isid/ms/2013/01
April 2, 2013
http://www.isid.ac.in/~statmath/eprints

Pólya Urn Schemes with Infinitely Many Colors

Antar Bandyopadhyay and Debleena Thacker

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi–110 016, India

Pólya Urn Schemes with Infinitely Many Colors

Antar Bandyopadhyay*
Debleena Thacker†

Theoretical Statistics and Mathematics Unit
Indian Statistical Institute
7 S. J. S. Sansanwal Marg
New Delhi 110016
INDIA

April 3, 2013

Abstract

In this work we introduce a new urn model with infinite but countably many colors indexed by an appropriate infinite set. We mainly focus on d-dimensional integer lattice and replacement matrix associated with bounded increment random walks on it. We prove central and local limit theorems for the expected configuration of the urn and show that irrespective of the null recurrent or transient behavior of the underlying random walk, the urn models have universal scaling and centering giving appropriate normal distribution at the limit. The work also provides similar results for urn models corresponding to other infinite lattices.

Keywords: Central limit theorem, infinite color urn, local limit theorem, random walk, universality, urn models.

AMS 2010 Subject Classification: Primary: 60B10, 60B12, 60B05; Secondary: 60C05, 60G50.

1 Introduction

Pólya urn schemes and its various generalizations with finitely many colors have been widely studied in literature [16, 11, 10, 1, 2, 14, 12, 13, 9, 5, 6, 8], also see [15] for an extensive survey of many of the known results. The model is described as follows:

> We start with an urn containing finitely many balls of different colors. At any time $n \geq 1$, a ball is selected uniformly at random from the urn, the color of the selected ball is noted, the selected ball is returned to the urn along with a set of balls of various colors which may depend on the color of the selected ball.

*E-Mail: `antar@isid.ac.in`
†E-Mail: `thackerdebleena@gmail.com`

The goal is to study the asymptotic properties of the configuration of the urn. Suppose there are $K \geq 1$ different colors and we denote the configuration of the urn at time n by $U_n = (U_{n,1}, U_{n,2} \ldots, U_{n,K})$, where $U_{n,j}$ denotes the number of balls of color j, $1 \leq j \leq K$. The dynamics of the urn model depends on the *replacement policy* which can be presented by a $K \times K$ matrix, say R whose $(i,j)^{\text{th}}$ entry is the number of balls of color j which are to be added to the urn if the selected color is i. In literature R is termed as *replacement matrix*. The dynamics of the model can then be written as

$$U_{n+1} = U_n + R_i \tag{1.1}$$

where R_i is the i^{th} row of the replacement matrix R if the (random) color of the ball selected at the $(n+1)^{\text{th}}$ draw is i.

A replacement matrix is said to be *balanced*, if the row sums are constant. In that case the asymptotic behavior of the proportion of balls of different colors will be same if we change the replacement matrix R by R/s where s is the row sum. Note that the later matrix is a *stochastic matrix* and hence the entries are not necessarily non-negative integers. Since we will be interested mostly in the asymptotic behavior of the configuration of a balanced urn model, so without loss of any generality, we assume that the replacement matrix R is a stochastic matrix. In that case it is also customary to assume that U_0 is a probability distribution on the set of colors, which is to be interpreted as the probability distribution of the selected color of the first ball drawn from the urn. Note in this case the entries of U_n indexed by the colors are no longer the number of balls of that color, but $U_n/(n+1)$ is the probability mass function associated with the the probability distribution of the color of the $(n+1)^{\text{th}}$ selected ball. In other words, if Z_n is the color of the ball selected at the $(n+1)^{\text{th}}$ draw then

$$\mathbb{P}\left(Z_n = i \,\middle|\, U_n\right) = \frac{U_{n,i}}{n+1}, \; 1 \leq i \leq K. \tag{1.2}$$

We can now consider a Markov chain with state space as the set of colors, the transition matrix as R and starting distribution as U_0. We call such a chain, a chain associated with the urn model and vice-versa. In other words given a balanced urn model we can associate with it a unique Markov chain on the set of colors and conversely given a Markov chain on a finite state space there is an associated urn model with balls of colors indexed by the state space.

It is well known [12, 13, 5, 6, 8] that the asymptotic properties of a balanced urn are often related to the qualitative properties of this associated Markov chain on finite state space. For example, if the associated finite state Markov chain is irreducible and aperiodic with stationary distribution π then in [12, 13] it has been proved that

$$\frac{U_n}{n+1} \longrightarrow \pi \;\; \text{a.s.} \tag{1.3}$$

The reducible case for the finite color urn model has been extensively studied in [5, 6, 8] and various different kind of limiting results have been derived based on the properties of the replacement/stochastic matrix R.

In this work we introduce a new urn model with countable but infinitely many colors. We will show that for such a generalization unlike in the finite color case, the asymptotic behavior of the configuration may not always depend only on the qualitative properties of the associated Markov chain.

1.1 Urn Model with Infinitely Many Colors

We consider the natural generalization of the urn model to infinitely many colors through the associated Markov chain on countably infinite state space. More precisely, let S be a countable and possibly infinite set and let R be a stochastic matrix on S. Starting with an initial configuration U_0 which is a probability distribution on S we consider the urn model with possibly infinitely many colors indexed by S. The dynamics of the model remains same as above, that is, at the $(n+1)^{\text{th}}$ trial if we choose a ball of color $u \in S$, then

$$U_{n+1} = U_n + R_u$$

where R_u is the u^{th} row of the replacement matrix R. Once again $U_n / (n+1)$ is the probability mass function associated with the the probability distribution of the color of the $(n+1)^{\text{th}}$ selected ball, that is, if Z_n is the color of the ball selected at the $(n+1)^{\text{th}}$ draw then

$$\mathbb{P}\left(Z_n = u \,\middle|\, U_n\right) = \frac{U_{n,u}}{n+1}, \;\; u \in S. \tag{1.4}$$

This basic recursion can be written in the matrix notation as follows

$$U_{n+1} = U_n + I_{n+1} R \tag{1.5}$$

where $I_n = (I_{n,v})_{v \in V}$ with $I_{n,u} = 1$ and $I_{n,v} = 0$ if $v \neq u$ where u is the color of the ball chosen at $(n+1)^{\text{th}}$ trial.

In this work, we mainly consider the special case of this generalization where the associated Markov chain is an i.i.d. bounded increment random walk on $\mathbb{Z}^d$. Although in Section 5 of the paper we also consider the case of general random walk on $\mathbb{R}^d$ with i.i.d. discrete bounded increments. Let $\{X_i\}_{i \geq 1}$ be a sequence of random d-dimensional i.i.d. vectors and $\emptyset \neq B \subseteq \mathbb{Z}^d$ be the support for X_1. We assume B is a finite set. Let the law of X_1 be given by the mass function

$$p(u) := \mathbb{P}(X_1 = u), \;\; u \in B,$$

where we assume $0 < p(u) \leq 1, \;\; u \in B$ and $\sum_{u \in B} p(u) = 1$. Let R be the transition matrix for the random walk $S_n = \sum_{k=1}^{n} X_k$, then it is easy to see that

$$R(u,v) = \begin{cases} p(v-u) & \text{if } v-u \in B \\ 0 & \text{otherwise} \end{cases} \tag{1.6}$$

for all $u, v \in \mathbb{Z}^d$. Following few special cases are of particular interest:

1. In one dimension we consider a trivial walk, namely, "*move one step to the right*". Formally, in this case $d = 1$ and $B = \{1\}$, the law of X_1 is given by $\mathbb{P}(X_1 = 1) = 1$. The associated Markov chain $S_n = S_0 + n$ is deterministic and trivially transient. The transition matrix R is given by

$$R(i,j) = \begin{cases} 1 & \text{if } j = i+1 \\ 0 & \text{otherwise.} \end{cases} \tag{1.7}$$

We call this R the *right-shift operator*.

2. The other special case is the *simple symmetric random walk* on $\mathbb{Z}^d$. For this $B = \{v \in \mathbb{Z}^d \,\big|\, \|v\|_1 = 1\}$ where $\|\cdot\|_1$ denotes the l_1 norm. The law of X_1 is given by $\mathbb{P}(X_1 = v) = \frac{1}{2d},\ v \in B$. The matrix R is given by

$$R(u,v) = \begin{cases} \frac{1}{2d} & \text{for } v - u \in B \\ 0 & \text{otherwise.} \end{cases} \tag{1.8}$$

We here note that by the famous result of Pólya for $d \le 2$, the simple symmetric random walk is null recurrent, while for $d \ge 3$, it is transient.

1.2 Notations

Following notations and conventions are used in the paper.

- For two sequences $\{a_n\}_{n\ge1}$ and $\{b_n\}_{n\ge1}$ of positive real numbers such that $b_n \ne 0$ for all $n \ge 1$, we will write $a_n \sim b_n$ if $\lim_{n\to\infty} \frac{a_n}{b_n} = 1$.
- Vectors are written as row vectors unless otherwise stated. For example, a finite dimensional vector $\mathbf{x} \in \mathbb{R}^d$ is written as $\mathbf{x} = \left(x^{(1)}, x^{(2)}, \ldots, x^{(d)}\right)$ where $x^{(i)}$ denotes the i^{th} coordinate. To be consistent with this notation matrices are multiplied to the right of the vectors. The infinite dimensional vectors are written as $y = (y_j)_{j\in\mathcal{J}}$ where y_j is the j^{th} coordinate and $\mathcal{J}$ is the indexing set. For any vector $\mathbf{x}$, $\mathbf{x}^t$ will denote the transpose of x.
- By $N_d(\mu, \varSigma)$ we denote the d-dimensional Gaussian distribution with mean vector $\mu \in \mathbb{R}^d$ and variance-covariance matrix $\varSigma$. The associated Gaussian measure on $\left(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right)\right)$ is denoted by Φ_d and by ϕ_d we denote the corresponding density function. For $d = 1$, we simply write $N(\mu, \sigma^2)$ with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$ and associated measure Φ and the density by ϕ.
- The symbol $\xrightarrow{w}$ will stand for weak convergence, $\xrightarrow{d}$ will denote convergence in distribution, while $\xrightarrow{p}$ will denote convergence in probability.
- For any two random variables/vectors X and Y, we will write $X \stackrel{d}{=} Y$ to denote that X and Y have the same law.
- The symbol d is used both for dimension and distribution, it will be clear from the context what it stands for.

1.3 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 d-dimensional random walk in $\mathbb{R}^d$ with i.i.d. discrete bounded increments, in particular we consider the two dimensional triangular lattice. Finally in Section 6 contains a technical result which is required for the proofs in Section 4.

2 Main Results

In this section, we present the main results for this paper. 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

In this subsection, we present the main results for the urns in which the colors of the balls are indexed $\mathbb{Z}^d,\ d \geq 1$. Let $(X_i)_{i\geq 1}$ be a sequence of i.i.d. random d-dimensional vectors with non-empty finite support $B \subset \mathbb{Z}^d$. Let the law of X_1 be given by $\mathbb{P}\left(X_1 = v\right) = p(v)$, for $v \in B$, where $0 < p(v) \leq 1$ and $\sum_{v\in B} p(v) = 1$. For $X_1 = \left(X_1^{(1)}, X_1^{(2)}, \ldots, X_1^{(d)}\right)$, let $\mu = \left(\mathbb{E}[X_1^{(1)}], \mathbb{E}[X_1^{(2)}], \ldots, \mathbb{E}[X_1^{(d)}]\right)$ and $\Sigma = ((\sigma_{ij}))_{1\leq i,j\leq d}$ where $\sigma_{ij} = \mathbb{E}[X_1^{(i)} X_1^{(j)}]$. We assume Σ to be positive definite. We write $\Sigma = \Sigma^{\frac{1}{2}} \Sigma^{\frac{1}{2}}$ where $\Sigma^{\frac{1}{2}}$ is the unique symmetric positive definite matrix known as the *positive square root* of Σ. Let $\{S_n\}_{n\geq 0}$ be the random walk with increments $\{X_j\}_{j\geq 1}$, that is $S_n = \sum_{j=1}^n X_j$. We also assume that the urn model starts with the initial configuration $U_0 = (U_{0,v})_{v\in\mathbb{Z}^d}$ such that $U_{0,v} = 0$ except for finitely many v.

Theorem 2.1. *Consider the urn model associated with the random walk $\{S_n\}_{n\geq 0}$. Suppose $U_0 = (U_{0,v})_{v\in\mathbb{Z}^d}$ be such that $U_{0,v} = 0$ except for finitely many v. Let $\overline{\Lambda}_n$ be the probability measure on $\mathbb{R}^d$ corresponding to the probability vector $\frac{1}{n+1}\left(\mathbb{E}[U_{n,v}]\right)_{v\in\mathbb{Z}^d}$ and $\overline{\Lambda}_n^{cs}$ be defined as $\overline{\Lambda}_n^{cs}(A) = \overline{\Lambda}_n\left(\sqrt{\log n} A \Sigma^{-1/2} + \mu \log n\right)$ for $A \in \mathcal{B}\left(\mathbb{R}^d\right)$. Then, as $n \to \infty$,*

$$\overline{\Lambda}_n^{cs} \xrightarrow{w} \Phi_d \tag{2.9}$$

where $\Phi_d\left(A\right) = \frac{1}{\left(\sqrt{2\pi}\right)^d} \int\limits_A e^{-\frac{xx^t}{2}}\, \mathrm{d}t$ for $A \in \mathcal{B}\left(\mathbb{R}^d\right)$.

Since all vectors are taken to be row vectors, we always multiply the matrices to the right of the vectors.

Remark 1. Theorem 2.1 states that if Z_n be the color of the randomly selected ball in the $(n+1)^{\text{th}}$ draw, then

$$\frac{Z_n - \mu \log n}{\sqrt{\log n}} \xrightarrow{d} N_d(\mathbf{0}, \Sigma) \text{ as } n \to \infty. \tag{2.10}$$

Corollary 2.2. *For $d = 1$, let $X_i \equiv 1$, that is the underlying Markov chain moves deterministically one step to the right, then as $n \to \infty$*

$$\frac{Z_n - \log n}{\sqrt{\log n}} \xrightarrow{d} N(0, 1).$$

Corollary 2.3. *Let S_n be the simple symmetric random walk on $\mathbb{Z}^d,\ d \geq 1$. Then, as $n \to \infty$,*

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

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

Remark 2. We know that the symmetric random walk on $\mathbb{Z}^d$ is null recurrent for $d \leq 2$ and transient for $d \geq 3$. For both cases, the asymptotic behavior of Z_n is similar upto centering and scaling. Furthermore, notice from (2.10) that even for the general case, the centering and scaling vary only upto multiplicative constants depending only on the law of X_1.

Remark 3. On $\mathbb{Z}$ the simple symmetric random walk is null recurrent while the right shift walk is transient. For both cases, the asymptotic behavior of Z_n is similar upto centering and scaling.

2.2 Weak Convergence of the Random Configuration

Let $\mathcal{M}_1$ be the space of probability measures on $\mathbb{R}^d$, $d \geq 1$ endowed with the topology of weak convergence. For $\omega \in \Omega$, let $\Lambda_n(\omega) \in \mathcal{M}_1$ be the random probability measure corresponding to the random probability vector $\frac{U_n(\omega)}{n+1}$.

Theorem 2.4. *For $\omega \in \Omega$, let $\Lambda_n^{cs}(\omega)(A) = \Lambda_n(\omega)\left(\sqrt{\log n} A \Sigma^{-1/2} + \mu \log n\right)$. Then, as $n \to \infty$,*

$$\Lambda_n^{cs} \xrightarrow{p} \Phi_d \text{ on } \mathcal{M}_1, \tag{2.11}$$

where $\Phi_d(A) = \frac{1}{\left(\sqrt{2\pi}\right)^d} \int_A e^{-\frac{xx^t}{2}} \, \mathrm{d}t$.

Remark 4. From Theorem 2.4 we can conclude that given any subsequence $\{n_k\}_{k=1}^{\infty}$ there exists a further subsequence $\{n_{k_j}\}_{j=1}^{\infty}$ such that as $j \to \infty$,

$$\Lambda_{n_{k_j}}^{cs} \xrightarrow{w} \Phi_d \text{ almost surely for all } \omega.$$

2.3 Local Limit Theorem Type Results for the Expected Configuration

The random variable Z_n corresponding to the probability vector $\frac{1}{n+1}\left(\mathbb{E}\left[U_{n,v}\right]\right)_{v \in \mathbb{Z}}$ admits local limit type results. In this section we present theorems demonstrating this. We will first present the local property for $d = 1$ and then show the same for higher dimensions.

2.3.1 Local Limit Type Results for One Dimension

In this subsection, we present the local limit theorems for urns with colors indexed by $\mathbb{Z}$. For $j \in \mathbb{N}$, X_j is a lattice random variable. Let $\mathbb{P}\left(X_j \in a + h\mathbb{Z}\right) = 1$, where $a \in \mathbb{R}$ and $h > 0$ is the span for X_j. Define $\mathcal{L}_n^{(1)} := \{x \colon x = \frac{n}{\sigma\sqrt{\log n}} a - \frac{\mu}{\sigma}\sqrt{\log n} + \frac{h}{\sigma\sqrt{\log n}}\mathbb{Z}\}$ where $\mu = \mathbb{E}\left[X_1\right]$ and $\sigma^2 = \mathbb{E}\left[X_1^2\right]$.

Theorem 2.5. *Consider the urn associated with the Markov chain of the random walk $\{S_n\}_{n \geq 0}$ and let Z_n be the color of the randomly selected ball at the $(n+1)^{th}$ trial. We further assume that $\mathbb{P}\left[X_1 = 0\right] > 0$. Then, as $n \to \infty$*

$$\sup_{x \in \mathcal{L}_n^{(1)}} \left| \sigma \frac{\sqrt{\log n}}{h} \mathbb{P}\left(\frac{Z_n - \mu \log n}{\sigma\sqrt{\log n}} = x\right) - \phi(x) \right| \longrightarrow 0 \tag{2.12}$$

where $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$.

Remark 5. The assumption $\mathbb{P}\left[X_1 = 0\right] > 0$ implies that for each $j \in \mathbb{N}$, X_j and $I_j X_j$ are supported on the same set.

Let $\mathbb{P}[X_1 = 0] = 0$ and $\tilde{h}$ be the span for X_j. Let $\mathbb{P}(I_j X_j \in a + h\mathbb{Z}) = 1$, where $a \in \mathbb{R}$ and $h > 0$ is the span for $I_j X_j$. It is easy to see that $h \leq \tilde{h}$. Define $\mathcal{L}_n^{(1)} := \{x\colon x = \frac{n}{\sigma\sqrt{\log n}} a - \frac{\mu}{\sigma}\sqrt{\log n} + \frac{h}{\sigma\sqrt{\log n}}\mathbb{Z}\}$ where $\mu = \mathbb{E}[X_1]$ and $\sigma^2 = \mathbb{E}\left[X_1^2\right]$.

Theorem 2.6. *Consider the urn associated with the Markov chain of the random walk $\{S_n\}_{n\geq 0}$ and let Z_n be the color of the randomly selected ball at the $(n+1)^{th}$ trial. We further assume that $\tilde{h} < 2h$. Then, as $n \to \infty$*

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

where $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$.

In the next theorem we present the local limit theorem when the urn is associated with the simple symmetric random walk which is not covered by either Theorem 2.5 or Theorem 2.6. Let $\{X_i\}_{i\geq 1}$ be an i.i.d. sequence of ± 1 Bernoulli random variables, such that $\mathbb{P}[X_1 = 1] = \mathbb{P}[X_1 = -1] = \frac{1}{2}$. In this case, the span of X_1 is 2. The random variables $I_j X_j$ is supported on the set $\{0, 1, -1\}$ and the span of $I_j X_j$ is 1. Therefore, in this X_j and $I_j X_j$ are supported on different sets. Let $\mathbb{P}(I_j X_j \in a + \mathbb{Z}) = 1$. One possible choice of a is 0. We define $\mathcal{L}_n^{(1)} := \{x\colon x = \frac{1}{\sqrt{\log n}}\mathbb{Z}\}$.

Theorem 2.7. *Consider the urn model associated with the simple symmetric random walk on $\mathbb{Z}$ and let Z_n be the color of the randomly selected ball at the $(n+1)^{th}$ trial. Then, as $n \to \infty$*

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

where $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$.

2.3.2 Local Limit Type Results for Higher Dimensions

Throughout this subsection we assume $d \geq 2$. For $j \in \mathbb{N}$, X_j is a lattice random vector with $\mathcal{L}$ as the minimal lattice (see pages 226 – 227 of [3] for a formal definition). We also define $l = |\det(\mathcal{L})|$ where $\det(\mathcal{L})$ (see the pages 228 – 229 of [3] for more details). Let $\mathbb{P}(X_1 \in x_0 + \mathcal{L}) = 1$, where $x_0 \in \mathbb{R}^d$. Let us define $\mathcal{L}_n^{(d)} = \{x\colon x = \frac{n}{\sqrt{\log n}} x_0 \Sigma^{-1/2} - \sqrt{\log n}\, \mu\, \Sigma^{-1/2} + \frac{1}{\sqrt{\log n}} \mathcal{L}\Sigma^{-1/2}\}$ where $\mu = \left(\mathbb{E}[X_1^{(1)}], \mathbb{E}[X_1^{(2)}], \dots \mathbb{E}[X_1^{(d)}]\right)$ and $\Sigma = ((\sigma_{ij}))_{1\leq i,j\leq d}$ with $\sigma_{ij} = \mathbb{E}[X_1^{(i)} X_1^{(j)}]$.

Theorem 2.8. *Consider the urn associated with the Markov chain of the random walk $\{S_n\}_{n\geq 0}$ and let Z_n be the color of the randomly selected ball at the $(n+1)^{th}$ trial. We further assume that $\mathbb{P}[X_1 = \mathbf{0}] > 0$, where $\mathbf{0}$ is the origin in $\mathbb{Z}^d$. Then, as $n \to \infty$*

$$\sup_{x \in \mathcal{L}_n^{(d)}} \left| \frac{det(\Sigma^{1/2}) \left(\sqrt{\log n}\right)^d}{l} \mathbb{P}\left(\frac{Z_n - \mu \log n}{\sqrt{\log n}} \Sigma^{-1/2} = x \right) - \phi_d(x) \right| \longrightarrow 0 \tag{2.13}$$

where $\phi_d(x) = \frac{1}{(2\pi)^{d/2}} e^{-\frac{xx^t}{2}}$.

The assumption $\mathbb{P}[X_1 = \mathbf{0}] > 0$ can be removed, for at least some cases, though we do not know the full generality under which the local limit theorem holds. We consider a special case of interest which is not covered by Theorem 2.8, namely the simple symmetric random walk on $\mathbb{Z}^d$.

Let $\{X_i\}_{i\geq 1}$ be i.i.d distributed as $\mathbb{P}[X_1 = \pm e_i] = \frac{1}{2d}$ where e_i is the d-dimensional row vector with 1 at the i^{th} coordinate and 0 elsewhere. Note that in this case, the mean increment vector $\mu = \mathbf{0}$ and the variance-covariance matrix $\Sigma = d^{-1}\mathbb{I}_d$. Also observe that $\mathbb{P}[X_1 = \mathbf{0}] = 0$. In the next theorem, we present the local limit behavior of the urn associated with the random walk $S_n = \sum_{j=1}^n X_j$. For $j \in \mathbb{N}$, $I_j X_j$ is a lattice random vector with minimal lattice as $\mathbb{Z}^d$ and we may choose of $x_0 = \mathbf{0}$, that is, $\mathbb{P}\left(I_j X_j \in \mathbb{Z}^d\right) = 1$. We define $\mathcal{L}_n^{(d)} = \{x\colon x = \frac{\sqrt{d}}{\sqrt{\log n}}\mathbb{Z}^d\}$.

Theorem 2.9. *Consider the urn model associated with the Markov chain of the simple symmetric random walk on $\mathbb{Z}^d$ and let Z_n be the color of the randomly selected ball at the $(n+1)^{th}$ trial. Then, as $n \to \infty$*

$$\sup_{x\in\mathcal{L}_n^{(d)}} \left| (d)^{\frac{d}{2}} \left(\sqrt{\log n}\right)^d \mathbb{P}\left(\frac{\sqrt{d}}{\sqrt{\log n}} Z_n = x\right) - \phi_d(x) \right| \longrightarrow 0, \tag{2.14}$$

where $\phi_d(x) = \frac{1}{(2\pi)^{d/2}} e^{-\frac{xx^t}{2}}$.

Note that here the the minimal lattice is $\mathbb{Z}^d$ and so its determinant is 1 (see page $228 - 229$ of [3]).

3 Auxiliary Results

In the following section, we present two important results which we will need to prove Theorems 2.1 and 2.4. Throughout this section we assume that the initial configuration $U_0 = (U_{0,v})_{v\in\mathbb{Z}^d}$ is such that $U_{0,v} = 0$ except for finitely many v.

Define $\Pi_n(z) = \prod_{j=1}^{n}\left(1 + \frac{z}{j}\right)$ for $z \in \mathbb{C}$. It is known from Euler product formula for gamma function which is also known as Gauss's formula (see page 178 of [7]) that

$$\lim_{n\to\infty} \frac{\Pi_n(z)}{n^z}\Gamma(z+1) = 1 \tag{3.15}$$

uniformly on compact subsets of $\mathbb{C} \setminus \{0, -1, -2, \ldots\}$.

For $\lambda \in \mathbb{R}^d$, let $e(\lambda) = \sum_{v\in B} e^{\langle\lambda,v\rangle} p(v)$ where $\langle u, v\rangle = uv^t$ for $u, v \in \mathbb{R}^d$ denote the moment generating function for X_1. It is easy to note that $e(\lambda)$ is an eigenvalue of R corresponding to the right eigenvector $x(\lambda) = \left(e^{\langle\lambda,v\rangle}\right)^t_{v\in\mathbb{Z}^d}$. Let $\mathcal{F}_n = \sigma(U_j\colon 0 \leq j \leq n)$. Define

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

From (1.5) we get,

$$U_{n+1}x(\lambda) = U_n x(\lambda) + I_{n+1} R x(\lambda)$$

Thus,

$$\mathbb{E}\left[U_{n+1}x(\lambda)\,\middle|\,\mathcal{F}_n\right] = U_n x(\lambda) + e(\lambda)\mathbb{E}\left[I_{n+1}x(\lambda)\,\middle|\,\mathcal{F}_n\right] = \left(1 + \tfrac{e(\lambda)}{n+1}\right) U_n x(\lambda).$$

Therefore, $\overline{M}_n(\lambda)$ is a non-negative martingale for every $\lambda \in \mathbb{R}^d$ and $\mathbb{E}\left[\overline{M}_n(\lambda)\right] = \overline{M}_0(\lambda)$. We will make use of the martingales $\overline{M}_n(\lambda)$ in the proof of the next theorem in which we present a representation of the marginal of Z_n in terms of the increments $X_i, i \in \mathbb{N}$.

Theorem 3.1. *Let Z_n be the color of the randomly selected ball in the $(n+1)^{th}$ draw, then for all $n \geq 1$*

$$Z_n \stackrel{d}{=} Z_0 + \sum_{j=1}^{n} I_j X_j. \tag{3.16}$$

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

Proof. We have already noted that $e(\lambda)$ denotes the moment generating function for X_1. The eigenvalues and the corresponding right eigenvectors for R are given by $e(\lambda)$ and $x(\lambda)$. If Z_n denotes the color of the randomly selected ball at the $(n+1)^{\text{th}}$ trial, then for $\lambda \in \mathbb{R}^d$, the moment generating function of Z_n is given by

$$\begin{aligned} \frac{1}{n+1}\sum_{v\in\mathbb{Z}^d} e^{\langle\lambda,v\rangle}\mathbb{E}\left[U_{n,v}\right] &= \frac{\Pi_n\left(e(\lambda)\right)}{n+1}\mathbb{E}\left[\overline{M}_n(\lambda)\right] \\ &= \frac{\Pi_n\left(e(\lambda)\right)}{n+1}\overline{M}_0(\lambda). \end{aligned} \tag{3.17}$$

We note that

$$\overline{M}_0(\lambda)\frac{1}{n+1}\Pi_n\left(e(\lambda)\right) = \overline{M}_0(\lambda)\prod_{j=1}^{n}\left(1 - \frac{1}{j+1} + \frac{e(\lambda)}{j+1}\right).$$

Therefore, from (3.17) it follows that

$$Z_n \stackrel{d}{=} Z_0 + \textstyle\sum_{j=1}^{n} I_j X_j,$$

where Z_0 is a random vector taking values in $\mathbb{Z}^d$ according to the probability vector U_0 and is independent of $\{I_j\}_{j\geq 1}$ and $\{X_j\}_{j\geq 1}$.

∎

Theorem 3.2. *There exists $\delta > 0$ such that*

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

Proof. From (1.5), we obtain

$$\mathbb{E}\left[\left(U_{n+1}x(\lambda)\right)^2 \Big| \mathcal{F}_n\right] = \left(U_n x(\lambda)\right)^2 + 2e(\lambda)\,U_n x(\lambda)\,\mathbb{E}\left[I_{n+1}x(\lambda)\Big|\mathcal{F}_n\right] + e^2(\lambda)\,\mathbb{E}\left[\left(I_{n+1}x(\lambda)\right)^2\Big|\mathcal{F}_n\right]$$

It is easy to see that

$$\mathbb{E}\left[I_{n+1}x(\lambda)\Big|\mathcal{F}_n\right] = \frac{1}{n+1}U_n x(\lambda) \text{ and } \mathbb{E}\left[\left(I_{n+1}x(\lambda)\right)^2\Big|\mathcal{F}_n\right] = \frac{1}{n+1}U_n x(2\lambda).$$

Therefore, we get the recursion

$$\mathbb{E}\left[\left(U_{n+1}x\left(\lambda\right)\right)^2\right] = \left(1+\frac{2e\left(\lambda\right)}{n+1}\right)\mathbb{E}\left[\left(U_n x\left(\lambda\right)\right)^2\right] + \frac{e^2\left(\lambda\right)}{n+1}\mathbb{E}\left[U_n x\left(2\lambda\right)\right]. \tag{3.18}$$

Dividing both sides of (3.18) by $\Pi_{n+1}^2\left(\lambda\right)$,

$$\mathbb{E}\left[\overline{M}_{n+1}^2\left(\lambda\right)\right] = \frac{\left(1+\frac{2e(\lambda)}{n+1}\right)}{\left(1+\frac{e(\lambda)}{n+1}\right)^2}\mathbb{E}\left[\overline{M}_n^2\left(\lambda\right)\right] + \frac{e^2\left(\lambda\right)}{n+1}\frac{\mathbb{E}\left[U_n x\left(2\lambda\right)\right]}{\Pi_{n+1}^2\left(\lambda\right)}. \tag{3.19}$$

$\overline{M}_n\left(2\lambda\right)$ being a martingale, we obtain $\mathbb{E}\left[U_n x\left(2\lambda\right)\right] = \Pi_n\left(e\left(2\lambda\right)\right)\overline{M}_0\left(2\lambda\right)$. Therefore from (3.19), we get

$$\mathbb{E}\left[\overline{M}_n^2\left(\lambda\right)\right] = \frac{\Pi_n\left(2e\left(\lambda\right)\right)}{\Pi_n\left(e\left(\lambda\right)\right)^2}\overline{M}_0^2\left(\lambda\right) + \sum_{k=1}^{n}\frac{e^2\left(\lambda\right)}{k}\left\{\prod_{j>k}^{n}\frac{\left(1+\frac{2e(\lambda)}{j}\right)}{\left(1+\frac{e(\lambda)}{j}\right)^2}\right\}\frac{\Pi_{k-1}\left(e\left(2\lambda\right)\right)}{\Pi_k^2\left(e\left(\lambda\right)\right)}\overline{M}_0\left(2\lambda\right). \tag{3.20}$$

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

Therefore,

$$\mathbb{E}\left[\overline{M}_n^2\left(\lambda\right)\right] \leq \overline{M}_0^2\left(\lambda\right) + e^2\left(\lambda\right)\overline{M}_0\left(2\lambda\right)\sum_{k=1}^{n}\frac{1}{k}\frac{\Pi_{k-1}\left(e\left(2\lambda\right)\right)}{\Pi_k^2\left(e\left(\lambda\right)\right)}. \tag{3.21}$$

Using (3.15), we know that

$$\Pi_n\left(e\left(2\lambda\right)\right) \sim \frac{n^{e(2\lambda)}}{\Gamma\left(e\left(2\lambda\right)+1\right)}$$

and

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

It is easy to note that $e\left(\lambda\right) > 0$ for all $\lambda \in \mathbb{R}^d$ and $e\left(\lambda\right)$ is continuous as a function of λ and $e\left(\lambda\right) > 0$. So given $\eta > 0$ there exists $0 < K_1, K_2 < \infty$, such that for all $\lambda \in [-\eta, \eta]^d$, $K_1 \leq e\left(\lambda\right) \leq K_2$. Since the convergence in (3.15) 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 $\lambda \in [-\eta, \eta]^d$,

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

Recall that $e\left(\lambda\right) = \sum_{v\in B} e^{\langle\lambda,v\rangle}p(v)$. Since $|B| < \infty$, so we can choose a $\delta_0 > 0$ such that for every $\lambda \in [-\delta_0, \delta_0]^d$, $2e\left(\lambda\right) - e\left(2\lambda\right) > 0$. Choose $\delta = \min\{\eta, \delta_0\}$. Since $2e\left(\lambda\right) - e\left(2\lambda\right)$ is continuous as a function of λ, 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 λ are bounded for $\lambda \in [-\delta, \delta]^d$. Choose $N = \max\{N_1, N_2\}$. From (3.21) we obtain for all $n \geq N$

$$\mathbb{E}\left[\overline{M}_n^2(\lambda)\right] \leq \overline{M}_0^2(\lambda) + C_1 \sum_{k=1}^{N} \frac{1}{k} \frac{\Pi_{k-1}(e(2\lambda))}{\Pi_k^2(e(\lambda))} + \epsilon \tag{3.22}$$

for an appropriate positive constant C_1.

$\sum_{k=1}^{N} \frac{1}{k} \frac{\Pi_{k-1}(e(2\lambda))}{\Pi_k^2(e(\lambda))}$ and $\overline{M}_0^2(\lambda)$ being continuous as functions of λ, are bounded for $\lambda \in [-\delta, \delta]^d$. Therefore, from (3.22) 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 that

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

□

Remark 6. From Theorem 3.2 we can conclude that there exists a random variable $\overline{M}(\lambda)$ such that as $n \to \infty$

$$\overline{M}_n(\lambda) \longrightarrow \overline{M}(\lambda)$$

almost surely $\mathbb{P}$ and in $L_2(\mathbb{P})$.

4 Proofs of the Main Results

In this section we provide the proofs of the main results. Some of the notations used here have been defined in the previous section.

4.1 Proofs for the Expected Configuration

This subsection contains the proofs for Theorems 2.1.

Proof of Theorem 2.1. Firstly we note by Theorem 3.1 without loss of generality, we may assume that the process starts with a single ball of color indexed by $\mathbf{0}$, in other words $Z_0 = \mathbf{0}$ in (3.16), then

$$Z_n \stackrel{d}{=} \sum_{j=1}^{n} I_j X_j.$$

It is easy to see that $\mathbb{E}\left[\sum_{j=1}^{n} I_j X_j\right] - \mu \log n = \sum_{j=1}^{n} \frac{1}{j}\mu - \mu \log n \longrightarrow \gamma\mu$ where γ is the Euler's constant.

Case I: $d = 1$**.** Let $s_n^2 = Var\left[\sum_{j=1}^{n} I_j X_j\right]$. It is easy to note that $s_n^2 = \sum_{j=1}^{n} \frac{1}{j+1}\mathbb{E}\left[X_1^2\right] - \frac{\mu^2}{(j+1)^2} \sim \sigma^2 \log n$. Since $|B| < \infty$, given $\epsilon > 0$, we have $\frac{1}{s_n^2}\sum_{j=1}^{n} \mathbb{E}\left[I_j X_j^2 1_{\{I_j X_j > \epsilon s_n\}}\right] \longrightarrow 0$ as $n \to \infty$. Therefore, by the Lindeberg Feller Central Limit theorem, we get as $n \to \infty$

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

This completes the proof for $d = 1$.

Case II: $d \geq 2$**.** Let $\Sigma_n = \left[\sigma_{k,l}(n)\right]_{d\times d}$ 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 for all $k, l \in \{1, 2, \ldots d\}$ as $n \to \infty$

$$\frac{\sigma_{k,l}(n)}{(\log n)\sigma_{k,l}} \longrightarrow 1.$$

Therefore for every $\theta \in \mathbb{R}^d$, by Lindeberg Feller Central Limit Theorem in one dimension,

$$\frac{\langle \theta, \sum_{j=1}^{n} I_j X_j\rangle - \langle \theta, \mu \log n\rangle}{\sqrt{\log n}\left(\theta \Sigma \theta^t\right)^{1/2}} \stackrel{d}{\longrightarrow} N(0,1) \text{ as } n \to \infty.$$

Therefore by Cramer-Wold device, it follows that as $n \to \infty$

$$\frac{\sum_{j=1}^{n} I_j X_j - \mu \log n}{\sqrt{\log n}} \stackrel{d}{\longrightarrow} N_d\left(\mathbf{0}, \Sigma\right).$$

Hence, as $n \to \infty$

$$\frac{Z_n - \mu \log n}{\sqrt{\log n}} \stackrel{d}{\longrightarrow} N_d\left(\mathbf{0}, \Sigma\right).$$

This completes the proof. ☐

4.2 Proofs for Random Configuration

In this subsection we will present the proof of Theorem 2.4. To do so, we present a lemma which we will need in the proof of Theorem 2.4.

Lemma 4.1. *Let δ be as in Theorem 3.2, then for every $\lambda \in [-\delta, \delta]^d$*

$$\overline{M}_n\left(\frac{\lambda}{\sqrt{\log n}}\right) \xrightarrow{p} 1 \tag{4.23}$$

as $n \to \infty$.

Proof. Without loss of generality, we assume that the process starts with a single ball of color indexed by $\mathbf{0}$. From equation (3.20) we get

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

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

$$\mathbb{E}\left[\overline{M}_n^2(\lambda_n)\right] = \frac{\Pi_n(2e(\lambda_n))}{\Pi_n^2(e(\lambda_n))} + \frac{\Pi_n(2e(\lambda_n))}{\Pi_n^2(e(\lambda_n))}\sum_{k=1}^{n}\frac{e^2(\lambda_n)}{k}\frac{\Pi_{k-1}(e(2\lambda_n))}{\Pi_k(2e(\lambda_n))} \tag{4.24}$$

Since the convergence in formula (3.15) is uniform on compact sets of $[0, \infty)$, we observe that for $\lambda \in [-\delta, \delta]^d$

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

We observe that $\lim_{n\to\infty} e(\lambda_n) = 1$ and $\lim_{n\to\infty}\frac{\Pi_n(2e(\lambda_n))}{\Pi_n^2(e(\lambda_n))}\frac{e^2(\lambda_n)}{k}\frac{\Pi_{k-1}(e(2\lambda_n))}{\Pi_k(2e(\lambda_n))} = \frac{1}{2}\frac{1}{k}\frac{\Pi_{k-1}(1)}{\Pi_k(2)}$. Since $\sup_{n\geq 1}\sup_{\lambda\in[-\delta,\delta]^d}\mathbb{E}\left[\overline{M}_n^2(\lambda)\right] < \infty$, by dominated convergence theorem, we get

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

Therefore, from (4.24) we obtain

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

Hence, it follows that

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

□

Proof of Theorem 2.4. Without loss of generality, we may assume that the urn process starts with a single ball of color indexed by $\mathbf{0}$. Let Λ_n be the random probability measure on $\mathbb{R}^d$ corresponding to the random probability vector $\frac{1}{n+1}U_n$. For $\lambda \in \mathbb{R}^d$ the corresponding random moment generating function is given by

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

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

$$\frac{1}{n+1}e^{-\langle \lambda, \mu\sqrt{\log n}\rangle} U_n x\left(\frac{\lambda}{\sqrt{\log n}}\right) = \frac{1}{n+1}e^{-\langle \lambda, \mu\sqrt{\log n}\rangle}\overline{M}_n\left(\frac{\lambda}{\sqrt{\log n}}\right)\Pi_n\left(e(\frac{\lambda}{\sqrt{\log n}})\right).$$

To show (2.11) 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 $j \to \infty$

$$\frac{e^{-\langle \lambda, \mu\sqrt{\log n_{k_j}}\rangle}}{n_{k_j}+1}\overline{M}_{n_{k_j}}\left(\frac{\lambda}{\sqrt{\log n_{k_j}}}\right)\Pi_n\left(e\left(\frac{\lambda}{\sqrt{\log n_{k_j}}}\right)\right) \longrightarrow e^{\frac{\lambda\Sigma\lambda^t}{2}} \text{ for all } \lambda \in [-\delta, \delta]^d \text{ almost surely} \tag{4.25}$$

where δ is as in Lemma 4.1. From Theorem 2.1 we know that $\frac{Z_n - \log n\mu}{\sqrt{\log n}} \xrightarrow{d} N_d\left(\mathbf{0}, \mathbb{I}_d\right)$. Therefore using (3.17) as $n \to \infty$ we obtain,

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

Using Theorem 6.1 it is enough to show (4.25) 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 $j \to \infty$

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

From Lemma 4.1 we know that for all $\lambda \in [-\delta, \delta]^d$

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

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 $\lambda \in \mathbb{Q}^d \cap [-\delta, \delta]^d$

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

This completes the proof. ☐

Remark 7. It is worth noting that the proofs of Theorems 2.1 and 2.4 go through if we assume U_0 to be non random probability vector such that there exists $r > 0$ such that for all $\lambda \in \{\lambda \colon \|\lambda\| < r\}$, $\sum_{v\in\mathbb{Z}^d} e^{\langle \lambda, v\rangle} U_{0,v} < \infty$ and $\lim_{n\to\infty}\sum_{v\in\mathbb{Z}^d} e^{\frac{\langle \lambda, v\rangle}{\sqrt{\log n}}} U_{0,v}$ exists finitely.

4.3 Proofs of the Local Limit Type Results

In this section, we present the proofs for the local limit type results for Z_n. As before, we present the proof for urn with colors indexed by $\mathbb{Z}$ first and then proceed towards the proofs for $\mathbb{Z}^d$, $d \geq 2$.

4.3.1 Proof for the Local Limit Theorems for d=1

Proof of Theorem 2.5. Without loss of generality we may assume $\mu = 0$ and $\sigma = 1$. We further assume that the process begins with a single ball of color indexed by 0. From Theorem 3.1, we know that $Z_n \stackrel{d}{=} \sum_{k=1}^{n} I_j X_j$. X_j is a lattice random variable, therefore $I_j X_j$ is also so. Furthermore $0 \in B$, therefore $I_j X_j$ and X_j have the same lattice structure. Therefore Z_n is a lattice random variable with lattice $\mathcal{L}_n^{(1)}$. Applying Fourrier inversion formula, for all $x \in \mathcal{L}_n^{(1)}$ we obtain

$$\mathbb{P}\left(\frac{Z_n}{\sqrt{\log n}} = x\right) = \frac{h}{2\pi\sqrt{\log n}} \int_{-\frac{\pi\sqrt{\log n}}{h}}^{\frac{\pi\sqrt{\log n}}{h}} e^{-itx}\psi_n(t)\,\mathrm{d}t$$

where $\psi_n(t) = \mathbb{E}\left[e^{it\frac{Z_n}{\sqrt{\log n}}}\right]$. Without loss of generality, we may assume $h = 1$. Also by Fourrier inversion formula, for all $x \in \mathbb{R}$

$$\phi(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}e^{\frac{-t^2}{2}}\,\mathrm{d}t.$$

Given $\epsilon > 0$, there exists N large enough such that for all $n \geq N$

$$\begin{aligned}\left|\sqrt{\log n}\,\mathbb{P}\left(\frac{Z_n}{\sqrt{\log n}} = x\right) - \phi(x)\right| &\leq \int_{-\pi\sqrt{\log n}}^{\pi\sqrt{\log n}}\left|\psi_n(t) - e^{\frac{-t^2}{2}}\right|\mathrm{d}t + 2\int_{[-\pi\sqrt{\log n},\pi\sqrt{\log n}]^c}\phi(t)\,\mathrm{d}t \\ &= \int_{-\pi\sqrt{\log n}}^{\pi\sqrt{\log n}}\left|\psi_n(t) - e^{\frac{-t^2}{2}}\right|\mathrm{d}t + \epsilon.\end{aligned}$$

Given $M > 0$, we can write for all n large enough

$$\begin{aligned}\int_{-\pi\sqrt{\log n}}^{\pi\sqrt{\log n}}\left|\psi_n(t) - e^{\frac{-t^2}{2}}\right|\mathrm{d}t \leq \int_{-M}^{M}\left|\psi_n(t) - e^{\frac{-t^2}{2}}\right|\mathrm{d}t + \int_{M}^{\pi\sqrt{\log n}}\left|\psi_n(t)\right|\mathrm{d}t \\ +2\int_{M}^{\pi\sqrt{\log n}} e^{\frac{-t^2}{2}}\,\mathrm{d}t. \qquad (4.26)\end{aligned}$$

Given $\epsilon > 0$ we choose an $M > 0$ such that

$$\int_{[-M,M]^c} e^{\frac{-t^2}{2}}\,\mathrm{d}t < \epsilon.$$

Therefore,

$$\int_{M}^{\pi\sqrt{\log n}} e^{-\frac{t^2}{2}}\,\mathrm{d}t \leq \int_{[-M,M]^c} e^{\frac{-t^2}{2}}\,\mathrm{d}t < \epsilon. \qquad (4.27)$$

We know from Theorem 2.1 that as $n \to \infty$, $\frac{Z_n}{\sqrt{\log n}} \xrightarrow{d} N(0,1)$. Hence for all $t \in \mathbb{R}$, $\psi_n(t) \longrightarrow e^{\frac{-t^2}{2}}$. Therefore, for the chosen $M > 0$ by bounded convergence theorem we get as $n \to \infty$

$$\int_{-M}^{M} \left| \psi_n(t) - e^{\frac{-t^2}{2}} \right| \mathrm{d}t \longrightarrow 0.$$

Let

$$I(n) = \int_{M}^{\pi\sqrt{\log n}} \left| \psi_n(t) \right| \mathrm{d}t.$$

We will show that as $n \to \infty$, $I(n) \longrightarrow 0$. Since $Z_n \stackrel{d}{=} \sum_{j=1}^{n} I_j X_j$, therefore

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

where $e\,(it) = \mathbb{E}\left[e^{itX_1}\right]$. Therefore,

$$\psi_n(t) = \mathbb{E}\left[e^{it\frac{Z_n}{\sqrt{\log n}}}\right] = \frac{1}{n+1} \Pi_n\left(e(it/\sqrt{\log n})\right).$$

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

$$I(n) = \sqrt{\log n} \int_{M/\sqrt{\log n}}^{\pi} \left| \psi_n\left(w\sqrt{\log n}\right) \right| \mathrm{d}w. \tag{4.28}$$

Now there exists $\delta > 0$ such that for all $t \in (0, \delta)$

$$\left| e\,(it) \right| \le 1 - \frac{t^2}{4}. \tag{4.29}$$

Therefore using the inequality $1 - x \le e^{-x}$, we obtain $1 - \frac{1}{j+1} + \frac{|e(it)|}{j+1} \le e^{-\frac{1}{j+1}\frac{t^2}{4}}$. Hence, for all $t \in (0, \delta)$

$$\frac{1}{n+1} \left| \Pi_n\left(e\,(it)\right) \right| \le e^{-\frac{t^2}{4} \sum_{j=1}^{n} \frac{1}{j+1}}. \tag{4.30}$$

We observe from (4.28) that we can write

$$I(n) = \sqrt{\log n} \int_{M/\sqrt{\log n}}^{\delta} \left| \psi_n\left(w\sqrt{\log n}\right) \right| \mathrm{d}w + \sqrt{\log n} \int_{\delta}^{\pi} \left| \psi_n\left(w\sqrt{\log n}\right) \right| \mathrm{d}w.$$

Let us write

$$I_1(n) = \sqrt{\log n} \int_{M/\sqrt{\log n}}^{\delta} \left| \psi_n\left(w\sqrt{\log n}\right) \right| \mathrm{d}w$$

and

$$I_2(n) = \sqrt{\log n} \int_{\delta}^{\pi} \left| \psi_n \left(w \sqrt{\log n} \right) \right| \, \mathrm{d}w.$$

From (4.30) we have $I_1(n) \longrightarrow 0$ as $n \to \infty$.

As 0 is in the support of X_j, therefore $I_j X_j$ and X_j will have the same lattices. Therefore for all $t \in [\delta, 2\pi)$, $|e\,(it)| < 1$ and the characteristic function being continuous in t, there exists $0 < \eta < 1$ such that $|e\,(it)| \leq \eta$ for all $t \in [\delta, \pi]$. Therefore

$$1 - \frac{1}{j+1} + \frac{|e\,(it)|}{j+1} \leq 1 - \frac{1}{j+1} + \frac{\eta}{j+1} \leq e^{-\frac{1-\eta}{j+1}}.$$

It follows that

$$\frac{1}{n+1} |\Pi_n\left(e\,(it)\right)| \leq e^{-\sum\limits_{j=1}^{n} \frac{1-\eta}{j+1}} \leq C_2 e^{-(1-\eta) \log n}$$

where C_2 is some positive constant. So as $n \to \infty$

$$I_2(n) \leq C_2 e^{-(1-\eta) \log n} \left(\pi - \delta\right) \sqrt{\log n} \longrightarrow 0.$$

So combining the facts that $I_1(n) \longrightarrow 0$, $I_2(n) \longrightarrow 0$ as $n \to \infty$ and from (4.26) and (4.27), the proof is complete. $\square$

Proof of Theorem 2.6. Without loss of generality we may assume that $\mu = 0$, $\sigma = 1$ and the process begins with a single ball of color indexed by 0. The proof is similar to the proof of Theorem 2.5 and therefore we omit certain details. Since the span of $I_j X_j$ is h, for all $x \in \mathcal{L}_n^{(1)}$ we obtain by Fourrier inversion formula

$$\mathbb{P}\left(\frac{Z_n}{\sqrt{\log n}} = x \right) = \frac{h}{2\pi \sqrt{\log n}} \int_{-\frac{\pi\sqrt{\log n}}{h}}^{\frac{\pi\sqrt{\log n}}{h}} e^{-itx} \psi_n(t) \, \mathrm{d}t$$

where $\psi_n\,(t) = \mathbb{E}\left[e^{it \frac{Z_n}{\sqrt{\log n}}} \right]$. Without loss of generality, we may assume $h = 1$. Also by Fourrier inversion formula, for all $x \in \mathbb{R}$

$$\phi(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} e^{\frac{-t^2}{2}} \, \mathrm{d}t.$$

The bounds for $\left| \sqrt{\log n} \mathbb{P}\left(\frac{Z_n}{\sqrt{\log n}} = x \right) - \phi(x) \right|$ are similar to that in the proof of Theorem 2.5 except for that of $I_2(n)$ where

$$I_2(n) = \sqrt{\log n} \int_{\delta}^{\pi} \left| \psi_n \left(w \sqrt{\log n} \right) \right| \, \mathrm{d}w$$

and δ is chosen as in (4.29). We have to show

$$I_2(n) \longrightarrow 0 \text{ as } n \to \infty.$$

The span of X_1 being $\tilde{h}$, for all $t \in \left[\delta, \frac{2\pi}{\tilde{h}}\right)$, $|e\,(it)| < 1$. Since $\tilde{h} < 2$ and the characteristic function is continuous in t, there exists $0 < \eta < 1$ such that $|e\,(it)| \leq \eta$ for all $t \in [\delta, \pi] \subset \left[\delta, \frac{2\pi}{\tilde{h}}\right)$. Therefore $I_2(n) \longrightarrow 0$ as $n \to \infty$ using the similar bounds as in the proof of Theorem 2.5. $\square$

Proof of Theorem 2.7. In this case, $\mu = 0$ and $\sigma = 1$. Without loss of generality, we may assume that the process begins with a single ball of color indexed by 0. In this case, the span of I_jX_j is 1 and $\mathbb{P}(I_jX_j \in \mathbb{Z}) = 1$. Therefore by the Fourrier inversion formula, for all $x \in \mathcal{L}_n^{(1)} = \{x\colon x = \frac{1}{\sqrt{\log n}}\mathbb{Z}\}$ we obtain

$$\mathbb{P}\left(\frac{Z_n}{\sqrt{\log n}} = x\right) = \frac{1}{2\pi\sqrt{\log n}} \int_{-\pi\sqrt{\log n}}^{\pi\sqrt{\log n}} e^{-itx}\psi_n(t)\,\mathrm{d}t$$

where $\psi_n(t) = \mathbb{E}\left[e^{it\frac{Z_n}{\sqrt{\log n}}}\right]$. Furthermore, by Fourrier inversion formula, for all $x \in \mathbb{R}$

$$\phi(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}e^{\frac{-t^2}{2}}\,\mathrm{d}t.$$

Like in the case of Theorem 2.6, the proof of this theorem is also very similar to that of Theorem 2.5. The bounds for $\left|\sqrt{\log n}\mathbb{P}\left(\frac{Z_n}{\sqrt{\log n}} = x\right) - \phi(x)\right|$ are similar to that in the proof of Theorem 2.5 except for that of $I_2(n)$ where

$$I_2(n) = \sqrt{\log n}\int_{\delta}^{\pi}\left|\psi_n\left(w\sqrt{\log n}\right)\right|\,\mathrm{d}w$$

and δ is chosen as in (4.29). To show that $I_2(n) \longrightarrow 0$ as $n \to \infty$, we observe that

$$\begin{aligned}\mathbb{E}\left[e^{itZ_n}\right] &= \prod_{j=1}^{n}\left(1 - \frac{1}{j+1} + \frac{\cos t}{j+1}\right)\\ &= \frac{1}{n+1}\Pi_n(\cos t)\end{aligned}$$

since $\mathbb{E}\left[e^{itX_1}\right] = \cos t$. Therefore,

$$\psi_n(w\sqrt{\log n}) = \mathbb{E}\left[e^{iwZ_n}\right] = \frac{1}{n+1}\Pi_n(\cos w).$$

We note that $\cos w$ is decreasing in $\left[\frac{\pi}{2}, \pi\right]$ and for all $w \in \left[\frac{\pi}{2}, \pi\right]$, $-1 \leq \cos w \leq 0$. Therefore, there exists $\eta > \frac{\pi}{2}$ such that for all $w \in [\pi - \eta, \pi)$ we have $-1 < \cos(\pi - \eta) < 0$ and

$$\left|\psi_n(w\sqrt{\log n})\right| \leq \frac{1}{n+1}\Pi_n(\cos(\pi - \eta)).$$

Since $-1 < \cos(\pi - \eta) < 0$, so for $j \geq 1$, $\left(1 + \frac{\cos(\pi-\eta)}{j}\right) < 1$. Therefore,

$$\Pi_n(\cos(\pi - \eta)) \leq 1. \tag{4.31}$$

Let us write

$$I_2(n) = J_1(n) + J_2(n)$$

where

$$J_1(n) = \sqrt{\log n}\int_{\delta}^{\pi-\eta}\left|\psi_n\left(w\sqrt{\log n}\right)\right|\,\mathrm{d}w \tag{4.32}$$

and

$$J_2(n) = \sqrt{\log n} \int_{\pi-\eta}^{\pi} \left| \psi_n \left(w\sqrt{\log n} \right) \right| \, \mathrm{d}w.$$

It is easy to see from (4.31) that

$$J_2(n) \le \frac{\eta}{n+1} \sqrt{\log n} \longrightarrow 0 \text{ as } n \to \infty.$$

For all $t \in [\delta, \pi - \eta]$, $0 \le |\cos t| < 1$, so there exists $0 < \alpha < 1$ such that $0 \le |\cos t| \le \alpha$ for all $t \in [\delta, \pi - \eta]$. Recall that

$$\psi_n(w\sqrt{\log n}) = \prod_{j=1}^{n} \left(1 - \frac{1}{j+1} + \frac{\cos w}{j+1} \right).$$

Using the inequality $1 - x \le e^{-x}$, it follows that for all $t \in [\delta, \pi - \eta]$

$$1 - \frac{1}{j+1} + \frac{|\cos t|}{j+1} \le 1 - \frac{1}{j+1} + \frac{\alpha}{j+1} \le e^{-\frac{1-\alpha}{j+1}}$$

and hence

$$\frac{1}{n+1} |\Pi_n(\cos t)| \le e^{-\sum\limits_{j=1}^{n} \frac{1-\alpha}{j+1}} \le Ce^{-(1-\eta)\log n}$$

where C is some positive constant. Therefore from (4.32) we obtain as $n \to \infty$

$$J_1(n) \le Ce^{-(1-\alpha)\log n} (\pi - \eta - \delta) \sqrt{\log n} \longrightarrow 0.$$

$\square$

4.3.2 Proofs for the Local Limit Type Results for $d \ge 2$

Proof of Theorem 2.8 . Without loss of generality we may assume that $\mu = 0$ and $\Sigma = \mathbb{I}_d$ and the process begins with a single ball of color indexed by $\mathbf{0}$. From Theorem 3.1, we obtain $Z_n \stackrel{d}{=} \sum_{k=1}^{n} I_j X_j$. X_j being a lattice random variable, $I_j X_j$ is also so. Furthermore, $\mathbf{0} \in B$, therefore X_j and $I_j X_j$ are supported on the same lattice. Therefore, Z_n has lattice $\mathcal{L}_n^{(d)}$. For $A \subset \mathbb{R}^d$ and $x \in \mathbb{R}$, we define $xA = \{xy \colon y \in A\}$. By Fourrier inversion formula (21.28 on page 230 of [3]), we get for $x \in \mathcal{L}_n^{(d)}$

$$\mathbb{P}\left(\frac{Z_n}{\sqrt{\log n}} = x \right) = \frac{l}{(2\pi\sqrt{\log n})^d} \int\limits_{(\sqrt{\log n}\mathcal{F}^*)} \psi_n(t) e^{-i\langle t, x \rangle} \, \mathrm{d}t$$

where $\psi_n(t) = \mathbb{E}\left[e^{i\langle t, \frac{Z_n}{\sqrt{\log n}} \rangle} \right]$, $l = |\det(\mathcal{L})|$ and $\mathcal{F}^*$ is the fundamental domain for X_1 as defined in equation(21.22) on page 229 of [3]. Also by Fourrier inversion formula

$$\phi_d(x) = \frac{1}{(2\pi)^d} \int\limits_{\mathbb{R}^d} e^{-i\langle t, x \rangle} e^{-\frac{\|t\|^2}{2}} \, \mathrm{d}t.$$

Given $\epsilon > 0$, there exists $N > 0$ such that $n \geq N$ large enough,

$$\begin{aligned}\left|\frac{(\sqrt{\log n})^d}{l}\mathbb{P}\left(\frac{Z_n}{\sqrt{\log n}} = x\right) - \phi_d(x)\right| &\leq \frac{1}{(2\pi)^d}\int\limits_{(\sqrt{\log n}\mathcal{F}^*)}\left|\psi_n(t) - e^{-\frac{\|t\|^2}{2}}\right|\mathrm{d}t + \frac{1}{(2\pi)^d}\int\limits_{\mathbb{R}^d\setminus\sqrt{\log n}\mathcal{F}^*} e^{-\frac{\|t\|^2}{2}}\,\mathrm{d}t \\ &\leq \frac{1}{(2\pi)^d}\int\limits_{(\sqrt{\log n}\mathcal{F}^*)}\left|\psi_n(t) - e^{-\frac{\|t\|^2}{2}}\right|\mathrm{d}t + \epsilon.\end{aligned}$$

Given any compact set $A \subset \mathbb{R}^d$ for all n large enough

$$\int\limits_{(\sqrt{\log n}\mathcal{F}^*)}\left|\psi_n(t) - e^{-\frac{\|t\|^2}{2}}\right|\mathrm{d}t \leq \int\limits_A\left|\psi_n(t) - e^{-\frac{\|t\|^2}{2}}\right|\mathrm{d}t + \int_{(\sqrt{\log n}\mathcal{F}^*)\setminus A}\left|\psi_n(t)\right|\mathrm{d}t + \int_{\mathbb{R}^d\setminus A} e^{-\frac{\|t\|^2}{2}}\,\mathrm{d}t.$$

By Theorem 2.1, we know that $\frac{Z_n}{\sqrt{\log n}} \xrightarrow{d} N_d(\mathbf{0}, \mathbb{I}_d)$ as $n \to \infty$. Therefore, for any compact set $A \subset \mathbb{R}^d$ by bounded convergence theorem,

$$\int\limits_A\left|\psi_n(t) - e^{-\frac{\|t\|^2}{2}}\right|\mathrm{d}t \longrightarrow 0 \text{ as } n \to \infty.$$

Choose A such that

$$\int\limits_{A^c} e^{-\frac{\|t\|^2}{2}}\,\mathrm{d}t < \epsilon.$$

Let us write

$$I(n) = \int\limits_{(\sqrt{\log n}\mathcal{F}^*)\setminus A}\left|\psi_n(t)\right|\mathrm{d}t. \tag{4.33}$$

For the above choice of A, we will show that

$$I(n) \longrightarrow 0 \text{ as } n \to \infty.$$

Since $Z_n \stackrel{d}{=} \sum_{j=1}^n I_j X_j$, we have

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

where $e(it) = \mathbb{E}\left[e^{i\langle t, X_1\rangle}\right]$. So,

$$\psi_n(t) = \mathbb{E}\left[e^{i\langle t, \frac{Z_n}{\sqrt{\log n}}\rangle}\right] = \frac{1}{n+1}\Pi_n\left(e\left(\frac{1}{\sqrt{\log n}}it\right)\right).$$

Applying a change of variables $t = \frac{1}{\sqrt{\log n}} w$ to (4.33), we obtain

$$I(n) = (\sqrt{\log n})^d \int\limits_{\mathcal{F}^* \setminus \frac{1}{\sqrt{\log n}} A} \left| \psi_n \left(\sqrt{\log n} w \right) \right| \mathrm{d}w. \tag{4.34}$$

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

$$|e(iw)| \le 1 - \frac{b\|w\|^2}{2}. \tag{4.35}$$

Therefore, using the inequality $1 - x \le e^{-x}$

$$\begin{aligned} |\psi_n(\sqrt{\log n} w)| &= \frac{1}{n+1} |\Pi_n \left(e(iw) \right)| \\ &\le \prod_{j=1}^{n+1} \left(1 - \frac{1}{j+1} + \frac{|e(iw)|}{j+1} \right) \\ &\le e^{-\sum\limits_{j=1}^{n} \frac{b}{j+1} \frac{\|w\|^2}{2}} \le C_1 e^{-b \frac{\|w\|^2}{2} \log n} \end{aligned} \tag{4.36}$$

for some positive constant C_1. From (4.34) we can write

$$I(n) = I_1(n) + I_2(n)$$

where

$$I_1(n) = (\sqrt{\log n})^d \int\limits_{\left(B(0,\delta) \setminus \frac{1}{\sqrt{\log n}} A \right) \cap \mathcal{F}^*} |\psi_n \left(\sqrt{\log n} w \right)| \,\mathrm{d}w$$

and

$$I_2(n) = (\sqrt{\log n})^d \int_{\mathcal{F}^* \setminus B(0,\delta)} |\psi_n \left(\sqrt{\log n} w \right)| dw.$$

Since (4.36) holds, given $\epsilon > 0$, we have for all n large enough

$$I_1(n) \le (\sqrt{\log n})^d \int\limits_{B(0,\delta) \setminus \frac{A}{\sqrt{\log n}}} C_1 e^{-b \frac{\|w\|^2}{2} \log n} \,\mathrm{d}w \le \epsilon. \tag{4.37}$$

Since $\mathbf{0} \in B$, the lattices for X_j and $I_j X_j$ are same. Therefore, for all $w \in \mathcal{F}^* \setminus B(0, \delta)$ we get $|e(iw)| < 1$, so there exists an $0 < \eta < 1$, such that $|e(iw)| \le \eta$. Therefore, using the inequality $1 - x \le e^{-x}$, we obtain

$$|\psi_n(\sqrt{\log n} w)| \le e^{-\sum_{j=i}^{n} \frac{1}{j+1} (1-\eta)} \le C_2 e^{-(1-\eta) \log n} \tag{4.38}$$

for some positive constant C_2. Therefore, using equation (21.25) on page 230 of [3] we obtain

$$I_2(n) \le C_2' (\sqrt{\log n})^d e^{-(1-\eta) \log n} \longrightarrow 0 \text{ as } n \to \infty$$

where C_2' is an appropriate positive constant. $\square$

Proof of the Theorem 2.9. It is easy to note that S_n is the simple symmetric random walk on $\mathbb{Z}^d$, $\mu = 0$ and $\Sigma = d^{-1}\mathbb{I}_d$. We further assume, without loss of generality that the process starts with a single ball of color indexed by $\mathbf{0}$.

For notional simplicity we write a proof for $d = 2$, the general case can be written similarly.

Now when $d = 2$, 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 (2.14), it is equivalent to proving

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

where $\phi_{2,\frac{1}{2}\mathbb{I}_2}(x) = \frac{1}{\pi} e^{-xx^t}$ is the bivariate normal density with mean vector $\mathbf{0}$ and variance-covariance matrix $\frac{1}{2}\mathbb{I}_2$ and $\frac{1}{\sqrt{2}}\mathcal{L}_n^{(2)} = \{x \colon x = \frac{1}{\sqrt{\log n}}\mathbb{Z}^2\}$. By Fourrier inversion formula (21.28 on page 230 of [3]), we get for $x \in \frac{1}{\sqrt{2}}\mathcal{L}_n^{(2)}$

$$\mathbb{P}\left(\frac{Z_n}{\sqrt{\log n}} = x \right) = \frac{1}{(2\pi)^2 \log n} \int\limits_{\left(-\sqrt{\log n}\pi, \sqrt{\log n}\pi\right)^2} \psi_n(t) e^{-i\langle t, x\rangle}\, \mathrm{d}t$$

Also by Fourrier inversion formula

$$\phi_{2,\frac{1}{2}\mathbb{I}_2}(x) = \frac{1}{(2\pi)^2} \int\limits_{\mathbb{R}^2} e^{-i\langle t,x\rangle} e^{-\frac{\|t\|^2}{4}}\, \mathrm{d}t.$$

Let us write $H_n = \left(-\sqrt{\log n}\pi, \sqrt{\log n}\pi\right)^2$. Given $\epsilon > 0$, there exists $N > 0$ such that $n \geq N$ large enough,

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

Given any compact set $A \subset \mathbb{R}^2$ for all n large enough we have

$$\int\limits_{H_n} \left| \psi_n(t) - e^{-\frac{\|t\|^2}{4}} \right| \mathrm{d}t \leq \int\limits_{A} \left| \psi_n(t) - e^{-\frac{\|t\|^2}{4}} \right| \mathrm{d}t + \int_{H_n \setminus A} \left| \psi_n(t) \right| \mathrm{d}t + \int_{\mathbb{R}^2 \setminus A} e^{-\frac{\|t\|^2}{4}}\, \mathrm{d}t.$$

By Theorem 2.1, we know that $\frac{Z_n}{\sqrt{\log n}} \xrightarrow{d} N_2(\mathbf{0}, 2^{-1}\mathbb{I}_2)$ as $n \to \infty$. Therefore, for any compact set $A \subset \mathbb{R}^2$ by bounded convergence theorem,

$$\int\limits_{A} \left| \psi_n(t) - e^{-\frac{\|t\|^2}{4}} \right| \mathrm{d}t \longrightarrow 0 \text{ as } n \to \infty.$$

Choose A such that

$$\int\limits_{A^c} e^{-\frac{\|t\|^2}{4}} \, \mathrm{d}t < \epsilon.$$

Let us write

$$I(n) = \int\limits_{H_n \setminus A} \Big| \psi_n(t) \Big| \, \mathrm{d}t.$$

For the above choice of A, we will show that

$$I(n) \longrightarrow 0 \text{ as } n \to \infty.$$

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

$$I(n) = \log n \int\limits_{(-\pi,\pi)^2 \setminus \frac{1}{\sqrt{\log n}} A} \Big| \psi_n \left(\sqrt{\log n} w \right) \Big| \, \mathrm{d}w.$$

We can write

$$I(n) = I_1(n) + I_2(n)$$

where

$$I_1(n) = \log n \int\limits_{\left(B(0,\delta) \setminus \frac{1}{\sqrt{\log n}} A \right) \cap (-\pi,\pi)^2} | \psi_n \left(\sqrt{\log n} w \right) | \, \mathrm{d}w$$

and

$$I_2(n) = \log n \int\limits_{(-\pi,\pi)^2 \setminus B(0,\delta)} | \psi_n \left(\sqrt{\log n} w \right) | dw.$$

where δ is as in (4.35). Using arguments similar to (4.37), we can show that $I_1(n) \longrightarrow 0$ as $n \to \infty$. Therefore it is enough to show that $I_2(n) \longrightarrow 0$ as $n \to \infty$. 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 $e\,(it) = \frac{1}{2}(\cos t_1 + \cos t_2)$. So if $t \in [-\pi, \pi]^2$ be such that $|e\,(it)| = 1$, then $t = (\pi, \pi), (-\pi, \pi), (\pi, -\pi), (-\pi, -\pi)$. The function $\cos t$ is continuous and decreasing as a function of t for $t \in \left[\frac{\pi}{2}, \pi\right]$. Choose $\eta > \frac{\pi}{2}$ 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

$$I_2(n) = J_1(n) + J_2(n)$$

where

$$J_1(n) = \log n \int_{A_1} | \psi_n \left(\sqrt{\log n} w \right) | dw$$

and

$$J_2(n) = \log n \int_D |\psi_n\left(\sqrt{\log n} w\right)| dw.$$

It is easy to note that

$$J_1(n) \leq \log n \int\limits_{\overline{A}_1} |\psi_n\left(\sqrt{\log n} w\right)| dw$$

where $\overline{A}_1$ denotes the closure of A_1. For $w \in \overline{A}_1$ there exists some $0 < \alpha < 1$ such that $|e(it)| \leq \alpha$. Therefore using bounds similar to that in (4.38) we can show that

$$J_1(n) \longrightarrow 0 \text{ as } n \to \infty.$$

We observe that

$$J_2(n) \leq 4 \log n \int_{[\pi-\eta,\pi]^2} |\psi_n\left(\sqrt{\log n} w\right)| dw.$$

Hence, it is enough to show that $\log n \int_{[\pi-\eta,\pi]^2} |\psi_n\left(\sqrt{\log n} w\right)| dw \longrightarrow 0$ as $n \to \infty$. 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

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

So,

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

□

5 Urn on General Graphs on $\mathbb{R}^d$

We can further generalize the urn models to a large variety of graphs with vertex set a countable subset of $\mathbb{R}^d$ and an appropriate edge set. Let $\{X_i\}_{i\geq 1}$ be a sequence of random d-dimensional i.i.d. vectors with non empty support set $B \subset \mathbb{R}^d$. We assume the cardinality of B to be finite. Consider $V := \left\{\sum_{i=1}^{k} n_i b_i \colon n_1, n_2, \ldots, n_k \in \mathbb{N}, b_1, b_2, \ldots, b_k \in B\right\} \subseteq \mathbb{R}^d$ and $E := \{\{u,v\} \colon v, u \in B \text{ and } v - u \in B\}$. Let $G = (V, E)$ be the graph with vertex set $V \subseteq \mathbb{R}^d$ and edge set E. Let the law of X_1 be given by

$$\mathbb{P}(X_1 = w) = p(w) \text{ for all } w \in B.$$

where $0 < p(w) \leq 1$ for all $w \in B$ and $\sum_{w \in B} p(w) = 1$. Define the random walk $S_n = \sum_{k=1}^{n} X_k$. In this section, we consider urn model with replacement matrix given by the stochastic matrix associated with the bounded increment random walk S_n on G. Let the urn evolve according to the random walk S_n with the replacement matrix R given by

$$R(u,v) = \begin{cases} p(v-u), & \text{if } v - u \in B \\ 0 & \text{otherwise} \end{cases} \tag{5.39}$$

for all $u, v \in V$.

The urn models as defined in the Subsection 1.1 with vertex set $\mathbb{Z}^d$ and edge set being the nearest neighbor links are special cases of urns on general graphs.

5.1 Main Results for General Graphs

Throughout this and the remaining subsections V, $(X_i)_{i\geq 1}$ and B will remain as defined in Section 5.

For $X_1 = \left(X_1^{(1)}, X_1^{(2)}, \ldots, X_1^{(d)}\right)$, let $\mu = \left(\mathbb{E}[X_1^{(1)}], \mathbb{E}[X_1^{(2)}], \ldots, \mathbb{E}[X_1^{(d)}]\right)$ and $\Sigma = [\sigma_{ij}]_{d\times d}$ where $\sigma_{ij} = \mathbb{E}[X_1^{(i)} X_1^{(j)}]$. We assume Σ to be positive definite and as earlier let $\Sigma = \Sigma^{1/2}\Sigma^{1/2}$. Let $S_n = \sum_{j=1}^n X_j$. Let U_n denote the configuration of the urn at time n and the process begin with $U_0 = (U_{0,v})_{v\in V}$ such that $U_{0,v} = 0$ except for finitely many v.

Theorem 5.1. *Let the urn model be associated with the random walk $\{S_n\}_{n\geq 0}$. Let $\overline{\Lambda}_n$ be the probability measure corresponding to the probability vector $\frac{1}{n+1}\left(\mathbb{E}[U_{n,v}]\right)_{v\in V}$. For all $A \in \mathcal{B}\left(\mathbb{R}^d\right)$, define $\overline{\Lambda}_n^{cs} = \overline{\Lambda}_n\left(\sqrt{\log n} A \Sigma^{-1/2} + \mu \log n\right)$. Then, as $n \to \infty$*

$$\overline{\Lambda}_n^{cs} \xrightarrow{w} \Phi_d.$$

where for $\Phi_d\left(A\right) = \frac{1}{(\sqrt{2\pi})^d}\int\limits_A e^{-\frac{xx^t}{2}}\,\mathrm{d}t$ for $A \in \mathcal{B}\left(\mathbb{R}^d\right)$.

Remark 8. Theorem 5.1 states that if Z_n be the color of the randomly selected ball in the $(n+1)^{\text{th}}$ draw, that is Z_n is a random d-dimensional vector corresponding to the probability vector $\frac{1}{n+1}\left(\mathbb{E}[U_{n,v}]\right)_{v\in V}$, then

$$\frac{Z_n - \mu \log n}{\sqrt{\log n}} \xrightarrow{d} N_d(\mathbf{0}, \Sigma) \text{ as } n \to \infty. \tag{5.40}$$

For $\omega \in \Omega$, let $\Lambda_n(\omega) \in \mathcal{M}_1$ be the random probability measure corresponding to the random probability vector $\frac{U_n(\omega)}{n+1}$ where $\mathcal{M}_1$ is the space of probability measures on $\mathbb{R}^d$, $d \geq 1$ endowed with the topology of weak convergence.

Theorem 5.2. *For $\omega \in \Omega$, let $\Lambda_n^{cs}(\omega)\left(A\right) = \Lambda_n(\omega)\left(\sqrt{\log n} A \Sigma^{-1/2} + \mu \log n\right)$. Then, as $n \to \infty$,*

$$\Lambda_n^{cs} \xrightarrow{p} \Phi_d \textit{ on } \mathcal{M}_1, \tag{5.41}$$

where $\Phi_d\left(A\right) = \frac{1}{(\sqrt{2\pi})^d}\int\limits_A e^{-\frac{xx^t}{2}}\,\mathrm{d}t$ for $A \in \mathbb{R}^d$.

Remark 9. From Theorem 5.2 we can conclude that given any subsequence $\{n_k\}_{k=1}^\infty$ there exists a further subsequence $\{n_{k_j}\}_{j=1}^\infty$ such that as $j \to \infty$

$$\Lambda_{n_{k_j}}^{cs} \xrightarrow{w} \Phi_d \text{ almost surely for all } \omega.$$

Let the support set for the i.i.d. sequence of increment vectors $\{X_i\}_{i\geq 1}$ be

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

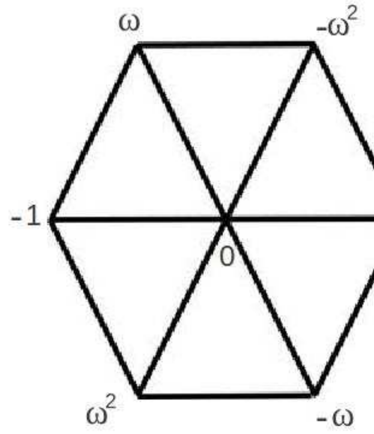

Figure 1: Triangular Lattice

where ω, ω^2 are the complex cube roots of unity and the law be given by $\mathbb{P}(X_1 = v) = \frac{1}{6}$ for every $v \in B$. Let $G = (V, E)$ be the corresponding graph. It is called the triangular lattice in $\mathbb{R}^2$. Let $S_n = \sum_{j=1}^{n} I_j X_j$ be the random walk on the triangular lattice G. The urn with colors indexed by the vertices of the triangular lattice and replacement matrix given by the stochastic matrix of the increment vectors X_i is an example of urn on general graphs.

Corollary 5.3. *Let the colors of the balls in the urn be indexed by the vertex set of the triangular lattice on $\mathbb{R}^2$ and the urn model be associated with the random walk $\{S_n\}_{n\geq 0}$ on triangular lattice. Let the process begin with a single ball of color indexed by $\mathbf{0}$, then as $n \to \infty$*

$$\frac{Z_n}{\sqrt{\log n}} \xrightarrow{d} N_2\left(\mathbf{0}, \frac{1}{2}\mathbb{I}_2\right). \tag{5.42}$$

5.2 Proofs

The proofs of Theorems 5.1 and 5.2 are exactly similar to that of Theorems 2.1 and 2.4 and hence omitted. However, we present the proof of Corollary 5.3.

Proof of Corollary 5.3. In this case $B = \{(1,0), (-1,0), \omega, -\omega, \omega^2, -\omega^2\}$ where ω and ω^2 are the complex cube roots of unity. For a complex number z, $Re(z)$ and $Im(z)$ will denote respectively the real and imaginary parts of z.

Since $1 + \omega + \omega^2 = 0$, therefore it is immediate that $\mu = 0$. Also we know that $\omega = \frac{1}{2} + i\frac{\sqrt{3}}{2}$. Observe that $\mathbb{E}\left[\left(X_1^{(1)}\right)^2\right] = \frac{2}{6}\left(1 + (Re\ \omega)^2 + \left(Re\ \omega^2\right)^2\right)$ and $Re\ \omega = Re\ \omega^2$, 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, since $Im(\omega) = -Im(\omega^2)$, $\mathbb{E}\left[\left(X_1^{2}\right)^2\right] = \frac{2}{6}\left((Im(\omega))^2 + \left(Im(\omega^2)\right)^2\right) = \frac{1}{2}$. Furthermore $\mathbb{E}\left[X_1^{(1)}X_1^{(2)}\right] = -\frac{2}{6}Im\left(1 + \omega + \omega^2\right) = 0$. $\square$

6 Technical Results

In this section we present the following technical result which we have used in the proof of Theorem 2.4.

Theorem 6.1. *Let $\hat{\mu}_n$ be a sequence of probability measures on $\left(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\right)$, $d \geq 1$ and let $m_n(\cdot)$ be the corresponding moment generating functions. For some $\delta > 0$ if $m_n(\lambda) \longrightarrow e^{\frac{\|\lambda\|^2}{2}}$ as $n \to \infty$ for every $\lambda \in [-\delta, \delta]^d \cap \mathbb{Q}^d$ then as $n \to \infty$*

$$\hat{\mu}_n \xrightarrow{w} \Phi_d,$$

where $\Phi_d(A) = \frac{1}{(\sqrt{2\pi})^d} \int\limits_A e^{-\frac{xx^t}{2}} \, \mathrm{d}t$ *for* $A \in \mathcal{B}\left(\mathbb{R}^d\right)$.

Proof. For notational simplicity we present the proof when $d = 1$. The proof when $d \geq 2$ is similar with appropriate modifications. From Theorem 22.2 of [4], we know that a distribution is uniquely determined by its moment generating function in a neighborhood of 0. Therefore it is enough to prove that sequence of probability measures $\hat{\mu}_n$ is tight. Choose a $\delta' \in \mathbb{Q}$ such that $0 < \delta' < \delta$. To prove tightness, we observe that for $a > 0$

$$\hat{\mu}_n\left[(-a, a)^c\right] \leq e^{-\delta' a}\left(m_n(-\delta') + m_n(\delta')\right).$$

Since $m_n(\delta') \to m(\delta')$ and $m_n(-\delta') \to m(-\delta')$ as $n \to \infty$, we can say that

$$\sup_{n \geq 1} \hat{\mu}_n\left[(-a, a)^c\right] \longrightarrow 0 \text{ as } a \to \infty.$$

Therefore, for every subsequence $\{n_k\}_{k \geq 1}$ there exists a further subsequence $\{n_{k_j}\}_{j \geq 1}$ and a probability measure $\hat{\mu}$ such that $\hat{\mu}_{n_{k_j}} \xrightarrow{d} \mu$ as $n \to \infty$. We already know that $m_{n_{k_j}}(\lambda) \longrightarrow e^{\frac{\lambda^2}{2}}$ as $j \to \infty$ for every $\lambda \in [-\delta, \delta] \cap \mathbb{Q}$. Since the moment generating function is continuous as a function of its argument, therefore the moment generating function for $\hat{\mu}$ is $e^{\frac{\lambda^2}{2}}$. Therefore this completes the proof. □

Acknowledgement

The authors are grateful to Krishanu Maulik for various discussions they had with him regarding general urn models and also for looking through the first draft of this work.

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] 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.

[4] 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.

[5] Arup Bose, Amites Dasgupta, and Krishanu Maulik. Multicolor urn models with with reducible replacement matrices. *Bernoulli*, 15(1):279–295, 2009.

[6] Arup Bose, Amites Dasgupta, and Krishanu Maulik. Strong laws for balanced triangular urns. *J. Appl. Probab.*, 46(2):571–584, 2009.

[7] John B. Conway. *Functions of one complex variable*, volume 11 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1978.

[8] Amites Dasgupta and Krishanu Maulik. Strong laws for urn models with balanced replacement matrices. *Electron. J. Probab.*, 16:no. 63, 1723–1749, 2011.

[9] 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.

[10] David A. Freedman. Bernard Friedman's urn. *Ann. Math. Statist*, 36:956–970, 1965.

[11] Bernard Friedman. A simple urn model. *Comm. Pure Appl. Math.*, 2:59–70, 1949.

[12] Raúl Gouet. Strong convergence of proportions in a multicolor Pólya urn. *J. Appl. Probab.*, 34(2):426–435, 1997.

[13] Svante Janson. Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Process. Appl.*, 110(2):177–245, 2004.

[14] Robin Pemantle. A time-dependent version of Pólya's urn. *J. Theoret. Probab.*, 3(4):627–637, 1990.

[15] Robin Pemantle. A survey of random processes with reinforcement. *Probab. Surv.*, 4:1–79, 2007.

[16] G. Pólya. Sur quelques points de la théorie des probabilités. *Ann. Inst. H. Poincaré*, 1(2):117–161, 1930.