbb{1}_{\{W_w=v\}} \mid \mathcal{T}_n \right) \leq C \rho^{\frac{d(u,w)}{2}},$$

where $d(u,w)$ denotes the graph distance between u and w , and C is a suitable positive constant. Therefore, from (3.25)–(3.26), the contribution to $J_{n,v}$ from the first part of (3.26) is at most

$$(3.27) \quad C \mathbb{E} \left(\sum_{u,w \in \mathcal{T}'_n} \rho^{d(u,w)/2} \right) = CB(\rho^{1/2}),$$

where we recall (3.4) and take $r := \rho^{1/2}$.

For the second part on the RHS of (3.26), we have that, given \mathcal{T}_n , the distribution of W_u is $(U_0/t)R^{d(u)}$. Hence,

$$(3.28) \quad \mathbb{E} \left(\mathbb{1}_{\{W_u=v\}} \mid \mathcal{T}_n \right) = (U_0/t)R^{d(u)}(v),$$

and thus it follows from the uniform ergodicity assumption (1.4) that

$$(3.29) \quad \left| \mathbb{E} \left(\mathbb{1}_{\{W_u=v\}} \mid \mathcal{T}_n \right) - \pi_v \right| \leq C \rho^{d(u)}.$$

Consequently,

$$\begin{aligned} \sum_{u,w \in \mathcal{T}'_n} \text{Cov} \left(\mathbb{E} \left[\mathbb{1}_{\{W_u=v\}} \mid \mathcal{T}_n \right], \mathbb{E} \left[\mathbb{1}_{\{W_w=v\}} \mid \mathcal{T}_n \right] \right) &= \text{Var} \left(\sum_{u \in \mathcal{T}'_n} \mathbb{E} \left[\mathbb{1}_{\{W_u=v\}} \mid \mathcal{T}_n \right] \right) \\ &= \text{Var} \left(\sum_{u \in \mathcal{T}'_n} \left(\mathbb{E} \left[\mathbb{1}_{\{W_u=v\}} \mid \mathcal{T}_n \right] - \pi_v \right) \right) \leq \mathbb{E} \left(\sum_{u \in \mathcal{T}'_n} \left(\mathbb{E} \left[\mathbb{1}_{\{W_u=v\}} \mid \mathcal{T}_n \right] - \pi_v \right) \right)^2 \\ &\leq \mathbb{E} \left(C \sum_{u \in \mathcal{T}'_n} \rho^{d(u)} \right)^2 = C \mathbb{E} \sum_{u,w \in \mathcal{T}'_n} \rho^{d(u)+d(w)} \leq C \mathbb{E} \sum_{u,w \in \mathcal{T}'_n} \rho^{d(u,w)} \\ &= CB_n(\rho), \end{aligned}$$

where we use the fact that $d(u) + d(w) \geq d(u, w)$ and $0 < \rho < 1$ to obtain the last inequality. Hence, the contribution to $J_{n,v}$ from the second part of (3.26) is at most $CB_n(\rho)$.

Combining the contributions from the two parts of (3.26), we thus have shown that, recalling $0 < \rho < 1$,

$$(3.30) \quad J_{n,v} \leq CB_n(\rho^{1/2}) + CB_n(\rho) \leq CB_n(\rho^{1/2}).$$

Hence, we can use Lemma 3.2 and conclude (3.24), which completes the proof. \square

4. RANDOM WALK WITH LINEAR REINFORCEMENT ON THE STAR GRAPH

In this section, we consider a linearly reinforced random walk model on the infinite (countable) *star graph*. We will show that the almost sure convergence for the local times for this walk can be derived using our main result stated in Section 1.2.

Let us consider a special type of vertex-reinforced nearest neighbor random walk $(X_n)_{n \geq 0}$ on an infinite star graph, with a loop at the root. We denote the root by v_0 and the other vertices by $v_i, i \geq 1$. Each edge is regarded as a pair of directed edges in opposite directions; the notation (v_i, v_j) indicates that the edge is from v_i to v_j . We impose the following condition on the walk that v_0 is a special vertex, in the sense that, whenever the walker takes the edge (v_j, v_0) , for any j , it puts an additional weight of $\alpha_j := (\alpha_{j,i})_{i \geq 0}$ on the vertices, such that, $\sum_i \alpha_{j,i} < \infty$. If the edge taken is (v_0, v_j) , $j \neq 0$ then no vertex is reinforced.

Initially, $X_0 \equiv v_0$, the walker is at the root, and jumps to one of the adjacent neighbors with probability proportional to the given weights δ_i , such that, $\delta := \sum_{i \geq 0} \delta_i < \infty$. At any time $n \geq 1$, the transition probabilities for the random walk is governed by

$$(4.1) \quad \mathbb{P}(X_{n+1} = v_j | X_n = v_i) = \begin{cases} \frac{\Delta_{n,j}}{\sum_k \Delta_{n,k}}, & \text{when } i = 0, \\ \mathbb{1}_{\{j=0\}}, & \text{for } i \geq 1. \end{cases}$$

where $\Delta_{n,j}$ denotes the weight at the vertex v_j at time n .

Observe that, if we denote by σ_k , the random time at which the weights are updated for the k -th time, then $\sigma_{k+1} = \sigma_k + Y_{k+1}$, where $Y_{k+1} \in \{1, 2\}$ is a random variable, such that,

$$\mathbb{P}(Y_{k+1} = 1 | \Delta_0, \Delta_{\sigma_1}, \dots, \Delta_{\sigma_k}) = \frac{\Delta_{\sigma_k, 0}}{\sum_j \Delta_{\sigma_k, j}}.$$

Therefore, the weight sequence at these updating random times can be coupled with an infinite color urn model, as described below.

Consider the urn model with colors indexed by $S := \{0, 1, 2, \dots\}$, and an initial composition $U_0 = (\delta_i)_{i \geq 0}$. The replacement matrix is such that the j -th row of the matrix is α_j . Since the graph is a star graph, for the random walk to take a step along (v_i, v_0) , $i \neq 0$, it implies that the walker has jumped along the edge (v_0, v_i) according to the transition probabilities given by (4.1). So, if we consider the sequence of weights at time $\sigma_1, \sigma_2, \dots$, then the processes are coupled such that

$$(4.2) \quad (\Delta_{\sigma_n})_{n \geq 0} = (U_n)_{n \geq 0}.$$

In particular, if $\sum_i \alpha_{j,i} = 1$ for each j , then the replacement matrix is a stochastic matrix. Henceforth, we assume that α_j is a probability vector for every $j \geq 0$.

We also assume that α_j are such that, the Markov chain corresponding to the replacement matrix is irreducible, aperiodic and uniformly ergodic.

A particular example of such a matrix is when $\alpha_0 = (p_j)_{j \geq 0}$, with $p_j > 0$ and $\sum_j p_j = 1$, and, for $j \neq 0$, $\alpha_{j,i} = 1$ if $i = 0$, and 0 otherwise. (Our conditions, including uniform ergodicity, are easily verified.)

Theorem 4.1. *Let X_n be a vertex reinforced random walk on an infinite star graph with a loop at the root, such that the replacement matrix is an irreducible, aperiodic and uniformly ergodic stochastic matrix. Let the transition probabilities of X_n be as in (4.1). If we denote by σ_n , the n -th update time, then as $n \rightarrow \infty$,*

$$(4.3) \quad \frac{\sigma_n}{n+1} \longrightarrow 2 - \pi_0, \text{ a.s. and in } L_1.$$

Furthermore, for any $j \geq 0$, as $n \rightarrow \infty$

$$(4.4) \quad \frac{\Delta_{n,j}}{n+\delta} \longrightarrow \frac{\pi_j}{2-\pi_0}, \text{ a.s.},$$

where π is the stationary distribution of the coupled urn process as defined in (4.2).

Proof. As observed earlier $\sigma_{k+1} = \sigma_k + Y_{k+1}$, where $Y_{k+1} \in \{1, 2\}$, and

$$\mathbb{P}(Y_{k+1} = 1 | \Delta_0, \Delta_{\sigma_1}, \dots, \Delta_{\sigma_k}) = \frac{\Delta_{\sigma_k, 0}}{\sum_j \Delta_{\sigma_k, j}}.$$

Let us denote by $\tilde{\sigma}_k := \sum_{i=0}^k \mathbb{1}_{\{Y_i=1\}}$. Then from the conditional distribution of Y_k above and from (4.2), we have, using the coupling above,

$$(4.5) \quad \tilde{\sigma}_n = \sum_{k=0}^n \mathbb{1}_{\{Z_k=0\}} = N_{n,0},$$

where Z_k denotes the random color of the ball selected in the coupled urn model. From Theorem 1.2(ii), we know that as $n \rightarrow \infty$, $N_{n,0}/(n+1) \rightarrow \pi_0$ a.s. Thus, as $n \rightarrow \infty$,

$$(4.6) \quad \frac{\tilde{\sigma}_n}{n+1} \longrightarrow \pi_0, \text{ a.s.}$$

Since $\sigma_n = \tilde{\sigma}_n + 2(n - \tilde{\sigma}_n)$, so (4.3) follows immediately. Since $0 \leq \frac{\sigma_n}{n+1} \leq 1$, the L_1 convergence in (4.3) follows by dominated convergence theorem.

Denote by $m(n) := \sup\{k : \sigma_k \leq n\}$. Then it follows from (4.3) that, as $n \rightarrow \infty$, $\frac{m(n)}{n+1} \rightarrow \frac{1}{2-\pi_0}$, a.s. From (4.2) and Theorem 1.2(i), we have as $n \rightarrow \infty$,

$$(4.7) \quad \frac{\Delta_{n,j}}{n+\delta} = \frac{\Delta_{\sigma_{m(n)},j}}{m(n)} \frac{m(n)}{n+\delta} = \frac{U_{m(n),j}}{m(n)} \frac{m(n)}{n+\delta} \longrightarrow \frac{\pi_j}{2-\pi_0}, \text{ a.s.}$$

□

ACKNOWLEDGEMENT

The authors are grateful to Colin Desmarais and Cecilia Holmgren for various discussions they had with them over the course of working on this paper. We are grateful for their insightful ideas and remarks, which helped us improve the quality of this paper. We also thank the anonymous referees for their comments and suggestions, which prompted us to add more references. This improved the overall quality of the paper.

REFERENCES

- [1] David J. Aldous. Exchangeability and related topics. In *École d'été de probabilités de Saint-Flour, XIII—1983*, volume 1117 of *Lecture Notes in Math.*, pages 1–198. Springer, Berlin, 1985.
- [2] Krishna B. Athreya and Samuel Karlin. Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist.*, 39:1801–1817, 1968.
- [3] Zhi-Dong Bai and Feifang Hu. Asymptotics in randomized urn models. *Ann. Appl. Probab.*, 15(1B):914–940, 2005.
- [4] Antar Bandyopadhyay and Debleena Thacker. Rate of convergence and large deviation for the infinite color Pólya urn schemes. *Statist. Probab. Lett.*, 92:232–240, 2014.
- [5] Antar Bandyopadhyay and Debleena Thacker. A new approach to pólya urn schemes and its infinite color generalization. (<https://arxiv.org/abs/1606.05317>), 2016.
- [6] Antar Bandyopadhyay and Debleena Thacker. Pólya urn schemes with infinitely many colors. *Bernoulli*, 23(4B):3243–3267, 2017.
- [7] David Blackwell and James B. MacQueen. Ferguson distributions via Pólya urn schemes. *Ann. Statist.*, 1:353–355, 1973.
- [8] Arup Bose, Amites Dasgupta, and Krishanu Maulik. Multicolor urn models with with reducible replacement matrices. *Bernoulli*, 15(1):279–295, 2009.
- [9] Arup Bose, Amites Dasgupta, and Krishanu Maulik. Strong laws for balanced triangular urns. *J. Appl. Probab.*, 46(2):571–584, 2009.
- [10] May-Ru Chen, Shoou-Ren Hsiau, and Ting-Hsin Yang. A new two-urn model. *J. Appl. Probab.*, 51(2):590–597, 2014.
- [11] May-Ru Chen and Markus Kuba. On generalized Pólya urn models. *J. Appl. Probab.*, 50(4):1169–1186, 2013.
- [12] Amites Dasgupta and Krishanu Maulik. Strong laws for urn models with balanced replacement matrices. *Electron. J. Probab.*, 16:no. 63, 1723–1749, 2011.
- [13] Peter Donnelly and Thomas G. Kurtz. The asymptotic behavior of an urn model arising in population genetics. *Stochastic Process. Appl.*, 64(1):1–16, 1996.
- [14] Michael Drmota. *SpringerWienNewYork, Vienna*, 2009.
- [15] William Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons, Inc., New York-London-Sydney, 1968.
- [16] Thomas S. Ferguson. A Bayesian analysis of some nonparametric problems. *Ann. Statist.*, 1:209–230, 1973.
- [17] Philippe Flajolet, Philippe Dumas, and Vincent Puyhaubert. Some exactly solvable models of urn process theory. In *Fourth Colloquium on Mathematics and Computer Science Algorithms, Trees, Combinatorics and Probabilities*, Discrete Math. Theor. Comput. Sci. Proc., AG, pages 59–118. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2006.
- [18] Raúl Gouet. Strong convergence of proportions in a multicolor Pólya urn. *J. Appl. Probab.*, 34(2):426–435, 1997.
- [19] Olle Häggström. On the central limit theorem for geometrically ergodic Markov chains. *Probab. Theory Related Fields*, 132(1):74–82, 2005.
- [20] Ulrich Martin Hirth. From GEM back to Dirichlet via Hoppe’s urn. *Combin. Probab. Comput.*, 6(2):185–195, 1997.
- [21] Lars Holst. The number of two consecutive successes in a Hoppe-Pólya urn. *J. Appl. Probab.*, 45(3):901–906, 2008.
- [22] Fred M. Hoppe. Pólya-like urns and the Ewens’ sampling formula. *J. Math. Biol.*, 20(1):91–94, 1984.
- [23] Fred M. Hoppe. The sampling theory of neutral alleles and an urn model in population genetics. *J. Math. Biol.*, 25(2):123–159, 1987.
- [24] Svante Janson. Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Process. Appl.*, 110(2):177–245, 2004.
- [25] Svante Janson. Limit theorems for triangular urn schemes. *Probab. Theory Related Fields*, 134(3):417–452, 2006.
- [26] Svante Janson. A.s. convergence for infinite colour Pólya urns associated with random walks. (<https://arxiv.org/abs/1803.04207>), 2018.

- [27] Svante Janson. Random replacements in Pólya urns with infinitely many colours. *Electron. Commun. Probab.*, 24:Paper No. 23, 11, 2019.
- [28] Sophie Laruelle and Gilles Pagès. Randomized urn models revisited using stochastic approximation. *Ann. Appl. Probab.*, 23(4):1409–1436, 2013.
- [29] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009.
- [30] Russell Lyons. Strong laws of large numbers for weakly correlated random variables. *Michigan Math. J.*, 35(3):353–359, 1988.
- [31] Cécile Mailler and Jean-François Marckert. Measure-valued Pólya urn processes. *Electron. J. Probab.*, 22:Paper No. 26, 33, 2017.
- [32] Sean Meyn and Richard L. Tweedie. *Markov chains and stochastic stability*. Cambridge University Press, Cambridge, second edition, 2009.
- [33] Robin Pemantle. A survey of random processes with reinforcement. *Probab. Surv.*, 4:1–79, 2007.
- [34] J. Pitman. *Combinatorial stochastic processes*, volume 1875 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006.
- [35] G. Pólya. Sur quelques points de la théorie des probabilités. *Ann. Inst. H. Poincaré*, 1(2):117–161, 1930.
- [36] Eugene Seneta. *Non-negative matrices and Markov chains*. Springer Series in Statistics. Springer, New York, 2006. Revised reprint of the second (1981) edition [Springer-Verlag, New York; MR0719544].
- [37] Debleena Thacker. *Infinite Color Urn Models*. 2015. Thesis (Ph.D.)—Indian Statistical Institute, Delhi Centre.
- [38] Erik Thörnblad. The dominating colour of an infinite Pólya urn model. *J. Appl. Probab.*, 53(3):914–924, 2016.

(Antar Bandyopadhyay) THEORETICAL STATISTICS AND MATHEMATICS UNIT, INDIAN STATISTICAL INSTITUTE, DELHI CENTRE, 7 S. J. S. SANSANWAL MARG, NEW DELHI 110016, INDIA

THEORETICAL STATISTICS AND MATHEMATICS UNIT, INDIAN STATISTICAL INSTITUTE, KOLKATA; 203 B. T. ROAD, KOLKATA 700108, INDIA

Email address: antar@isid.ac.in

(Svante Janson) DEPARTMENT OF MATHEMATICS, UPPSALA UNIVERSITY, BOX 480, 751 06 UPPSALA, SWEDEN,

Email address: svante@math.uu.se

(Debleena Thacker) DEPARTMENT OF MATHEMATICS, UPPSALA UNIVERSITY, BOX 480, 751 06 UPPSALA, SWEDEN,

Email address: thackerdebleena@gmail.com