

# PÓLYA URN SCHEMES WITH INFINITELY MANY COLORS

ANTAR BANDYOPADHYAY AND DEBLEENA THACKER

ABSTRACT. In this work, we introduce a class of *balanced* urn schemes with infinitely many colors indexed by  $\mathbb{Z}^d$ , where the replacement schemes are given by the transition matrices associated with bounded increment random walks. We show that the color of the  $n$ -th selected ball follows a Gaussian distribution on  $\mathbb{R}^d$  after  $\mathcal{O}(\log n)$  centering and  $\mathcal{O}(\sqrt{\log n})$  scaling irrespective of whether the underlying walk is null recurrent or transient. We also provide finer asymptotic similar to local limit theorems for the expected configuration of the urn. The proofs are based on a novel representation of the color of the  $n$ -th selected ball as “slowed down” version of the underlying random walk.

## 1. INTRODUCTION

**1.1. Background and Motivation.** In recent years, there has been a wide variety of work on *random reinforcement models* of various kind [3, 11, 13, 15, 17–20, 24, 29, 30, 33, 34, 36]. In particular, there has been several work on different kinds of *urn models* and their generalizations [3, 11, 13, 15, 18, 19, 24, 29–31]. The so called *urn schemes* was started by the seminal work of Pólya [38]. Since then various generalizations with finitely many colors have been studied in the literature [1–3, 11–14, 19, 24–29, 35]. See [36] for an extensive survey of the known results. However, other than the classical work by Blackwell and MacQueen [10], there has not been much development of infinite color generalization of the Pólya urn scheme. In this paper, we introduce and analyze a new Pólya type urn scheme with infinitely many colors indexed by  $\mathbb{Z}^d$ .

In this work, we will only consider *balanced* urn schemes. More precisely, if  $R := ((R(i, j)))_{i, j \in \mathbb{Z}^d}$  denotes the *replacement matrix*, that is,  $R(i, j) \geq 0$  is the *number of balls of color  $j$  to be placed in the the urn when the color of the selected ball is  $i$* , then for a balanced urn, all row sums of  $R$  are constant. In this case, we may assume  $R$  is a *stochastic matrix*. We will also assume that the starting configuration  $U_0 := (U_{0, j})_{j \in \mathbb{Z}^d}$  is a probability distribution on the set of colors, namely,  $\mathbb{Z}^d$ . As we will see from the proofs of our main results, this apparent loss of generality can easily be removed if  $\sum_{j \in \mathbb{Z}^d} U_{0, j} < \infty$ . Since  $R$  is a stochastic matrix and  $U_0$  a probability distribution on  $\mathbb{Z}^d$ , we can now consider a Markov chain on  $\mathbb{Z}^d$  with transition matrix  $R$  and initial distribution  $U_0$ . We will call such a chain, a chain associated with the urn model and vice-versa.

In this work, we are interested to study the process when  $R$  is the transition matrix of a bounded increment random walk on  $\mathbb{Z}^d$ . This is a novel generalization of the Pólya urn scheme, which we term as *urn model with infinitely many colors*.

---

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

*Key words and phrases*. Central limit theorem, infinite color urn, local limit theorem, random walk, reinforcement processes, urn models.

Our main motivations to study such a process have been twofold. One is to investigate, whether the asymptotic properties of a (balanced) urn with infinitely many colors, depend on the recurrence/transience of the underlying states. This is indeed the case for finitely many colors [11, 12, 19, 27–29]. The bounded increment random walks on  $\mathbb{Z}^d$  is a rich class of examples of Markov chains on infinite states covering both the transient and (null) recurrent cases. We will show that the asymptotic configuration is approximately Gaussian, irrespective of whether the underlying walk is transient or recurrent. Our other motivation comes from the work of Blackwell and MacQueen [10], where the authors introduced a possibly infinite color generalization of Pólya urn scheme, with only a diagonal replacement scheme. The model then described a process whose limiting distribution is the *Ferguson distribution* [9, 10], also known as the *Dirichlet process prior* in the Bayesian statistics literature [23]. Our model complements this work where we consider replacement mechanisms with non-zero off-diagonal entries. We would also like to point out that due to the presence of off-diagonal entries in the replacement matrix our models do not exhibit exchangeability of the sequence of colors. Hence, the techniques used in this paper are entirely different and new.

Finally, we would like to note here that our model is also a generalization of a subclass of models studied in [15], namely the class of *linearly reinforced models*. In [15] the authors prove that for such models cardinality of all the colors will grow to infinity. As we will see in Section 2, our results will not only show that the cardinality of all colors will grow to infinity, but will also provide the exact rates of their growths.

**1.2. Model.** Let  $\{X_j\}_{j \geq 1}$  be i.i.d. random vectors taking values in  $\mathbb{Z}^d$  with probability mass function  $p(u) := \mathbb{P}(X_1 = u)$ ,  $u \in \mathbb{Z}^d$ . We assume that the distribution of  $X_1$  is bounded, that is there exists a non-empty finite subset  $B \subseteq \mathbb{Z}^d$  such that  $p(u) = 0$  for all  $u \notin B$ . It is worthwhile to note that the assumption that  $B$  is finite may be removed. Instead, if we assume  $X_1$  has moment generating function on an open interval around 0, then all the results of this paper hold. But for simplicity, we will assume  $B$  to be finite.

Throughout this paper we take the convention of writing all vectors as row vectors. Thus for a vector  $x \in \mathbb{R}^d$  we will write  $x^T$  to denote it as a column vector. The notation  $\langle \cdot, \cdot \rangle$  will denote the usual Euclidean inner product on  $\mathbb{R}^d$  and  $\|\cdot\|$  the Euclidean norm. We shall always write

$$(1) \quad \begin{aligned} \mu &:= \mathbb{E}[X_1] \\ \Sigma &:= \mathbb{E}[X_1^T X_1] \\ e(\lambda) &:= \mathbb{E}[\exp(\langle \lambda, X_1 \rangle)], \lambda \in \mathbb{R}^d. \end{aligned}$$

We shall write  $\Sigma := ((\sigma_{ij}))_{1 \leq i, j \leq d}$  and assume that it is a positive definite matrix. Also  $\Sigma^{\frac{1}{2}}$  will denote the unique *positive definite square root* of  $\Sigma$ , that is,  $\Sigma^{\frac{1}{2}}$  is a positive definite matrix such that  $\Sigma = \Sigma^{\frac{1}{2}} \Sigma^{\frac{1}{2}}$ . When  $d = 1$ , we will denote the mean and the second moment (and not the variance) simply by  $\mu$  and  $\sigma^2$  respectively and in that case we assume  $\sigma^2 > 0$ .

Let  $S_n := X_0 + X_1 + \cdots + X_n$ ,  $n \geq 0$  be the random walk on  $\mathbb{Z}^d$  starting at  $X_0$  distributed according to  $U_0$  and with increments  $\{X_j\}_{j \geq 1}$  which are independent. Needless to say that  $\{S_n\}_{n \geq 0}$  is Markov chain with state space  $\mathbb{Z}^d$ , initial

distribution given by the distribution of  $X_0$  and the transition matrix

$$R := ((p(v-u))_{u,v \in \mathbb{Z}^d}.$$

In this work, we consider the following infinite color generalization of Pólya urn scheme where the colors are indexed by  $\mathbb{Z}^d$ . Let  $U_n := (U_{n,v})_{v \in \mathbb{Z}^d} \in [0, \infty)^{\mathbb{Z}^d}$  denote the configuration of the urn at time  $n$ , that is,

$$\mathbb{P} \left( (n+1)^{\text{th}} \text{ selected ball has color } v \mid U_n, U_{n-1}, \dots, U_0 \right) \propto U_{n,v}, v \in \mathbb{Z}^d.$$

Starting with  $U_0$  which is a probability distribution we define  $(U_n)_{n \geq 0}$  recursively as follows

$$(2) \quad U_{n+1} = U_n + \chi_{n+1} R$$

where  $\chi_{n+1} = (\chi_{n+1,v})_{v \in \mathbb{Z}^d}$  is such that  $\chi_{n+1, Z_n} = 1$  and  $\chi_{n+1,u} = 0$  if  $u \neq Z_n$ , where  $Z_n$  is the random color chosen from the configuration  $U_n$ . In other words,

$$U_{n+1} = U_n + R_{Z_n}$$

where  $R_{Z_n}$  is the  $Z_n$ -th row of the replacement matrix  $R$ . We will call the process  $(U_n)_{n \geq 0}$  as the *infinite color urn model* with initial configuration  $U_0$  and replacement matrix  $R$ . We will also refer to it as the *infinite color urn model associated with the random walk  $\{S_n\}_{n \geq 0}$  on  $\mathbb{Z}^d$* . Throughout this paper, we will assume  $U_0$  to be such that there exists  $r > 0$ , with

$$(3) \quad \sum_{v \in \mathbb{Z}^d} \exp(\langle \lambda, v \rangle) U_{0,v} < \infty,$$

whenever  $\|\lambda\| < r$ . In particular, (3) is satisfied when  $U_{0,v} = 0$ , for all but finitely many  $v \in \mathbb{Z}^d$ .

**Random configuration of the urn:** Observe that from definition

$$(4) \quad \sum_{v \in \mathbb{Z}^d} U_{n,v} = n + 1,$$

for all  $n \geq 0$ . Thus the *random configuration of the urn*, namely,  $\frac{U_n}{n+1}$  is a probability mass function on the set of colors. In fact,

$$(5) \quad \mathbb{P} \left( Z_n = v \mid U_n, U_{n-1}, \dots, U_0 \right) = \frac{U_{n,v}}{n+1}, v \in \mathbb{Z}^d.$$

In other words, the  $n$ -th random configuration of the urn is the conditional distribution of the  $(n+1)$ -th selected color, given  $U_0, U_1, \dots, U_n$ .

**Expected configuration of the urn:** As the entries of  $U_n$  are non-negative so by taking expectation in equation (4), we get

$$(6) \quad \sum_{v \in \mathbb{Z}^d} \mathbb{E}[U_{n,v}] = n + 1,$$

for all  $n \geq 0$ . Thus  $\left( \frac{\mathbb{E}[U_{n,v}]}{n+1} \right)_{v \in \mathbb{Z}^d}$  is also a probability mass function. In fact, it is the marginal distribution of  $Z_n$ , the  $(n+1)$ -th selected color. This follows by taking expectation on both sides of equation (5),

$$(7) \quad \mathbb{P}(Z_n = v) = \frac{\mathbb{E}[U_{n,v}]}{n+1}.$$

For the rest of this work, we will be interested in the asymptotic properties of the random and expected configurations.

**1.3. Notations.** Most of the notations used in this paper are consistent with the literature on finite color urn models. However, we use few specific notations as well, which are given below.

- (i) As mentioned earlier, all vectors are written as row vectors unless otherwise stated. Column vectors are denoted by  $x^T$ , where  $x$  is a row vector.
- (ii) For any vector  $x$ ,  $x^2$  will denote a vector with the coordinates squared.
- (iii) The standard Gaussian measure on  $\mathbb{R}^d$  will be denoted by  $\Phi_d$  with its density given by

$$\phi_d(x) := \frac{1}{(2\pi)^{d/2}} \exp\left(-\frac{\|x\|^2}{2}\right), x \in \mathbb{R}^d.$$

For  $d = 1$ , we will simply write  $\Phi$  for the standard Gaussian measure on  $\mathbb{R}$  and  $\phi$  for its density.

**1.4. Outline.** In the following section we state the main results, which we prove in Section 4. In Section 3, we state and prove two important results, which we use in the proofs of the main results. In Section 5, we further generalize our results for urns with infinitely many colors, where the color sets are indexed by other countable lattices on  $\mathbb{R}^d$ . An elementary technical result which is needed in the proofs of the main results is deferred to the appendix.

## 2. MAIN RESULTS

Throughout this paper we assume that  $(\Omega, \mathcal{F}, \mathbb{P})$  is a probability space on which all the random processes are defined.

**2.1. Weak Convergence of the Expected Configuration.** We present in this subsection the central limit theorem for the randomly selected color. The centering and scaling of the central limit theorem are of the order  $\mathcal{O}(\log n)$  and  $\mathcal{O}(\sqrt{\log n})$  respectively. Such centering and scalings are available because the marginal distribution of the randomly selected color behaves like that of a delayed random walk, where the delay is of the order  $\mathcal{O}(\log n)$ , see Theorem 7.

**Theorem 1.** *Let  $\bar{\Lambda}_n$  be the probability measure on  $\mathbb{R}^d$  corresponding to the probability vector  $\frac{1}{n+1} (\mathbb{E}[U_{n,v}])_{v \in \mathbb{Z}^d}$ . For any Borel subset  $A$  of  $\mathbb{R}^d$  and  $x \in \mathbb{R}$ , we define*

$$xA\Sigma^{1/2} := \{xy\Sigma^{1/2} : y \in A\},$$

and

$$\bar{\Lambda}_n^{cs}(A) := \bar{\Lambda}_n\left(\sqrt{\log n}xA\Sigma^{1/2} + \mu \log n\right).$$

Then, as  $n \rightarrow \infty$ ,

$$(8) \quad \bar{\Lambda}_n^{cs} \Rightarrow \Phi_d.$$

Recall that if  $Z_n$  denotes the  $(n+1)$ -th selected color then its probability mass function is given by  $\left(\frac{\mathbb{E}[U_{n,v}]}{n+1}\right)_{v \in \mathbb{Z}^d}$ . Thus  $\bar{\Lambda}_n$  is the probability distribution of  $Z_n$ . So (8) is equivalent to the following

$$(9) \quad \frac{Z_n - \mu \log n}{\sqrt{\log n}} \Rightarrow N_d(0, \Sigma), \text{ as } n \rightarrow \infty.$$

The following result is an immediate application of Theorem 1.

**Corollary 2.** *Consider the urn model associated with the simple symmetric random walk on  $\mathbb{Z}^d$ ,  $d \geq 1$ . Then, as  $n \rightarrow \infty$ ,*

$$\frac{Z_n}{\sqrt{\log n}} \Rightarrow N_d(0, d^{-1}\mathbb{I}_d),$$

where  $\mathbb{I}_d$  is the  $d \times d$  identity matrix.

The above result essentially shows that irrespective of the recurrent or transient behavior of the underlying random walk, the associated urn models have similar asymptotic behavior. In particular, the limiting distribution is always Gaussian with universal orders for centering and scaling, namely,  $\mathcal{O}(\log n)$  and  $\mathcal{O}(\sqrt{\log n})$  respectively.

**2.2. Weak Convergence of the Random Configuration.** In this subsection we will present an asymptotic result for the random configuration of the urn. Let  $\mathcal{M}_1$  be the space of probability measures on  $\mathbb{R}^d$ ,  $d \geq 1$ , endowed with the topology of weak convergence. Let  $\Lambda_n$  be the random probability measure on  $\mathbb{Z}^d \subset \mathbb{R}^d$  corresponding to the random probability vector  $\frac{U_n}{n+1}$ . It is easy to see that the function  $\Lambda_n$  is measurable.

**Theorem 3.** *Let*

$$\Lambda_n^{cs}(A) = \Lambda_n\left(\sqrt{\log n}A\Sigma^{1/2} + \mu \log n\right).$$

Then, as  $n \rightarrow \infty$ ,

$$(10) \quad \Lambda_n^{cs} \xrightarrow{P} \Phi_d \text{ in } \mathcal{M}_1.$$

We note that Theorem 3 is a stronger version of Theorem 1.

**2.3. Local Limit Theorem Type Results for the Expected Configuration.**

It turns out that under certain assumptions the expected configuration of the urn at time  $n$ , namely,  $\left(\frac{\mathbb{E}[U_n]}{n+1}\right)_{n \geq 0}$  satisfies a *local limit theorem*.

Note that  $X_1$  is a lattice random vector taking values in  $\mathbb{Z}^d$ . Let  $\mathcal{L}$  be its *minimal lattice*, that is,

$$(11) \quad \mathbb{P}(X_1 \in x + \mathcal{L}) = 1$$

for every  $x \in \mathbb{Z}^d$ , such that,  $\mathbb{P}(X_1 = x) > 0$ , and if  $\mathcal{L}'$  is any closed subgroup of  $\mathbb{R}^d$ , with  $\mathbb{P}(X_1 \in y + \mathcal{L}') = 1$  for some  $y \in \mathbb{Z}^d$ , then  $\mathcal{L} \subseteq \mathcal{L}'$  and the rank of  $\mathcal{L}$  is  $d$ . We refer to pages 226 – 227 of [4] for formal definitions of minimal lattice of a  $d$ -dimensional lattice random variable and its rank. Let  $l = \det(\mathcal{L})$  (see pages 228 – 229 of [4] for more details). Now let  $x_0$  be such that  $\mathbb{P}(X_1 \in x_0 + \mathcal{L}) = 1$  and we define

$$(12) \quad \mathcal{L}_n^{(d)} := \left\{ x: x = \frac{n}{\sqrt{\log n}}x_0\Sigma^{-1/2} - \sqrt{\log n}\mu\Sigma^{-1/2} + \frac{1}{\sqrt{\log n}}z\Sigma^{-1/2}, z \in \mathcal{L} \right\}.$$

**Theorem 4.** *Assume that  $\mathbb{P}(X_1 = 0) > 0$ . Then, as  $n \rightarrow \infty$*

$$(13) \quad \sup_{x \in \mathcal{L}_n^{(d)}} \left| \frac{\det(\Sigma^{1/2})(\sqrt{\log n})^d}{l} \mathbb{P}\left(\frac{Z_n - \mu \log n}{\sqrt{\log n}}\Sigma^{-1/2} = x\right) - \phi_d(x) \right| \rightarrow 0.$$

For  $d = 1$ , it is well known that (11) can be explicitly written as

$$(14) \quad \mathbb{P}(X_1 \in a + h\mathbb{Z}) = 1,$$

where  $a \in \mathbb{R}$ , and  $h > 0$ , is the maximum value such that (14) holds.  $h$  is called the span for  $X_1$  (see Section 3.5 of [22]). Then from (12) we obtain

$$(15) \quad \mathcal{L}_n^{(1)} := \left\{ x : x = \frac{n}{\sigma\sqrt{\log n}}a - \frac{\mu}{\sigma}\sqrt{\log n} + \frac{h}{\sigma\sqrt{\log n}}z, z \in \mathbb{Z} \right\}.$$

Therefore, in one dimension, (13) can be stated as

$$(16) \quad \sup_{x \in \mathcal{L}_n^{(1)}} \left| \sigma \frac{\sqrt{\log n}}{h} \mathbb{P} \left( \frac{Z_n - \mu \log n}{\sigma\sqrt{\log n}} = x \right) - \phi(x) \right| \rightarrow 0.$$

Observe that Theorem 4 covers only the case when  $\mathbb{P}(X_1 = 0) > 0$ . The next theorem is for the special case when the urn is associated with the simple symmetric random walk on  $\mathbb{Z}^d$ ,  $d \geq 1$ , which is not covered by Theorem 4.

**Theorem 5.** *Assume that  $\mathbb{P}(X_1 = \pm e_i) = \frac{1}{2d}$  for  $1 \leq i \leq d$ , where  $e_i$  is the  $i$ -th unit vector in direction  $i$ . Then, as  $n \rightarrow \infty$*

$$(17) \quad \sup_{x \in \mathcal{L}_n^{(d)}} \left| (d)^{\frac{d}{2}} \left( \sqrt{\log n} \right)^d \mathbb{P} \left( \frac{\sqrt{d}}{\sqrt{\log n}} Z_n = x \right) - \phi_d(x) \right| \rightarrow 0,$$

where  $\mathcal{L}_n^{(d)}$  is as defined in (12) with  $\mu = 0 = x_0$ ,  $\Sigma = \mathbb{I}_d$  and  $\mathcal{L} = \sqrt{d}\mathbb{Z}^d$ .

The following result is immediate from the above theorem.

**Corollary 6.** *Assume that  $\mathbb{P}(X_1 = \pm e_i) = \frac{1}{2d}$  for  $1 \leq i \leq d$ , where  $e_i$  is the  $i$ -th unit vector in direction  $i$ . Then, as  $n \rightarrow \infty$*

$$(18) \quad \mathbb{P}(Z_n = 0) \sim \frac{1}{(\sqrt{2\pi d \log n})^d}.$$

As we see in Theorem 5, the assumption  $\mathbb{P}(X_1 = 0) > 0$  of Theorem 4 can be removed. We present Theorem 10 in Section 4.3, which though only in dimension  $d = 1$ , is another such example. As we will see later that Theorem 10 is slightly technical. Unfortunately, we do not know the full generality under which the local limit theorem holds, though we conjecture that it holds for all cases.

### 3. AUXILIARY RESULTS

In this section, we present two results which we need to prove our main results. These results are two very important tools for studying infinite color urn models associated with random walks on  $\mathbb{Z}^d$  and hence presented separately.

For  $z \in \mathbb{C}$ , define  $\Pi_n(z) = \prod_{j=1}^n \left( 1 + \frac{z}{j} \right)$ , for  $n \geq 1$ , and set  $\Pi_0(z) = 1$ . It is

known from Euler product formula for gamma function, which is also referred to as Gauss's formula (see page 178 of [16]), that

$$(19) \quad \lim_{n \rightarrow \infty} \frac{\Pi_n(z)}{n^z} \Gamma(z+1) = 1$$

uniformly on compact subsets of  $\mathbb{C} \setminus \{-1, -2, -3, \dots\}$ .

Recall  $e(\lambda) := \sum_{v \in B} \exp(\langle \lambda, v \rangle) p(v)$  is the moment generating function of  $X_1$ . It is easy to note that  $e(\lambda)$  is an eigenvalue of  $R$  corresponding to the right eigenvector  $x(\lambda) = (\exp(\langle \lambda, v \rangle))_{v \in \mathbb{Z}^d}^T$ . Let  $\mathcal{F}_n = \sigma(U_j : 0 \leq j \leq n)$ ,  $n \geq 0$  be the natural filtration. Define

$$(20) \quad \overline{M}_n(\lambda) = \frac{U_n x(\lambda)}{\Pi_n(e(\lambda))}$$

From the fundamental recursion (2) we get,

$$(21) \quad \begin{aligned} U_{n+1}x(\lambda) &= U_nx(\lambda) + \chi_{n+1}Rx(\lambda) \\ &= U_nx(\lambda) + e(\lambda)\chi_{n+1}x(\lambda) \end{aligned}$$

Thus,

$$\mathbb{E} \left[ U_{n+1}x(\lambda) \middle| \mathcal{F}_n \right] = U_nx(\lambda) + e(\lambda) \mathbb{E} \left[ \chi_{n+1}x(\lambda) \middle| \mathcal{F}_n \right] = \left( 1 + \frac{e(\lambda)}{n+1} \right) U_nx(\lambda).$$

Therefore,  $\overline{M}_n(\lambda)$  is a non-negative martingale for every  $\lambda \in \mathbb{R}^d$ . In particular  $\mathbb{E} [\overline{M}_n(\lambda)] = \overline{M}_0(\lambda)$ .

**3.1. A Marginal Representation.** We now present a representation of the marginal distribution of  $Z_n$  in terms of the increments  $(X_j)_{j \geq 1}$ . This representation is interesting and non-trivial, as it necessarily demonstrates that the marginal distribution of the randomly selected color behaves like a delayed random walk.

**Theorem 7.** *For each  $n \geq 1$ ,*

$$(22) \quad Z_n \stackrel{d}{=} Z_0 + \sum_{j=1}^n I_j X_j,$$

where  $\{I_j\}_{j \geq 1}$  are independent Bernoulli random variables such that  $\mathbb{E}[I_j] = \frac{1}{j+1}$ ,  $j \geq 1$  and are independent of  $\{X_j\}_{j \geq 1}$ ; and  $Z_0$  is a random vector taking values in  $\mathbb{Z}^d$  distributed according to the probability vector  $U_0$  and is independent of  $(\{I_j\}_{j \geq 1}; \{X_j\}_{j \geq 1})$ .

A formal proof of this result is given below, using the martingale sequence  $(\overline{M}_n(\lambda))_{n \geq 0}$ , as defined in (20). However, one can also establish the representation given in (22) intuitively. Let  $\mathcal{T}_{n+1}$  be the *random recursive tree* on  $n+1$  vertices, say,  $\{v_0, v_1, v_2, \dots, v_n\}$  [21, 39]. We put i.i.d. weights on each of the  $n$  edges of  $\mathcal{T}_{n+1}$  given by  $\{X_j\}_{j \geq 1}$ , as defined in Theorem 7. Let  $Z_0$  with distribution given by  $U_0$  be as defined in Theorem 7. Then from the equations (2) and (4), it follows that the distribution of  $Z_n$  is given by

$$Z_0 + \text{weighted length of the path from } v_0 \text{ to } v_n \text{ in } \mathcal{T}_{n+1}.$$

This shows (22) must hold.

It may be noted here that the above representation is similar to a discrete time *branching random walk (BRW)*, with i.i.d. increments given by  $\{X_j\}_{j \geq 1}$ , but on the random recursive tree  $(\mathcal{T}_n)_{n \geq 1}$ . Unlike the usual BRW [5–7], in this case the underlying tree is not a *Galton-Watson branching process tree*. Also as described above, in this case we are interested in studying the position of a randomly chosen particle, instead of the maximum displacement.

We would like to point out here that our approach of coupling with underlying Markov chain as in the above theorem is entirely new. It helps us to completely bypass the technical difficulties which one may face in using the eigenvalue techniques in the infinite color case.

*Proof.* As noted before, the probability mass function for the color of the  $(n+1)$ -th selected ball, namely  $Z_n$ , is  $\left(\frac{\mathbb{E}[U_{n,v}]}{n+1}\right)_{v \in \mathbb{Z}^d}$ . So for  $\lambda \in \mathbb{R}^d$  with  $\|\lambda\| < r$ , where  $r$  is as in (3), the moment generating function of  $Z_n$  is given by

$$\begin{aligned}
\frac{1}{n+1} \sum_{v \in \mathbb{Z}^d} \exp(\langle \lambda, v \rangle) \mathbb{E}[U_{n,v}] &= \frac{\Pi_n(e(\lambda))}{n+1} \mathbb{E}[\overline{M}_n(\lambda)] \\
&= \frac{\Pi_n(e(\lambda))}{n+1} \overline{M}_0(\lambda) \\
(23) \qquad \qquad \qquad &= \overline{M}_0(\lambda) \prod_{j=1}^n \left(1 - \frac{1}{j+1} + \frac{e(\lambda)}{j+1}\right).
\end{aligned}$$

The equation (22) follows from (23).  $\square$

Observe that in the proof of Theorem 7 we require the assumption (3), which guarantees that  $\overline{M}_0(\lambda)$  is finite.

**3.2. Uniform  $\mathcal{L}_2$  boundedness.** Our next theorem states that around a non-trivial closed neighborhood of 0 the martingales  $(\overline{M}_n(\lambda))_{n \geq 0}$  are uniformly (in  $\lambda$ )  $\mathcal{L}_2$  bounded.

**Theorem 8.** *There exists  $\delta > 0$  such that*

$$(24) \qquad \sup_{\lambda \in [-\delta, \delta]^d} \sup_{n \geq 1} \mathbb{E}[\overline{M}_n^2(\lambda)] < \infty.$$

*Proof.* From (21) we obtain

$$\begin{aligned}
\mathbb{E}[(U_{n+1}x(\lambda))^2 | \mathcal{F}_n] &= (U_nx(\lambda))^2 + 2e(\lambda) U_nx(\lambda) \mathbb{E}[\chi_{n+1}x(\lambda) | \mathcal{F}_n] \\
&\quad + e^2(\lambda) \mathbb{E}[(\chi_{n+1}x(\lambda))^2 | \mathcal{F}_n]
\end{aligned}$$

It is easy to see that

$$(25) \qquad \mathbb{E}[\chi_{n+1}x(\lambda) | \mathcal{F}_n] = \frac{1}{n+1} U_nx(\lambda) \quad \text{and} \quad \mathbb{E}[(\chi_{n+1}x(\lambda))^2 | \mathcal{F}_n] = \frac{1}{n+1} U_nx(2\lambda).$$

Therefore, we get the recursion

$$\begin{aligned}
\mathbb{E}[(U_{n+1}x(\lambda))^2] &= \left(1 + \frac{2e(\lambda)}{n+1}\right) \mathbb{E}[(U_nx(\lambda))^2] \\
(26) \qquad \qquad \qquad &\quad + \frac{e^2(\lambda)}{n+1} \mathbb{E}[U_nx(2\lambda)].
\end{aligned}$$

Dividing both sides of (26) by  $\Pi_{n+1}^2(e(\lambda))$ ,

$$(27) \qquad \mathbb{E}[\overline{M}_{n+1}^2(\lambda)] = \frac{\left(1 + \frac{2e(\lambda)}{n+1}\right)}{\left(1 + \frac{e(\lambda)}{n+1}\right)^2} \mathbb{E}[\overline{M}_n^2(\lambda)] + \frac{e^2(\lambda) \mathbb{E}[U_nx(2\lambda)]}{n+1 \Pi_{n+1}^2(e(\lambda))}.$$



Recall that  $\overline{M}_n(2\lambda)$  is a martingale, thus  $\mathbb{E}[U_n x(2\lambda)] = \Pi_n(e(2\lambda)) \overline{M}_0(2\lambda)$ . Therefore from (27), we get

$$(28) \quad \mathbb{E}[\overline{M}_n^2(\lambda)] = \frac{\Pi_n(2e(\lambda))}{\Pi_n(e(\lambda))^2} \overline{M}_0^2(\lambda) + \sum_{k=1}^n \frac{e^2(\lambda)}{k} \left\{ \prod_{j>k}^n \frac{\left(1 + \frac{2e(\lambda)}{j}\right)}{\left(1 + \frac{e(\lambda)}{j}\right)^2} \right\} \frac{\Pi_{k-1}(e(2\lambda))}{\Pi_k^2(e(\lambda))} \overline{M}_0(2\lambda).$$

We observe that as  $e(\lambda) > 0$ , so  $\frac{1 + \frac{2e(\lambda)}{j}}{\left(1 + \frac{e(\lambda)}{j}\right)^2} < 1$  and hence  $\frac{\Pi_n(2e(\lambda))}{\Pi_n^2(e(\lambda))} < 1$ . Thus

$$(29) \quad \mathbb{E}[\overline{M}_n^2(\lambda)] \leq \overline{M}_0^2(\lambda) + e^2(\lambda) \overline{M}_0(2\lambda) \sum_{k=1}^n \frac{1}{k} \frac{\Pi_{k-1}(e(2\lambda))}{\Pi_k^2(e(\lambda))}.$$

Using (19), we know that

$$(30) \quad \Pi_n^2(e(\lambda)) \sim \frac{n^{2e(\lambda)}}{\Gamma^2(e(\lambda) + 1)}.$$

Note that  $e(0) = 1$  and  $e(\lambda)$  is continuous as a function of  $\lambda$ . So given  $\eta > 0$ , there exists  $0 < K_1, K_2 < \infty$ , such that for all  $\lambda \in [-\eta, \eta]^d$ ,  $K_1 \leq e(\lambda) \leq K_2$ . Since the convergence in (19) is uniform on compact subsets of  $[0, \infty)$ , given  $\epsilon > 0$  there exists  $N_1 > 0$  such that for all  $n \geq N_1$  and  $\lambda \in [-\eta, \eta]^d$ ,

$$\begin{aligned} & (1 - \epsilon) \frac{\Gamma^2(e(\lambda) + 1)}{\Gamma(e(2\lambda) + 1)} \sum_{k \geq N_1}^n \frac{1}{k^{1+2e(\lambda)-e(2\lambda)}} \\ & \leq \sum_{k \geq N_1}^n \frac{1}{k} \frac{\Pi_{k-1}(e(2\lambda))}{\Pi_k^2(e(\lambda))} \\ & \leq (1 + \epsilon) \frac{\Gamma^2(e(\lambda) + 1)}{\Gamma(e(2\lambda) + 1)} \sum_{k \geq N_1}^n \frac{1}{k^{1+2e(\lambda)-e(2\lambda)}}. \end{aligned}$$

Recall that  $e(\lambda) = \sum_{v \in B} \exp(\langle \lambda, v \rangle) p(v)$ . Since the cardinality of  $B$  is finite, we can choose a  $\delta_0 > 0$  such that for every  $\lambda \in [-\delta_0, \delta_0]^d$ ,  $2e(\lambda) - e(2\lambda) > 0$ . Choose  $\delta = \min\{\eta, \delta_0\}$ . Since  $2e(\lambda) - e(2\lambda)$  is continuous as a function of  $\lambda$ , there exists a  $\lambda_0 \in [-\delta, \delta]^d$  such that  $\min_{\lambda \in [-\delta, \delta]^d} 2e(\lambda) - e(2\lambda) = 2e(\lambda_0) - e(2\lambda_0) > 0$ . Therefore,

$$\sum_{k=1}^{\infty} \frac{1}{k^{1+2e(\lambda)-e(2\lambda)}} \leq \sum_{k=1}^{\infty} \frac{1}{k^{1+2e(\lambda_0)-e(2\lambda_0)}}.$$

Therefore given  $\epsilon > 0$  there exists  $N_2 > 0$  such that  $\forall \lambda \in [-\delta, \delta]^d$

$$\sum_{k > N_2}^{\infty} \frac{1}{k^{1+2e(\lambda)-e(2\lambda)}} \leq \sum_{k > N_2}^{\infty} \frac{1}{k^{1+2e(\lambda_0)-e(2\lambda_0)}} < \epsilon.$$

$\frac{\Gamma^2(e(\lambda)+1)}{\Gamma(e(2\lambda)+1)}$ ,  $e^2(\lambda)$  and  $\overline{M}_0(2\lambda)$  being continuous as functions of  $\lambda$ , they are bounded for  $\lambda \in [-\delta, \delta]^d$ . Choose  $N = \max\{N_1, N_2\}$ . From (29) we obtain for all  $n \geq N$ ,

$$(31) \quad \mathbb{E} \left[ \overline{M}_n^2(\lambda) \right] \leq \overline{M}_0^2(\lambda) + C_1 \sum_{k=1}^N \frac{1}{k} \frac{\Pi_{k-1}(e(2\lambda))}{\Pi_k^2(e(\lambda))} + C_2 \epsilon$$

for an appropriate positive constants  $C_1, C_2$ .

$\sum_{k=1}^N \frac{1}{k} \frac{\Pi_{k-1}(e(2\lambda))}{\Pi_k^2(e(\lambda))}$  and  $\overline{M}_0^2(\lambda)$  being continuous as functions of  $\lambda$ , they are bounded for  $\lambda \in [-\delta, \delta]^d$ . Therefore, from (31) we obtain that there exists  $C > 0$  such that for all  $\lambda \in [-\delta, \delta]^d$  and for all  $n \geq 1$

$$\mathbb{E} \left[ \overline{M}_n^2(\lambda) \right] \leq C.$$

This proves (24). □

#### 4. PROOFS OF THE MAIN RESULTS

For proving the central and local limit theorems, we will use the representation (22), which after the appropriate centering and scaling can be re-written as

$$(32) \quad \frac{Z_n - \mu \log n}{\sqrt{\log n}} \stackrel{d}{=} \frac{Z_0}{\sqrt{\log n}} + \frac{\sum_{j=1}^n I_j X_j - \mu \log n}{\sqrt{\log n}}.$$

So, without loss of generality, we may assume that  $Z_0 \equiv 0$ , that is the initial configuration of the urn consists of only one ball of color 0.

##### 4.1. Proofs for the Expected Configuration.

*Proof of Theorem 1.* Observe that

$$(33) \quad \mathbb{E} \left[ \sum_{j=1}^n I_j X_j \right] - \mu \log n = \sum_{j=1}^n \frac{1}{j+1} \mu - \mu \log n \longrightarrow \mu(\gamma - 1),$$

where  $\gamma$  is the Euler's constant.

**Case I:** Let  $d = 1$ . Let  $s_n^2 = \text{Var} \left( \sum_{j=1}^n I_j X_j \right)$ . It is easy to note that

$$s_n^2 = \sum_{j=1}^n \frac{1}{j+1} \mathbb{E} [X_1^2] - \frac{\mu^2}{(j+1)^2} \sim \sigma^2 \log n.$$

Recall  $\mathbb{P}(X_1 \in B) = 1$ , where  $B \subset \mathbb{Z}$  is a finite set, thus for any  $\epsilon > 0$ , we have

$$\frac{1}{s_n^2} \sum_{j=1}^n \mathbb{E} [I_j X_j^2 1_{\{|I_j X_j| > \epsilon s_n\}}] \longrightarrow 0$$

as  $n \rightarrow \infty$ . Therefore, by the *Lindeberg Central Limit Theorem* [22], we conclude that as  $n \rightarrow \infty$

$$\frac{Z_n - \mu \log n}{\sigma \sqrt{\log n}} \Rightarrow N(0, 1).$$

This completes the proof in this case.

**Case II:** Now suppose  $d \geq 2$ . Let  $\Sigma_n$  denote the variance-covariance matrix for  $\sum_{j=1}^n I_j X_j$ . Then by calculations similar to that in one-dimension it is easy to see that as  $n \rightarrow \infty$

$$(34) \quad \frac{1}{\log n} \Sigma_n \longrightarrow \Sigma,$$

where the matrix convergence is entry-wise. Further, similar to one-dimension, as  $\mathbb{P}(X_1 \in B) = 1$ , where  $B \subset \mathbb{Z}^d$  is a finite set,

$$(35) \quad \frac{1}{\log n} \sum_{j=1}^n \mathbb{E} [\|I_j X_j\|^2 1_{\{\|I_j X_j\| > \epsilon \log n\}}] \longrightarrow 0,$$

for any  $\epsilon > 0$ . The result then follows from a *multidimensional Lindberg-Feller Central Limit Theorem* (see Proposition 2.27 on page 20 of [40]).  $\square$

**4.2. Proofs for Random Configuration.** In this subsection we will present the proof of Theorem 3. We start with the following lemma which is needed in the proof of Theorem 3.

**Lemma 9.** *Let  $\delta$  be as in Theorem 8, then for every  $\lambda \in [-\delta, \delta]^d$  as  $n \rightarrow \infty$ ,*

$$(36) \quad \overline{M}_n \left( \frac{\lambda}{\sqrt{\log n}} \right) \xrightarrow{p} 1.$$

*Proof.* Observe that  $\overline{M}_0(\lambda) \equiv 1$ , since  $Z_0 \equiv 0$ . Therefore, from equation (28) we get

$$\mathbb{E} [\overline{M}_n^2(\lambda)] = \frac{\Pi_n(2e(\lambda))}{\Pi_n^2(e(\lambda))} + \frac{\Pi_n(2e(\lambda))}{\Pi_n^2(e(\lambda))} \sum_{k=1}^n \frac{e^2(\lambda)}{k} \frac{\Pi_{k-1}(e(2\lambda))}{\Pi_k(2e(\lambda))}.$$

Replacing  $\lambda$  by  $\lambda_n = \frac{\lambda}{\sqrt{\log n}}$ , we obtain

$$(37) \quad \mathbb{E} [\overline{M}_n^2(\lambda_n)] = \frac{\Pi_n(2e(\lambda_n))}{\Pi_n^2(e(\lambda_n))} + \frac{\Pi_n(2e(\lambda_n))}{\Pi_n^2(e(\lambda_n))} \sum_{k=1}^n \frac{e^2(\lambda_n)}{k} \frac{\Pi_{k-1}(e(2\lambda_n))}{\Pi_k(2e(\lambda_n))}$$

We observe that

$$(38) \quad \lim_{n \rightarrow \infty} e(\lambda_n) = 1.$$

Since the convergence in formula (19) is uniform on compact sets of  $[0, \infty)$ , we observe that for  $\lambda \in [-\delta, \delta]^d$

$$\lim_{n \rightarrow \infty} \frac{\Pi_n(2e(\lambda_n))}{\Pi_n^2(e(\lambda_n))} = \frac{\Gamma^2(2)}{\Gamma(3)} = \frac{1}{2}.$$

Therefore,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\Pi_n(2e(\lambda_n))}{\Pi_n^2(e(\lambda_n))} \frac{e^2(\lambda_n)}{k} \frac{\Pi_{k-1}(e(2\lambda_n))}{\Pi_k(2e(\lambda_n))} &= \frac{1}{2} \frac{1}{k} \frac{\Pi_{k-1}(1)}{\Pi_k(2)} \\ &= \frac{1}{(k+2)(k+1)}. \end{aligned}$$

Now using Theorem 8 and the dominated convergence theorem, we get

$$\lim_{n \rightarrow \infty} \frac{\Pi_n(2e(\lambda_n))}{\Pi_n^2(e(\lambda_n))} \sum_{k=1}^n \frac{e^2(\lambda_n)}{k} \frac{\Pi_{k-1}(e(2\lambda_n))}{\Pi_k(2e(\lambda_n))} = \sum_{k=1}^{\infty} \frac{1}{(k+2)(k+1)} = \frac{1}{2}.$$

Therefore, from (37) we obtain

$$(39) \quad \mathbb{E} \left[ \overline{M}_n^2(\lambda_n) \right] \longrightarrow 1 \text{ as } n \rightarrow \infty.$$

Observing that  $\mathbb{E} \left[ \overline{M}_n(\lambda_n) \right] = 1$ , we get

$$(40) \quad \text{Var} \left( \overline{M}_n(\lambda_n) \right) \rightarrow 0,$$

as  $n \rightarrow \infty$ . This implies

$$\overline{M}_n(\lambda_n) \xrightarrow{p} 1 \text{ as } n \rightarrow \infty,$$

completing the proof of the lemma.  $\square$

*Proof of Theorem 3.* Note that  $\Lambda_n$  is the random probability measure on  $\mathbb{Z}^d \subset \mathbb{R}^d$ , corresponding to the random probability vector  $\frac{U_n}{n+1}$ . That is, for any Borel subset  $A$  of  $\mathbb{Z}^d$ ,

$$\Lambda_n(A) = \frac{1}{n+1} \sum_{v \in A} U_{n,v}.$$

For  $\lambda \in \mathbb{R}^d$  the corresponding moment generating function is given by

$$(41) \quad \frac{1}{n+1} \sum_{v \in \mathbb{Z}^d} \exp(\langle \lambda, v \rangle) U_{n,v} = \frac{1}{n+1} U_n x(\lambda) = \frac{1}{n+1} \overline{M}_n(\lambda) \Pi_n(e(\lambda)).$$

The moment generating function corresponding to the scaled and centered random measure  $\Lambda_n^{cs}$  is

$$\begin{aligned} & \sum_{v \in \mathbb{Z}^d} \exp\left(\left\langle \lambda, \frac{v - \mu \log n}{\sqrt{\log n}} \Sigma^{-1/2} \right\rangle\right) \frac{U_{n,v}}{n+1} \\ &= \frac{1}{n+1} \exp\left(-\langle \lambda, \mu \sqrt{\log n} \Sigma^{-1/2} \rangle\right) U_n x\left(\frac{\lambda \Sigma^{-1/2}}{\sqrt{\log n}}\right) \\ &= \frac{1}{n+1} \exp\left(-\langle \lambda, \mu \sqrt{\log n} \Sigma^{-1/2} \rangle\right) \overline{M}_n\left(\frac{\lambda \Sigma^{-1/2}}{\sqrt{\log n}}\right) \Pi_n\left(e\left(\frac{\lambda \Sigma^{-1/2}}{\sqrt{\log n}}\right)\right). \end{aligned}$$

To show (10) it is enough to show that for every subsequence  $\{n_k\}_{k \geq 1}$ , there exists a further subsequence  $\{n_{k_j}\}_{j=1}^\infty$  such that as  $j \rightarrow \infty$

$$(42) \quad \frac{\exp(-\langle \lambda, \mu \sqrt{\log n_{k_j}} \rangle)}{n_{k_j} + 1} \overline{M}_{n_{k_j}}\left(\frac{\lambda}{\sqrt{\log n_{k_j}}}\right) \Pi_n\left(e\left(\frac{\lambda}{\sqrt{\log n_{k_j}}}\right)\right) \longrightarrow \exp\left(\frac{\lambda \Sigma \lambda^T}{2}\right)$$

for all  $\lambda \in [-\delta, \delta]^d$  almost surely, where  $\delta$  is as in Theorem 8. From Theorem 1 we know that

$$\frac{Z_n - \mu \log n}{\sqrt{\log n}} \Rightarrow N_d(0, \mathbb{I}_d).$$

Therefore using (23) as  $n \rightarrow \infty$  we obtain,

$$\begin{aligned} & \exp\left(-\langle \lambda, \mu \sqrt{\log n} \rangle\right) \mathbb{E} \left[ \exp\left(\left\langle \lambda, \frac{Z_n}{\sqrt{\log n}} \right\rangle\right) \right] \\ &= \frac{1}{n+1} \exp\left(-\langle \lambda, \mu \sqrt{\log n} \rangle\right) \Pi_n\left(e\left(\frac{\lambda}{\sqrt{\log n}}\right)\right) \longrightarrow \exp\left(\frac{\lambda \Sigma \lambda^T}{2}\right). \end{aligned}$$

Now using Theorem 14 from the appendix it is enough to show (42) only for  $\lambda \in \mathbb{Q}^d \cap [-\delta, \delta]^d$  which is equivalent to proving that for every  $\lambda \in \mathbb{Q}^d \cap [-\delta, \delta]^d$  as  $j \rightarrow \infty$

$$\overline{M}_{n_{k_j}} \left( \frac{\lambda}{\sqrt{\log n_{k_j}}} \right) \rightarrow 1 \text{ almost surely.}$$

From Lemma 9 we know that for all  $\lambda \in [-\delta, \delta]^d$

$$\overline{M}_n \left( \frac{\lambda}{\sqrt{\log n}} \right) \xrightarrow{p} 1 \text{ as } n \rightarrow \infty.$$

Therefore using the standard diagonalization argument we can say that given a subsequence  $\{n_k\}_{k \geq 1}$  there exists a further subsequence  $\{n_{k_j}\}_{j=1}^\infty$  such that for every  $\lambda \in \mathbb{Q}^d \cap [-\delta, \delta]^d$

$$\overline{M}_{n_{k_j}} \left( \frac{\lambda}{\sqrt{\log n_{k_j}}} \right) \rightarrow 1 \text{ almost surely.}$$

This completes the proof.  $\square$

**4.3. Proofs of the Local Limit Type Results.** In this section, we present the proofs for the local limit theorems.

*Proof of Theorem 4.* Since without loss we have assumed  $U_0 = \delta_0$ , so from Theorem 7, we obtain  $Z_n \stackrel{d}{=} \sum_{j=1}^n I_j X_j$ . The random vector  $X_j$  is a lattice random vector. Therefore,  $I_j X_j$  is also a lattice random vector. By our assumption  $\mathbb{P}(X_1 = 0) > 0$ , so  $0 \in B$ , therefore,  $X_j$  and  $I_j X_j$  are supported on the same lattice.

Observe that  $Z_n$  is a lattice random vector, for every  $n \in \mathbb{N}$ . By Fourier inversion formula (see 21.28 on page 230 of [4]), we get for  $x \in \mathcal{L}_n^{(d)}$ ,

$$\mathbb{P} \left( \frac{Z_n - \mu \log n}{\sqrt{\log n}} \Sigma^{-1/2} = x \right) = \frac{l}{(2\pi \sqrt{\log n})^d \det(\Sigma^{1/2})} \int_{(\sqrt{\log n} \mathcal{F}^* \Sigma^{1/2})} \psi_n(t) \exp(-i\langle t, x \rangle) dt$$

where  $\psi_n(t) = \mathbb{E} \left[ \exp \left( i \langle t, \frac{Z_n - \mu \log n}{\sqrt{\log n}} \Sigma^{-1/2} \rangle \right) \right]$ ,  $l = |\det(\mathcal{L})|$  and  $\mathcal{F}^*$  is the fundamental domain for  $X_1$  as defined in equation (21.22) on page 229 of [4]. Also, by Fourier inversion formula we have

$$\phi_d(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \exp(-i\langle t, x \rangle) \exp \left( -\frac{\|t\|^2}{2} \right) dt.$$

Thus, given  $\epsilon > 0$ , there exists  $N > 0$  such that for all  $n \geq N$ , we get the estimate

$$\begin{aligned} & \left| \frac{\det(\Sigma^{1/2}) (\sqrt{\log n})^d}{l} \mathbb{P} \left( \frac{Z_n - \mu \log n}{\sqrt{\log n}} \Sigma^{-1/2} = x \right) - \phi_d(x) \right| \\ & \leq \frac{1}{(2\pi)^d} \int_{(\sqrt{\log n} \mathcal{F}^* \Sigma^{1/2})} \left| \psi_n(t) - \exp \left( -\frac{\|t\|^2}{2} \right) \right| dt + \epsilon. \end{aligned}$$

Therefore, it is enough to prove that as  $n \rightarrow \infty$

$$(43) \quad \int_{(\sqrt{\log n} \mathcal{F}^* \Sigma^{1/2})} \left| \psi_n(t) - \exp \left( -\frac{\|t\|^2}{2} \right) \right| dt \rightarrow 0.$$

Now we follow an argument similar to the proof of a local limit theorem for sum of i.i.d. lattice random variables in one dimension, as given on page 190 of [37]. We observe that to prove (43), it is then enough to show that for any  $\epsilon > 0$ , we can choose a compact subset  $A \subseteq \mathbb{R}^d$ , such that for all  $n$  large enough

$$(44) \quad \mathcal{I}(n, A) := \int_{(\sqrt{\log n} \mathcal{F}^* \Sigma^{1/2}) \setminus A} |\psi_n(t)| dt < \epsilon.$$

Fix  $\epsilon > 0$ , we will show that (44) holds for a suitable choice of a compact subset  $A$  of  $\mathbb{R}^d$ . For that we first note that as,  $Z_n \stackrel{d}{=} \sum_{j=1}^n I_j X_j$ , so for every  $t \in \mathbb{R}^d$ , and each  $n$

$$|\psi_n(t)| \leq |g_n(t)|,$$

where

$$(45) \quad g_n(t) := \mathbb{E} \left[ \exp \left( \left\langle it, \frac{\sum_{j=1}^n I_j X_j}{\sqrt{\log n}} \Sigma^{-1/2} \right\rangle \right) \right].$$

Therefore, from definition of  $\mathcal{I}(n, A)$  we obtain

$$(46) \quad \mathcal{I}(n, A) \leq \int_{(\sqrt{\log n} \mathcal{F}^* \Sigma^{1/2}) \setminus A} |g_n(t)| dt.$$

Observe that

$$\begin{aligned} \mathbb{E} \left[ \exp \left( \left\langle it, \sum_{j=1}^n I_j X_j \right\rangle \right) \right] &= \prod_{j=1}^n \left( 1 - \frac{1}{j+1} + \frac{e(it)}{j+1} \right) \\ &= \frac{1}{n+1} \Pi_n(e(it)) \end{aligned}$$

where  $e(it) = \mathbb{E}[\exp(i\langle t, X_1 \rangle)]$ . So, from (45)

$$(47) \quad g_n(t) = \frac{1}{n+1} \Pi_n \left( e \left( \frac{1}{\sqrt{\log n}} it \Sigma^{-1/2} \right) \right).$$

Applying a change of variables  $t = \frac{1}{\sqrt{\log n}} w$ , we obtain

$$(48) \quad \mathcal{I}(n, A) \leq (\sqrt{\log n})^d \int_{\mathcal{F}^* \Sigma^{1/2} \setminus \frac{1}{\sqrt{\log n}} A} |g_n(\sqrt{\log n} w)| dw.$$

We can choose  $\delta > 0$ , such that for all  $w \in B(0, \delta) \setminus \{0\}$ , there exists  $b > 0$ , such that

$$(49) \quad |e(iw)| \leq 1 - \frac{b\|w\|^2}{2},$$

(see Lemma 2.3.2(a) of [32] for a proof). Therefore, using the inequality  $1 - x \leq \exp(-x)$ ,

$$\begin{aligned}
 |g_n(\sqrt{\log nw})| &= \frac{1}{n+1} |\Pi_n(e(iw\Sigma^{-1/2}))| \\
 &\leq \prod_{j=1}^{n+1} \left( 1 - \frac{1}{j+1} + \frac{|e(iw\Sigma^{-1/2})|}{j+1} \right) \\
 (50) \quad &\leq \exp\left(-\sum_{j=1}^n \frac{b}{j+1} \frac{w\Sigma w^T}{2}\right) \leq C_1 \exp\left(-b \frac{w\Sigma w^T}{2} \log n\right)
 \end{aligned}$$

for some positive constant  $C_1$ . We write

$$(\sqrt{\log n})^d \int_{\mathcal{F}^* \Sigma^{1/2} \setminus \frac{1}{\sqrt{\log n}} A} |g_n(\sqrt{\log nw})| dw = \mathcal{I}_1(n, A) + \mathcal{I}_2(n)$$

where

$$\mathcal{I}_1(n, A) := (\sqrt{\log n})^d \int_{(B(0, \delta) \Sigma^{1/2} \setminus \frac{1}{\sqrt{\log n}} A) \cap \mathcal{F}^* \Sigma^{1/2}} |g_n(\sqrt{\log nw})| dw$$

and

$$\mathcal{I}_2(n) = (\sqrt{\log n})^d \int_{\mathcal{F}^* \Sigma^{1/2} \setminus B(0, \delta) \Sigma^{1/2}} |g_n(\sqrt{\log nw})| dw.$$

Now it is enough to show that for a suitable choice of a compact subset  $A$  of  $\mathbb{R}^d$  and  $n$  sufficiently large,  $\mathcal{I}_1(n, A) < \epsilon$  and  $\mathcal{I}_2(n) \rightarrow 0$  as  $n \rightarrow \infty$ .

To this end, we first use (50) and obtain

$$(51) \quad \mathcal{I}_1(n) \leq C_1 (\sqrt{\log n})^d \int_{B(0, \delta) \Sigma^{1/2} \setminus \frac{A}{\sqrt{\log n}}} \exp\left(-b \frac{w\Sigma w^T}{2} \log n\right) dw,$$

for an appropriate positive constant  $C_1$ . Applying a change of variables  $w(\sqrt{\log n}) = t$ , we obtain

$$(52) \quad \mathcal{I}_1(n, A) \leq C_1 \int_{B(0, \delta \sqrt{\log n}) \Sigma^{1/2} \setminus A} \exp\left(-b \frac{t\Sigma t^T}{2}\right) dt.$$

Recall  $\epsilon > 0$  is fixed, so we now choose a compact  $A \subseteq \mathbb{R}^d$ , such that

$$(53) \quad C_1 \int_{\mathbb{R}^d \setminus A} \exp\left(-b \frac{t\Sigma t^T}{2}\right) dt < \epsilon.$$

Therefore, for this choice of  $A$ ,

$$(54) \quad \mathcal{I}_1(n, A) < \epsilon.$$

Since the lattices for  $X_1$  and  $I_1 X_1$  are same, for all  $w \in \mathcal{F}^* \setminus B(0, \delta)$ , we get  $|e(iw)| < 1$ , so there exists an  $0 < \eta < 1$ , such that,  $|e(iw\Sigma^{-1/2})| \leq \eta$ , for all  $w \in$

$\mathcal{F}^* \Sigma^{1/2} \setminus B(0, \delta) \Sigma^{1/2}$  Therefore, using the inequality  $1 - x \leq \exp(-x)$ , we obtain

$$(55) \quad |g_n(\sqrt{\log n} w)| \leq \exp \left( - \sum_{j=i}^n \frac{1}{j+1} (1-\eta) \right) \leq C_2 \exp(-(1-\eta) \log n)$$

for some positive constant  $C_2$ . Therefore, using equation (21.25) on page 230 of [4], we obtain

$$(56) \quad \mathcal{I}_2(n) \leq C'_2 (\sqrt{\log n})^d \exp(-(1-\eta) \log n) \rightarrow 0 \text{ as } n \rightarrow \infty$$

where  $C'_2$  is an appropriate positive constant. This completes the proof.  $\square$

*Proof of the Theorem 5.* In this case  $\mathbb{P}(X_1 = \pm e_i) = \frac{1}{2d}$  for  $1 \leq i \leq d$ , where  $e_i$  is the  $i$ -th unit vector in direction  $i$ . Thus  $\mu = 0$  and  $\Sigma = \frac{1}{d} \mathbb{I}_d$ .

For notational simplicity, we consider the case  $d = 2$ , the general case can be written similarly.

Now for each  $j \in \mathbb{N}$ ,  $I_j X_j$  is a lattice random vector with the minimal lattice  $\mathbb{Z}^2$ . It is easy to note that  $2\pi\mathbb{Z} \times 2\pi\mathbb{Z}$  is the set of all periods for  $I_j X_j$  and its fundamental domain is given by  $(-\pi, \pi)^2$ . To prove (17), it is enough to show

$$\sup_{x \in \frac{1}{\sqrt{2}} \mathcal{L}_n^{(2)}} \left| (\log n) \mathbb{P} \left( \frac{Z_n}{\sqrt{\log n}} = x \right) - \phi_{2, \frac{1}{2} \mathbb{I}_2}(x) \right| \rightarrow 0 \text{ as } n \rightarrow \infty,$$

where  $\phi_{2, \frac{1}{2} \mathbb{I}_2}(x) = \frac{1}{\pi} \exp(-\|x\|^2)$  is the bivariate normal density with mean vector 0 and variance-covariance matrix  $\frac{1}{2} \mathbb{I}_2$ , and  $\frac{1}{\sqrt{2}} \mathcal{L}_n^{(2)} = \frac{1}{\sqrt{\log n}} \mathbb{Z}^2$ . We again use the Fourier inversion formula (see 21.28 on page 230 of [4]), and get that for  $x \in \frac{1}{\sqrt{2}} \mathcal{L}_n^{(2)}$ ,

$$\mathbb{P} \left( \frac{Z_n}{\sqrt{\log n}} = x \right) = \frac{1}{(2\pi)^2 \log n} \int_{(-\sqrt{\log n} \pi, \sqrt{\log n} \pi)^2} \psi_n(t) \exp(-i \langle t, x \rangle) dt.$$

Further, we also have

$$\phi_{2, \frac{1}{2} \mathbb{I}_2}(x) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \exp(-i \langle t, x \rangle) \exp\left(-\frac{\|t\|^2}{4}\right) dt.$$

Put  $H_n = (-\sqrt{\log n} \pi, \sqrt{\log n} \pi)^2$ , then given  $\epsilon > 0$ , there exists  $N > 0$ , such that for all  $n \geq N$ , we have the estimate

$$\begin{aligned} \left| \log n \mathbb{P} \left( \frac{Z_n}{\sqrt{\log n}} = x \right) - \phi_{2, \frac{1}{2} \mathbb{I}_2}(x) \right| &\leq \frac{1}{(2\pi)^2} \int_{H_n} \left| \psi_n(t) - \exp\left(-\frac{\|t\|^2}{4}\right) \right| dt \\ &\quad + \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2 \setminus H_n} \exp\left(-\frac{\|t\|^2}{4}\right) dt \\ &\leq \frac{1}{(2\pi)^2} \int_{H_n} \left| \psi_n(t) - \exp\left(-\frac{\|t\|^2}{4}\right) \right| dt + \epsilon. \end{aligned}$$

Similar to the proof of Theorem 4, it is enough to show that the first term in the above inequality tends to zero as  $n \rightarrow \infty$ . As discussed earlier to prove this it is



enough to show that for any  $\epsilon > 0$ , we can choose a compact subset  $A$  of  $\mathbb{R}^d$ , such that for all  $n$  large enough

$$(57) \quad \mathcal{I}(n, A) < \epsilon,$$

where

$$\mathcal{I}(n, A) := \int_{H_n \setminus A} |\psi_n(t)| dt.$$

For more details we refer the reader to the proof of Theorem 2 on page 190 of [37].

Applying a change of variables  $t = \frac{1}{\sqrt{\log n}} w$ , we obtain

$$\mathcal{I}(n, A) = \log n \int_{(-\pi, \pi)^2 \setminus \frac{1}{\sqrt{\log n}} A} |\psi_n(\sqrt{\log nw})| dw.$$

We can write

$$\mathcal{I}(n, A) = \mathcal{I}_1(n, A) + \mathcal{I}_2(n)$$

where

$$\mathcal{I}_1(n, A) := \log n \int_{(B(0, \delta) \setminus \frac{1}{\sqrt{\log n}} A) \cap (-\pi, \pi)^2} |\psi_n(\sqrt{\log nw})| dw,$$

and

$$\mathcal{I}_2(n) := \log n \int_{(-\pi, \pi)^2 \setminus B(0, \delta)} |\psi_n(\sqrt{\log nw})| dw.$$

where  $\delta$  is as in (49). Using arguments similar to (51), (52) and (53) one can easily show  $\mathcal{I}_1(n, A) < \epsilon$ . Therefore, it is enough to show that

$$\mathcal{I}_2(n) \longrightarrow 0 \text{ as } n \rightarrow \infty.$$

To do so, we first observe that for  $t = (t^{(1)}, t^{(2)}) \in \mathbb{R}^2$ , the characteristic function for  $X_1$  is given by  $e(it) = \frac{1}{2} (\cos t^{(1)} + \cos t^{(2)})$ . If  $t \in [-\pi, \pi]^2$  be such that  $|e(it)| = 1$ , then  $t \in \{(\pi, \pi), (-\pi, \pi), (\pi, -\pi), (-\pi, -\pi)\}$ . The function  $\cos \theta$  is continuous and decreasing as a function of  $\theta$  for  $t \in [\frac{\pi}{2}, \pi]$ . Choose  $0 < \eta < \frac{\pi}{2}$  such that for  $t \in A_1 = (-\pi, \pi)^2 \cap B^c(0, \delta) \cap D^c$ , we have  $|e(it)| < 1$ , where  $D = [\pi - \eta, \pi]^2 \cup [-\pi + \eta, -\pi] \times [\pi - \eta, \pi] \cup [-\pi + \eta, -\pi]^2 \cup [\pi - \eta, \pi] \times [-\pi + \eta, -\pi]$ . Let us write

$$\mathcal{I}_2(n) := \mathcal{J}_1(n) + \mathcal{J}_2(n),$$

where

$$\mathcal{J}_1(n) := \log n \int_{A_1} |\psi_n(\sqrt{\log nw})| dw,$$

and

$$\mathcal{J}_2(n) := \log n \int_D |\psi_n(\sqrt{\log nw})| dw.$$

It is easy to note that

$$\mathcal{J}_1(n) \leq \log n \int_{\bar{A}_1} |\psi_n(\sqrt{\log nw})| dw$$

where  $\bar{A}_1$  denotes the closure of  $A_1$ . For  $w \in \bar{A}_1$ , there exists  $0 < \alpha < 1$ , such that,  $|e(it)| \leq \alpha$ . Therefore using bounds similar to that in (55) we can show that

$$\mathcal{J}_1(n) \longrightarrow 0 \text{ as } n \rightarrow \infty.$$

We observe that

$$\mathcal{J}_2(n) \leq 4 \log n \int_{[\pi-\eta, \pi]^2} |\psi_n(\sqrt{\log nw})| dw.$$

Hence, it is enough to show that as  $n \rightarrow \infty$

$$\log n \int_{[\pi-\eta, \pi]^2} |\psi_n(\sqrt{\log nw})| dw \longrightarrow 0.$$

For  $w \in [\pi - \eta, \pi]^2$ , we have  $0 < \left|1 + \frac{e(iw)}{j}\right| \leq \left(1 + \frac{\cos(\pi-\eta)}{j}\right) \leq 1$ . Therefore,

$$|\psi_n(w)| = \frac{1}{n+1} \prod_{j=1}^n \left|1 + \frac{e(iw)}{j}\right| \leq \frac{1}{n+1}.$$

So,

$$\log n \int_{[\pi-\eta, \pi]^2} |\psi_n(\sqrt{\log nw})| dw \leq \frac{\eta^2}{n+1} \log n \longrightarrow 0 \text{ as } n \rightarrow \infty.$$

□

We now assume that  $d = 1$  and  $\mathbb{P}(X_1 = 0) = 0$ . Let  $\tilde{h}$  be the span for  $X_1$ . We can now write  $\mathbb{P}(I_1 X_1 \in a + h\mathbb{Z}) = 1$ , where  $a \in \mathbb{R}$  and  $h > 0$  is the span for  $I_1 X_1$ . It is easy to note that  $h \leq \tilde{h}$ . The following result gives a local limit theorem under the assumption that  $\tilde{h} < 2h$ . This assumption is non-trivial, and a concrete case is, when  $\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = 2) = 1/2$ . Then,  $\tilde{h} = 1$ . The support for  $I_1 X_1$  is  $\{0, 1, 2\}$  and  $h = 1$ , satisfying the assumption  $\tilde{h} < 2h$ .

**Theorem 10.** *Assume that  $\tilde{h} < 2h$ , then, as  $n \rightarrow \infty$*

$$(58) \quad \sup_{x \in \mathcal{L}_n^{(1)}} \left| \sigma \frac{\sqrt{\log n}}{h} \mathbb{P} \left( \frac{Z_n - \mu \log n}{\sigma \sqrt{\log n}} = x \right) - \phi(x) \right| \longrightarrow 0,$$

where  $\mathcal{L}_n^{(1)}$  is as defined in (15).

*Proof of Theorem 10.* This proof is similar to the proof of Theorem 4, except we now have  $d = 1$ , which is simpler. So we omit most of the details and only point out the key differences, where the assumption  $\tilde{h} < 2h$  is crucial.

We begin by observing that the bounds for  $\left| \frac{\sigma \sqrt{\log n}}{h} \mathbb{P} \left( \frac{Z_n - \mu \log n}{\sigma \sqrt{\log n}} = x \right) - \phi(x) \right|$  are similar to those in the proof of Theorem 4, except for that of  $\mathcal{I}_2(n)$ , where in

this case,

$$\mathcal{I}_2(n) := \sqrt{\log n} \int_{\sigma\delta}^{\frac{\sigma\pi}{h}} \left| g_n \left( w\sqrt{\log n} \right) \right| dw$$

and  $\delta$  is chosen as in (49) and  $g_n(\cdot)$  is as in (47). We have to show

$$\mathcal{I}_2(n) \longrightarrow 0 \text{ as } n \rightarrow \infty.$$

The span of  $X_1$  being  $\tilde{h}$ , for all  $t \in \left[ \delta, \frac{2\pi}{h} \right)$ ,  $|e(it)| < 1$ . Note that here we use the assumption  $\tilde{h} < 2h$ . The characteristic function being continuous in  $t$ , there exists  $0 < \eta < 1$ , such that,  $|e\left(\frac{it}{\sigma}\right)| \leq \eta$  for all  $t \in \left[ \delta\sigma, \frac{\pi\sigma}{h} \right) \subset \left[ \delta\sigma, \frac{2\pi\sigma}{h} \right)$ . So, as  $n \rightarrow \infty$ ,

$$(59) \quad \mathcal{I}_2(n) \leq C_2\sigma \exp\left(- (1 - \eta) \log n\right) (\pi - \delta) \sqrt{\log n} \longrightarrow 0.$$

□

## 5. URNS WITH COLORS INDEXED BY OTHER LATTICES ON $\mathbb{R}^d$

We can further generalize the urn models with colors indexed by certain countable lattices in  $\mathbb{R}^d$ . Such a model will be associated with the corresponding random walk on the lattice. To state the results rigorously we consider the following notations.

Let  $\{X_j\}_{j \geq 1}$  be a sequence of random  $d$ -dimensional i.i.d. vectors with non empty support set  $B \subseteq \mathbb{R}^d$  and probability mass function  $p$ . We assume that  $B$  is finite. Consider the countable subset

$$S^d := \left\{ \sum_{i=1}^k n_i b_i : n_1, n_2, \dots, n_k \in \mathbb{N}, b_1, b_2, \dots, b_k \in B \right\}$$

of  $\mathbb{R}^d$  which will index the set of colors.

As before we consider  $S_n := X_0 + X_1 + \dots + X_n$ ,  $n \geq 0$ , the random walk starting at  $X_0$  which is distributed as  $U_0$ . We say a process  $(U_n)_{n \geq 0}$  is a urn scheme with colors indexed by  $S^d$  and replacement matrix  $R$  and starting configuration  $U_0$ , if  $(U_n)_{n \geq 1}$  is defined recursively by (2), where now  $R$  is given by

$$R := ((p(v - u)))_{u, v \in S^d}.$$

Following the same nomenclature as done earlier, we will call this process the *infinite color urn model associated with the random walk  $\{S_n\}_{n \geq 0}$  on  $S^d$* . Naturally, when  $S^d = \mathbb{Z}^d$ , this process is exactly the one discussed earlier.

We will use same notations as earlier for the mean, non-centered second moment matrix and moment generating function for the increment  $X_1$  (see (1) for the definitions). As before, we denote by  $Z_n$  the  $(n + 1)$ -th selected color. Just like in the previous case, the expected proportion of colors in the urn at time  $n$  will be given by the distribution of  $Z_n$  but now on  $S^d$ .

From the proof of Theorem 7 it follows that the result holds also for this generalization. This enable us to generalize Theorem 1 and Theorem 3 as follows.

**Theorem 11.** *Let  $\bar{\Lambda}_n$  be the probability measure on  $\mathbb{R}^d$  corresponding to the probability vector  $\frac{1}{n+1} (\mathbb{E}[U_{n,v}])_{v \in S^d}$  and let*

$$\bar{\Lambda}_n^{cs}(A) := \bar{\Lambda}_n \left( \sqrt{\log n} A \Sigma^{1/2} + \mu \log n \right), \quad A \in \mathcal{B}(\mathbb{R}^d).$$

Then, as  $n \rightarrow \infty$ ,

$$(60) \quad \bar{\Lambda}_n^{cs} \Rightarrow \Phi_d.$$

**Theorem 12.** *Let  $\Lambda_n \in \mathcal{M}_1$  be the random probability measure corresponding to the random probability vector  $\frac{U_n}{n+1}$ . Let*

$$\Lambda_n^{cs}(A) = \Lambda_n \left( \sqrt{\log n} A \Sigma^{1/2} + \mu \log n \right).$$

where  $A$  is a Borel subset of  $\mathbb{R}^d$ . Then, as  $n \rightarrow \infty$ ,

$$(61) \quad \Lambda_n^{cs} \xrightarrow{p} \Phi_d \text{ in } \mathcal{M}_1.$$

The proofs of these two theorems are similar to Theorem 1 and Theorem 3 respectively and hence omitted.

As an application we now consider a specific example, namely, the triangular lattice in two dimensions. For this the support set for the i.i.d. increment vectors is given by

$$B = \{(1, 0), (-1, 0), \omega, -\omega, \omega^2, -\omega^2\},$$

where  $\omega, \omega^2$  are the complex cube roots of unity. The law of  $X_1$  is uniform on  $B$ . This gives the random walk on the triangular lattice in two dimensions. The following is an immediate corollary of Theorem 11.

**Corollary 13.** *Consider the urn model associated with the random walk on two dimensional triangular lattice then as  $n \rightarrow \infty$*

$$(62) \quad \frac{Z_n}{\sqrt{\log n}} \Rightarrow N_2 \left( 0, \frac{1}{2} \mathbb{I}_2 \right).$$

*Proof.* Since  $1 + \omega + \omega^2 = 0$ , therefore it is immediate that  $\mu = 0$ . Also we know that  $\omega = \frac{1}{2} + i\frac{\sqrt{3}}{2}$ . Writing  $\omega = (Re(\omega) + iIm(\omega))$ , we get

$$\mathbb{E} \left[ \left( X_1^{(1)} \right)^2 \right] = \frac{2}{6} \left( 1 + (Re(\omega))^2 + (Re(\omega^2))^2 \right).$$

Since  $Re(\omega) = Re(\omega^2)$ , therefore

$$\mathbb{E} \left[ \left( X_1^{(1)} \right)^2 \right] = \frac{2}{6} \left( 1 + 2(Re(\omega))^2 \right) = \frac{1}{2}.$$

Similarly,  $Im(\omega) = -Im(\omega^2)$ , and hence  $\mathbb{E} \left[ \left( X_1^{(2)} \right)^2 \right] = \frac{2}{6} \left( (Im(\omega))^2 + (Im(\omega^2))^2 \right) = \frac{1}{2}$ . Finally

$$\mathbb{E} \left[ X_1^{(1)} X_1^{(2)} \right] = -\frac{2}{6} Im(1 + \omega + \omega^2) = 0.$$

So  $\Sigma = \frac{1}{2} \mathbb{I}_2$ . The rest is just an application of Theorem 11.  $\square$

#### APPENDIX

We present here an elementary but technical result which we have used in the proof of Theorem 3. It is really a generalization of the classical result for Laplace transform, namely, Theorem 22.2 of [8].

**Theorem 14.** *Let  $\nu_n$  be a sequence of probability measures on  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$  and let  $m_n(\cdot)$  be the corresponding moment generating functions. Suppose there exists  $\delta > 0$  such that  $m_n(\lambda) \rightarrow \exp\left(\frac{\|\lambda\|^2}{2}\right)$  as  $n \rightarrow \infty$  for every  $\lambda \in [-\delta, \delta]^d \cap \mathbb{Q}^d$ , then as  $n \rightarrow \infty$*

$$(63) \quad \nu_n \Rightarrow \Phi_d.$$

*Proof.* Choose a  $\delta' \in \mathbb{Q}$  such that  $0 < \delta' < \delta$ , and observe that for every  $a > 0$

$$\nu_n \left( \left( [-a, a]^d \right)^c \right) \leq \sum_{i=1}^d \exp(-\delta' a) (m_n(-\delta' e_i) + m_n(\delta' e_i)),$$

where  $\{e_i\}_{i=1}^d$  are the  $d$ -unit vectors. Now for our assumption we get  $m_n(\delta' e_i) \rightarrow \exp\left(\frac{\delta'^2}{2}\right)$  and  $m_n(-\delta' e_i) \rightarrow \exp\left(\frac{\delta'^2}{2}\right)$  as  $n \rightarrow \infty$  for every  $1 \leq i \leq d$ . Thus we get

$$\sup_{n \geq 1} \nu_n \left( \left( [-a, a]^d \right)^c \right) \rightarrow 0 \text{ as } a \rightarrow \infty.$$

So the sequence of probability measures  $(\nu_n)_{n \geq 1}$  is tight. Therefore, for every subsequence  $\{n_k\}_{k \geq 1}$  there exists a further subsequence  $\{n_{k_j}\}_{j \geq 1}$  and a probability measure  $\nu$  such that as  $n \rightarrow \infty$ ,

$$\nu_{n_{k_j}} \Rightarrow \nu.$$

Then by dominated convergence theorem

$$m_{n_{k_j}}(\lambda) \rightarrow m_\infty(\lambda), \quad \forall \lambda \in (-\delta, \delta)^d \cap \mathbb{Q}^d$$

where  $m_\infty$  is the moment generating function of  $\nu$ . But from our assumption

$$m_{n_{k_j}}(\lambda) \rightarrow \exp\left(\frac{\|\lambda\|^2}{2}\right), \quad \forall \lambda \in [-\delta, \delta]^d \cap \mathbb{Q}^d.$$

So we conclude that

$$m_\infty(\lambda) = \exp\left(\frac{\|\lambda\|^2}{2}\right), \quad \forall \lambda \in (-\delta, \delta)^d \cap \mathbb{Q}^d.$$

Since both sides of the above equation are continuous functions on their respective domains, we get that  $m_\infty(\lambda) = \exp\left(\frac{\|\lambda\|^2}{2}\right)$  for every  $\lambda \in (-\delta, \delta)^d$ . But the standard Gaussian distribution is characterized by the values of its moment generating function in an open neighborhood of 0, so we conclude that every sub-sequential limit is standard Gaussian. This proves (63).  $\square$

#### ACKNOWLEDGEMENT

The authors are grateful to Krishanu Maulik and Codina Cotar for various discussions they had with them. We are also thankful to the anonymous referees for their valuable remarks which has significantly improved the exposition of the article.

## REFERENCES

- [1] 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.
- [2] A. Bagchi and A. K. Pal. Asymptotic normality in the generalized Pólya-Eggenberger urn model, with an application to computer data structures. *SIAM J. Algebraic Discrete Methods*, 6(3):394–405, 1985.
- [3] Zhi-Dong Bai and Feifang Hu. Asymptotics in randomized urn models. *Ann. Appl. Probab.*, 15(1B):914–940, 2005.
- [4] R. N. Bhattacharya and R. Ranga Rao. *Normal approximation and asymptotic expansions*. John Wiley & Sons, New York-London-Sydney, 1976. Wiley Series in Probability and Mathematical Statistics.
- [5] J. D. Biggins. Chernoff’s theorem in the branching random walk. *J. Appl. Probability*, 14(3):630–636, 1977.
- [6] J. D. Biggins. Martingale convergence in the branching random walk. *J. Appl. Probability*, 14(1):25–37, 1977.
- [7] J. D. Biggins. Lindley-type equations in the branching random walk. *Stochastic Process. Appl.*, 75(1):105–133, 1998.
- [8] Patrick Billingsley. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edition, 1995. A Wiley-Interscience Publication.
- [9] David Blackwell. Discreteness of Ferguson selections. *Ann. Statist.*, 1:356–358, 1973.
- [10] David Blackwell and James B. MacQueen. Ferguson distributions via Pólya urn schemes. *Ann. Statist.*, 1:353–355, 1973.
- [11] Arup Bose, Amites Dasgupta, and Krishanu Maulik. Multicolor urn models with reducible replacement matrices. *Bernoulli*, 15(1):279–295, 2009.
- [12] Arup Bose, Amites Dasgupta, and Krishanu Maulik. Strong laws for balanced triangular urns. *J. Appl. Probab.*, 46(2):571–584, 2009.
- [13] May-Ru Chen, Shoou-Ren Hsiau, and Ting-Hsin Yang. A new two-urn model. *J. Appl. Probab.*, 51(2):590–597, 2014.
- [14] May-Ru Chen and Markus Kuba. On generalized Pólya urn models. *J. Appl. Probab.*, 50(4):1169–1186, 2013.
- [15] Andrea Collecchio, Codina Cotar, and Marco LiCalzi. On a preferential attachment and generalized Pólya’s urn model. *Ann. Appl. Probab.*, 23(3):1219–1253, 2013.
- [16] John B. Conway. *Functions of one complex variable*, volume 11 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1978.
- [17] Codina Cotar and Vlada Limic. Attraction time for strongly reinforced walks. *Ann. Appl. Probab.*, 19(5):1972–2007, 2009.
- [18] Edward Crane, Nicholas Georgiou, Stanislav Volkov, Andrew R. Wade, and Robert J. Waters. The simple harmonic urn. *Ann. Probab.*, 39(6):2119–2177, 2011.
- [19] Amites Dasgupta and Krishanu Maulik. Strong laws for urn models with balanced replacement matrices. *Electron. J. Probab.*, 16:no. 63, 1723–1749, 2011.
- [20] Burgess Davis. Reinforced random walk. *Probab. Theory Related Fields*, 84(2):203–229, 1990.

- [21] Luc Devroye. Applications of the theory of records in the study of random trees. *Acta Inform.*, 26(1-2):123–130, 1988.
- [22] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition, 2010.
- [23] Thomas S. Ferguson. A Bayesian analysis of some nonparametric problems. *Ann. Statist.*, 1:209–230, 1973.
- [24] 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.
- [25] David A. Freedman. Bernard Friedman’s urn. *Ann. Math. Statist.*, 36:956–970, 1965.
- [26] Bernard Friedman. A simple urn model. *Comm. Pure Appl. Math.*, 2:59–70, 1949.
- [27] Raúl Gouet. Strong convergence of proportions in a multicolor Pólya urn. *J. Appl. Probab.*, 34(2):426–435, 1997.
- [28] Svante Janson. Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Process. Appl.*, 110(2):177–245, 2004.
- [29] Svante Janson. Limit theorems for triangular urn schemes. *Probab. Theory Related Fields*, 134(3):417–452, 2006.
- [30] Sophie Laruelle and Gilles Pagès. Randomized urn models revisited using stochastic approximation. *Ann. Appl. Probab.*, 23(4):1409–1436, 2013.
- [31] Mickaël Launay and Vlada Limic. Generalized Interacting Urn Models. available at <http://arxiv.org/abs/1201.3495>, 2012.
- [32] Gregory F. Lawler and Vlada Limic. *Random walk: a modern introduction*, volume 123 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.
- [33] Vlada Limic. Attracting edge property for a class of reinforced random walks. *Ann. Probab.*, 31(3):1615–1654, 2003.
- [34] Vlada Limic and Pierre Tarrès. Attracting edge and strongly edge reinforced walks. *Ann. Probab.*, 35(5):1783–1806, 2007.
- [35] Robin Pemantle. A time-dependent version of Pólya’s urn. *J. Theoret. Probab.*, 3(4):627–637, 1990.
- [36] Robin Pemantle. A survey of random processes with reinforcement. *Probab. Surv.*, 4:1–79, 2007.
- [37] V. V. Petrov. *Sums of independent random variables*. Springer-Verlag, New York-Heidelberg, 1975. Translated from the Russian by A. A. Brown, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82.
- [38] Georg Pólya. Sur quelques points de la théorie des probabilités. *Ann. Inst. H. Poincaré*, 1(2):117–161, 1930.
- [39] Robert T. Smythe and Hosam M. Mahmoud. A survey of recursive trees. *Teor. Īmovir. Mat. Stat.*, (51):1–29, 1994.
- [40] A. W. van der Vaart. *Asymptotic statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998.

(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

*E-mail address:* [antar@isid.ac.in](mailto:antar@isid.ac.in)

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

*E-mail address:* [thackerdebleena@gmail.com](mailto:thackerdebleena@gmail.com)