RANDOM ADDITIONS IN URNS OF INTEGERS

MACKENZIE SIMPER [ID],* *Stanford University*

Abstract

Consider an urn containing balls labeled with integer values. Define a discrete-time random process by drawing two balls, one at a time and with replacement, and noting the labels. Add a new ball labeled with the sum of the two drawn labels. This model was introduced by Siegmund and Yakir (2005) *Ann. Prob.* **33**, 2036 for labels taking values in a finite group, in which case the distribution defined by the urn converges to the uniform distribution on the group. For the urn of integers, the main result of this paper is an exponential limit law. The mean of the exponential is a random variable with distribution depending on the starting configuration. This is a novel urn model which combines multi-drawing and an infinite type of balls. The proof of convergence uses the contraction method for recursive distributional equations.

Keywords: urn model; exponential limit law; recursive distributional equation; contraction method

2010 Mathematics Subject Classification: Primary 60F05
Secondary 60K35

1. Introduction

In [19], the following urn model was introduced. Balls in an urn are labeled with elements of a finite group G. To update the urn, two balls are drawn (independently and with replacement) and their labels recorded. A new ball with the product of the two labels is then added to the urn. When the initial configuration of the urn contains generators of the group, it is proved that the composition of the urn converges to a uniform distribution on the group (an alternate proof is given in [1]). A natural extension of the model is to allow labels with values from an infinite group. Specifically, this paper considers the same dynamics for an urn containing integers with the addition operation, and calls such a process a $\mathbb{Z}$-urn.

The behavior of the urn should depend on the initial configuration of balls. Perhaps a natural first question to explore is the behavior of the urn initially started with an additive basis for $\mathbb{Z}$, e.g. $\{-1, 1\}$. Figure 1 shows the results from two different simulations of this model. An interesting phenomenon is observed with the starting configuration of $\{-1, 1\}$: the values in the urn either become almost all positive or almost all negative. Furthermore, the curve of the histogram appears exponential, reflected on either side of the vertical axis. That is, if μ_n is the empirical measure defined by the labels of the n balls in the urn, then μ_n converges to some
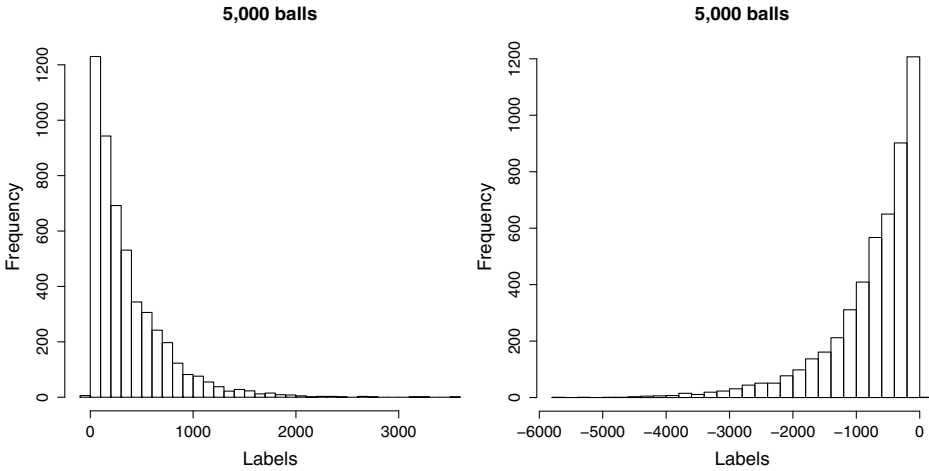
FIGURE 1: Histograms of two different trials of the urn model with initial configuration $\{-1, 1\}$, after 5000 balls have been added. Each histogram has 30 bins. The bin width of the left histogram is 100, with mean 356, and the width of the right histogram is 200, with mean -737.

scaled exponential. Another observation from simulations is that the mean of the distribution μ_n varies with different trials, i.e. there is not a deterministic limit.

The appropriately rescaled mean of the empirical measure defined by the urn process converges almost surely to some random variable. The distribution of the limiting random variable is dependent on the initial configuration of the urn. The distribution of a random draw from the urn then converges to an exponential. This is our main result, Theorem 1, described in the following section.

1.1. Model and results

Let τ_0 be the number of balls initially started in the urn. For convenience, we begin the time index at $\tau_0 + 1$, so that at time $n > \tau_0$ there are exactly n balls in the urn (and the number of additions is $n - \tau_0$). Let $X_n \in \mathbb{Z}^d$ denote the label of the nth ball in the urn. When $d > 1$, write $X_n(i)$ for the ith coordinate. Thus, the urn at time n is given by the set $\mathbf{U}_n = \mathbf{U}_0 \cup \{X_{\tau_0+1}, \ldots, X_n\}$.

The letter Z_n is used to denote a random draw from the urn, that is $Z_n \sim \mu_n$, where μ_n is the empirical distribution

$$\mu_n = \frac{1}{n}\left(\sum_{x \in \mathbf{U}_0} \delta_x + \sum_{i=\tau_0+1}^{n} \delta_{X_i} \right), \qquad n \geq \tau_0.$$

Then X_{n+1} is generated by $X_{n+1} = Z_n^1 + Z_n^2$, where $Z_n^1 \stackrel{\text{d}}{=} Z_n^2 \stackrel{\text{d}}{=} Z_n$ are two independent draws from the urn at step n. For convenience, assign an arbitrary indexing to the initial balls $x \in \mathbf{U}_0$, writing them $X_1, X_2, \ldots, X_{\tau_0}$. Let $S_n = \sum_{i=1}^{n} X_i$ be the sum of the labels of balls in the urn at time n. When studying Z_n and X_n we focus on the quenched values. That is, there is an implicit underlying urn process, and expectations are calculated conditional on the appropriate step of the urn. For instance, the mean of Z_n is $\mathbb{E}[Z_n \mid \mu_n] = \frac{S_n}{n}$.

Theorem 1. *Let* $\{\mu_n\}_{n\geq 0}$ *be a sequence of empirical measures defined by the* $\mathbb{Z}$-*urn with any initial configuration. Suppose* $Z_n \sim \mu_n$ *is a random draw from the urn. Let* $A = \lim_{n\to\infty} \frac{\mathbb{E}[Z_n|\mu_n]}{n+1}$. *Then* A *exists almost surely,* $\mathbb{P}(A \neq 0) > 0$, *and*

$$\lim_{n\to\infty} \mathcal{L}\left(\frac{Z_n}{n} \,\Big|\, \mu_n\right) = \mathcal{L}(Z),$$

where $Z \sim A \cdot \mathrm{Exp}(1)$.

In particular, this result implies that if the urn is started with $\{-1, 1\}$, the limiting distribution defined by the urn will be supported either on $(0, \infty)$ or $(-\infty, 0)$, depending on the result for A. As remarked before, Theorem 1 is a statement about the *quenched* version of the problem. This means a realization of the urn process $\{\mu_n\}$ is fixed and the random variable Z_n is sampled from the urn process at stage n. In contrast, let Z_n^* denote the *annealed* model. That is, Z_n^* is the random variable generated by first generating a new urn μ_n, and then sampling from μ_n. In other words, Z_n^* is the mixture of Z_n over all realizations of μ_n. The limiting distribution of $\frac{Z_n^*}{n}$ is the mixture distribution of exponential multiplied by the random variable A. At the moment, there is not an explicit form for the distribution of A, though we can observe how the expected value of A depends on the initial configuration of the urn.

All results are the same for urns with labels taking values in the real numbers. The motivation for introducing the process for integer-value labels comes from (a) the original desire to extend the model from [19] to a finitely generated group, and (b) to somewhat simplify the possible starting configurations that can be investigated. In addition, the methods can easily be extended to urns with vector labels. The exponential limit law for Z_n, a *draw* from the urn, follows from a limit law for X_n, the nth ball added to the urn. The key is that the limit of an appropriately scaled version of X_n satisfies the distributional identity

$$X \stackrel{\mathrm{d}}{=} U_1 \cdot X^1 + U_2 \cdot X^2, \tag{1}$$

where U_1, U_2 are independent and identically distributed (i.i.d.) Uniform([0, 1]) random variables and X^1, X^2 are i.i.d. copies of X. The unique random variable that satisfies this distributional identity is a Gamma with shape 2 and mean $\mathbb{E}[X]$. If $X \in \mathbb{R}^d$ is a vector, then it is interesting to observe that solutions to (1) are of the form $G \cdot \mathbb{E}[X]$, where G is a one-dimensional Gamma random variable. That is, each scaled coordinate of X_n marginally converges to a Gamma distribution, but as a joint process every coordinate converges to a multiple of the *same* Gamma random variable. The precise statement for urns with vector labels is given in the following theorem.

Theorem 2. *Let* $\{\mu_n^d\}_{n\geq 0}$ *be a sequence of empirical measures defined by the addition urn model on* $\mathbb{Z}^d$. *Suppose* $\mathbf{X}_n \in \mathbb{Z}^d$ *is the nth ball added to the urn. Let* $\mathbf{A} = \lim_{n\to\infty} \frac{\mathbb{E}[\mathbf{X}_{n+1}|\mu_n]}{2(n+1)}$. *Then* $\mathbf{A} = (A(1), \ldots, A(d)) \in \mathbb{R}^d$ *exists almost surely,* $\mathbb{P}(A(i) \neq 0) > 0$ *for all* $i = 1, \ldots, d$, *and*

$$\lim_{n\to\infty} \mathcal{L}\left(\frac{\mathbf{X}_n}{n} \,\Big|\, \mu_{n-1}^d\right) = \mathcal{L}(\mathbf{X}),$$

where $\mathbf{X} \sim G \cdot \mathbf{A}$, *and* $G \sim Gamma(2, 1)$.

Remark 1. A simple generalization of the model would be to draw k balls (with replacement) and add a new ball with label the sum of the k drawn balls. For any integer $k \geq 2$, the same

techniques used in this paper prove that the nth labeled added X_n has limiting distribution X satisfying the identity $X \stackrel{\mathrm{d}}{=} \sum_{i=1}^{k} U_i \cdot X^i$, where U_i are i.i.d. Uniform([0, 1]) and X^i are i.i.d. copies of X, for $1 \leq i \leq k$. Note that when $k > 2$, a Gamma distribution does *not* satisfy this distributional identity.

1.2. Related work

Random processes defined using urns can arise is many different settings. For a plethora of example models and applications, see [6] and [16]. 'Pólya-type' urn models refer to an urn containing different types of balls with some sort of replacement scheme. That is, balls are drawn randomly from the urn and replaced with some number and type of balls that depends on the type of the drawn balls [13].

Urn models are well studied for a finite number of colors (i.e. types of balls); an extended scheme allowing a continuum of colors was proposed in the classic paper by Blackwell and MacQueen [4]. More work in this setting has been done in the past few years, and a general infinite-color model was introduced in [21]. In [2], the model on urns containing countably infinite number of colors is proposed, and when the replacement schemes are balanced and associated with bounded increment random walks on the color space (e.g. $\mathbb{Z}^d$), central and local limit theorems for the random color of the nth selected ball can be proven. This has been extended to almost sure convergence, with assumptions on the replacement scheme [5]. In all of these works, the urns are updated after drawing a *single* ball. The replacement rule is a function of a single variable of the color space.

Much less work has been done for models in which *multiple* balls are drawn at each stage. Several specific cases have been analyzed, for example in Chapter 10 of [13]. The first general results for two-color urns were contained in [9,10]. The results were extended to arbitrary (but finite) d-color urns, with an additional assumption removed, in [11]. The method used in [11] is *stochastic approximation*, which is useful for a broad class of urn models. Though stochastic approximation processes can be defined for infinite-dimensional random processes, it is challenging to apply to the $\mathbb{Z}$-urn because new types are introduced as the process progresses. To the best of the author's knowledge, the $\mathbb{Z}$-urn is the first model to be analyzed that involves both multiple drawings and an infinite type of balls.

Motivation for such an urn model comes from philosophy: a reinforcement learning process for a signaling game which models the evolution of a simple language [20]. The reinforcement learning can be viewed as an interacting Pólya-type urn system, and each step involves multiple drawings. In [12], the reinforcement procedure is extended to include the feature of *invention*, in which new strategies are introduced as time progresses, and so the urn model contains an infinite type of balls. Mathematical analysis of the convergence of this process remains incomplete.

1.3. Outline

In Section 2, a martingale is used to prove that the rescaled mean converges. Though the rescaled mean is not bounded, the second moment can be bounded, thus applying Doob's L_2 martingale convergence theorem. In addition, recursive arguments prove that all moments of the limiting random variable exist. In Section 3, the contraction method is outlined and a recursive distributional equation is used to prove Theorems 1 and 2. The final section contains conclusions, future questions, and acknowledgments.

2. A martingale for the mean

In this section, a scaled version of the mean of the empirical distribution determined by the urn is studied. Specifically, recall that $S_n = \sum_{i=0}^{n} X_i$ is the sum of the labels of the first n balls in the urn, and let $A_n := \frac{S_n}{n(n+1)} \in \mathbb{R}^d$.

Lemma 1. *The process* $\{A_n\}_{n \geq \tau_0}$ *is a martingale and* $0 < \liminf_{n \geq \tau_0} \mathbb{E}[A_n(i)^2] \leq \limsup_{n \geq \tau_0} \mathbb{E}[A_n(i)^2] < \infty$ *for* $1 \leq i \leq d$. *Hence,* A_n *converges almost surely and in* $L_2(\mathbb{R}^d)$.

Proof. Conditional on the current configuration of the urn, the expectation of the $(n+1)$th ball added is $\mathbb{E}[X_{n+1} \mid \mu_n] = 2 \cdot \mathbb{E}[Z_n \mid \mu_n] = \frac{2S_n}{n}$, $n \geq \tau_0$. Then,

$$\mathbb{E}[A_{n+1} \mid \mu_n] = \frac{1}{(n+1) \cdot (n+2)} \cdot \mathbb{E}[S_n + X_{n+1} \mid \mu_n]$$

$$= \frac{1}{(n+1) \cdot (n+2)} \left(\frac{(n+2)S_n}{n} \right) = \frac{S_n}{n \cdot (n+1)} = A_n,$$

and thus $\{A_n\}_{n \geq \tau_0}$ is a martingale. Now, the *annealed* expectation of the sum S_n is

$$\mathbb{E}[S_n] = n \cdot (n+1) \, \mathbb{E}[A_n] = n \cdot (n+1) \, \mathbb{E}[A_{\tau_0}] = \frac{n(n+1)}{\tau_0(\tau_0+1)} \cdot S_{\tau_0},$$

for $n > \tau_0$. Note that, for a fixed n, each coordinate $S_n(i)$ is almost surely bounded. However, the size of the largest possible ball in the urn is growing exponentially. To prove convergence of the martingale A_n, we can prove that each coordinate $A_n(i)$ is L_2 bounded. For the following, omit writing the index so that each quantity is one dimensional. Define $R_n := S_n^2 = \sum_{i,j=1}^{n} X_i X_j$, and also define $Q_n := \sum_{i=1}^{n} X_i^2$, for the purpose of writing

$$\mathbb{E}\left[X_{n+1}^2 \mid \mu_n\right] = \frac{1}{n^2} \sum_{i,j=1}^{n} (X_i + X_j)^2 = \frac{2Q_n}{n} + \frac{2R_n}{n^2}.$$

Using this, calculate the conditional expectation

$$\mathbb{E}\left[R_{n+1} \mid \mu_n\right] = \mathbb{E}\left[(S_n + X_{n+1})^2 \mid \mu_n\right]$$

$$= S_n^2 + 2S_n \, \mathbb{E}[X_{n+1} \mid \mu_n] + \mathbb{E}[X_{n+1}^2 \mid \mu_n]$$

$$= R_n \left(1 + \frac{4}{n} + \frac{2}{n^2}\right) + \frac{2Q_n}{n}.$$

Similarly, we can compute

$$\mathbb{E}\left[Q_{n+1} \mid \mu_n\right] = \mathbb{E}\left[Q_n + X_{n+1}^2 \mid \mu_n\right]$$

$$= Q_n \left(1 + \frac{2}{n}\right) + \frac{2R_n}{n^2}.$$

After taking expectations, for $n > \tau_0$ we get the system of recursions

$$\begin{pmatrix} \mathbb{E}[R_{n+1}] \\ \mathbb{E}[Q_{n+1}] \end{pmatrix} = \begin{pmatrix} 1 + \frac{4}{n} + \frac{2}{n^2} & \frac{2}{n} \\ \frac{2}{n^2} & 1 + \frac{2}{n} \end{pmatrix} \cdot \begin{pmatrix} \mathbb{E}[R_n] \\ \mathbb{E}[Q_n] \end{pmatrix}.$$

This recursion works to prove that $\mathbb{E}[R_n] \leq Cn^4$ and $\mathbb{E}[Q_n] \leq 2Cn^3$, for some constant C. Indeed, assume for induction the statement is true for n. Then,

$$\mathbb{E}[R_{n+1}] \leq \left(1 + \frac{4}{n} + \frac{2}{n^2}\right)Cn^4 + \frac{2}{n}2Cn^3$$

$$= C\left(n^4 + 4n^3 + 6n^2\right) \leq C(n+1)^4,$$

and also

$$\mathbb{E}[Q_{n+1}] \leq \frac{2}{n^2}Cn^4 + \left(1 + \frac{2}{n}\right)2Cn^3$$

$$= 2C\left(n^3 + 3n^2\right) \leq 2C(n+1)^3.$$

The induction can be initialized for $C = R_{\tau_0}$.

Note that $A_n^2 = R_n/\left(n^2(n+1)^2\right)$, thus $\limsup_{n \geq 0} \mathbb{E}[A_n^2] \leq C$, and so by the L_2 martingale convergence theorem (e.g. [22, p. 111]) A_n converges almost surely and in L_2. To prove that $\inf_{n \geq \tau_0} \mathbb{E}[A_n^2] > 0$, observe that

$$\mathbb{E}[A_n^2] \geq \prod_{k=\tau_0}^{n-1} \frac{1}{(k+2)^2}\left(k^2 + 4k + 2\right) \cdot A_{\tau_0}$$

$$= \prod_{k=\tau_0}^{n-1}\left(1 - \frac{2}{k^2 + 4k + 4}\right) \cdot A_{\tau_0}$$

$$\geq \prod_{k=\tau_0}^{\infty}\left(1 - \frac{2}{k^2 + 4k + 4}\right) \cdot A_{\tau_0}.$$

Since $\sum_{k=0}^{\infty} 2/(k^2 + 4k + 4) < \infty$, the infinite product converges to a positive value. □

The previous result states that the limit $\lim_{n \to \infty} A_n =: A$ exists almost surely for each realization of the urn process, and furthermore since $\mathbb{E}[A^2] > 0$ the limit is non-trivial. Note that this result holds for *any* initial configuration of the urn. The exact distribution of the random variable A is undetermined.

3. Recursive distributional equations

The contraction method is a tool in the probabilistic analysis of algorithms for which a recursive distributional equation exists. It was developed in [18] to analyze the number of comparisons required by the Quicksort sorting algorithm. A more generalized development was presented in [17]. More limit theorems, along with numerous applications to recursive algorithms and random trees, were discussed in [15]. The contraction method has been used to prove convergence in distribution of general Pólya-type urns [7] with a finite number of colors, using recursive distributions for the number of balls of each color, started from some initial configuration.

The idea of the contraction method is to find a fixed-point equation from the recursive distributional equations. Just as in a deterministic recursion, one expects the sequence to converge to the fixed point of the transformation. However, in the setting of random variables care must be taken in choosing a metric in which the recursive equation is a contraction. The correct metric is introduced in Section 3.1, which also contains the necessary distributional identities. The main contraction method proof is in Section 3.2, and it is applied in Section 3.3 with the correct scaling for the $\mathbb{Z}$-urn.

3.1. Setup

Let $\mathcal{M}^{\mathbb{R}^d}$ denote the space of all probability distributions on $\mathbb{R}^d$ with the Borel σ-field. For a vector $\mathbf{x} \in \mathbb{R}^d$, let $\|\mathbf{x}\|$ denote the usual Euclidean norm. Let $\mu \in \mathbb{R}^d$, and define the following subspaces:

$$\mathcal{M}_s^{\mathbb{R}^d} := \left\{ \mathcal{L}(X) \in \mathcal{M}^{\mathbb{R}^d} : \mathbb{E}\left[\|X\|^s\right] < \infty \right\},$$

$$\mathcal{M}_s^{\mathbb{R}^d}(\mu) := \left\{ \mathcal{L}(X) \in \mathcal{M}_s^{\mathbb{R}^d} : \mathbb{E}[X] = \mu \right\}.$$

For $1 \le p < \infty$ and two probability distributions $\nu, \rho \in \mathcal{M}_p^{\mathbb{R}^d}$, the minimal L_p (or Wasserstein L_p) metric, ℓ_p, is defined as

$$\ell_p(\nu, \rho) := \inf\left\{ \|V - W\|_p : \mathcal{L}(V) = \nu, \mathcal{L}(W) = \rho \right\}, \tag{2}$$

where $\|V - W\|_p := (\mathbb{E}[\|V - W\|^p])^{1/p}$ is the usual L_p-distance. For random variables X, Y, we will write $\ell_p(X, Y)$ to mean $\ell_p(\mathcal{L}(X), \mathcal{L}(Y))$. The infimum in (2) is obtained for all $\nu, \rho \in \mathcal{M}_1^{\mathbb{R}}$, and the random variables $V \sim \nu$, $W \sim \rho$ which achieve the infimum are called the *optimal coupling*. For proofs of this fact and the following lemma, see [3, Section 8].

Lemma 2. *Suppose $\{X_n\}$ is a sequence of random variables in $\mathcal{M}_p^{\mathbb{R}^d}$ and $X \in \mathcal{M}_p^{\mathbb{R}^d}$. Then $\ell_p(X_n, X) \to 0$ if and only if $X_n \to X$ weakly and $\mathbb{E}[\|X_n\|^p] \to \mathbb{E}[\|X\|^p]$.*

The contraction method gives a distributional identity that the limiting random variable must satisfy. For the purpose of characterizing the limiting random variable, note the following distributional identity that exponential distributions satisfy. This can easily be calculated, or see [8].

Lemma 3. *Let Y be a random variable satisfying $Y \ge 0$ with probability 1 and $Y \overset{d}{=} U \cdot (Y^1 + Y^2)$, where U is uniformly distributed on $[0, 1]$ and $Y^1 \overset{d}{=} Y^2 \overset{d}{=} Y$, and the three random variables U, Y^1, Y^2 are independent. Then Y has an exponential distribution.*

In addition, we need the following characterization of a random vector in which each coordinate is a multiple of the same one-dimensional gamma random variable.

Lemma 4. *Let $\mathbf{X} \in \mathbb{R}^d$ be a random vector satisfying $\mathbf{X}(i) \ge 0$ with probability 1 for all $1 \le i \le d$ and $\mathbf{X} \overset{d}{=} U_1 \cdot \mathbf{X}^1 + U_2 \cdot \mathbf{X}^2$, where U_1, U_2 are Uniform($[0, 1]$) random variables and $\mathbf{X}^1 \overset{d}{=} \mathbf{X}^2 \overset{d}{=} \mathbf{X}$, and $U_1, U_2, \mathbf{X}^1$, and $\mathbf{X}^2$ are independent. Suppose $\mathbf{m} = \mathbb{E}[\mathbf{X}] \in \mathbb{R}^d$. Then $\mathbf{X} \sim G \cdot \mathbf{m}$, where $G \sim \text{Gamma}(2, 1)$.*

Proof. To determine the distribution of $\mathbf{X}$, we can calculate the joint characteristic function. For $\mathbf{t} = (t_1, \ldots, t_d) \in \mathbb{R}^d$, the result is

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \mathbb{E}\left[\exp\left(i \sum_{j=1}^d t_j \mathbf{X}(j)\right)\right] = \mathbb{E}\left[\exp\left(i \sum_{j=1}^d t_j \left(U_1 \mathbf{X}(j)^1 + U_2 \mathbf{X}(j)^2\right)\right)\right]$$

$$= \mathbb{E}\left[\exp\left(iU_1 \sum_{j=1}^d t_j \mathbf{X}(j)^1\right)\right] \cdot \mathbb{E}\left[\exp\left(iU_2 \sum_{j=1}^d t_j \mathbf{X}(j)^2\right)\right]$$

$$= \left(\int_0^1 \mathbb{E}\left[\exp\left(iu \sum_{j=1}^d t_j \mathbf{X}(j)\right)\right] du\right)^2 = \left(\int_0^1 \varphi_{\mathbf{X}}(u \cdot \mathbf{t}) \, du\right)^2.$$

The function $\varphi_{\mathbf{X}}$ also satisfies

$$\begin{cases} \varphi_{\mathbf{X}}(\mathbf{0}) = 1, \\ \frac{\partial}{\partial t_j} \varphi_{\mathbf{X}}(\mathbf{0}) = \mathbb{E}\left[i \cdot \mathbf{X}(j)\right] = 2i \cdot \mathbf{A}(j), \quad 1 \le j \le d. \end{cases}$$

The unique solution to this integral equation is

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \left(1 - i \cdot \sum_{j=1}^{d} t_j \mathbf{A}(j)\right)^{-2}. \tag{3}$$

Equation (3) is exactly the characteristic function for the random vector $G \cdot \mathbf{A}$, where G is a Gamma(2, 1) random variable. □

3.2. Contraction method result

To write down the defining recursive distributional equation for the $\mathbb{Z}$-urn, recall that the time index is started at τ_0. If the initial configuration is $\mathbf{U} = \{x_1, \dots, x_{\tau_0}\}$, define $X_1 = x_1$, $X_2 = x_2, \dots, X_{\tau_0} = x_{\tau_0}$. The newly added ball X_n is the sum of two randomly chosen balls that were added in the past. This gives the distributional identity

$$X_n \overset{\mathrm{d}}{=} X_{I_n^1}^1 + X_{I_n^2}^2, \qquad n > \tau_0, \tag{4}$$

where, for $i = 1, 2$, $\{X_n^i\}$ are i.i.d. copies of $\{X_n\}$ and I_n^i are indices drawn uniformly from $\{1, \dots, n-1\}$. To prove convergence, it is necessary to scale X_n appropriately. Define $\tilde{X}_n = \frac{X_n}{n}$. Then (4) becomes

$$\tilde{X}_n \overset{\mathrm{d}}{=} B_{I_n^1} \cdot \tilde{X}_{I_n^1}^1 + B_{I_n^2} \cdot \tilde{X}_{I_n^2}^2, \tag{5}$$

where $B_{I_n^i} = \frac{I_n^i}{n}$. The correct fixed-point equation from this recursion arises from the fact that $\frac{I_n^i}{n}$ converges to a Uniform([0, 1]) random variable in the ℓ_2 metric. This can be proven by constructing the variables as follows. Let U be Uniform([0, 1]) and define

$$I_n = \sum_{i=0}^{n-1} i \cdot \mathbf{1}\left\{\frac{i}{n} \le U < \frac{i+1}{n}\right\}.$$

Then $I_n \sim$ Uniform($\{1, \dots, n-1\}$). Clearly, $\frac{I_n}{n} \to U$ weakly, and since the second moment converges to $1/3$, the convergence is in ℓ_2 by Lemma 2. The following is a general result about random variables satisfying the distributional recursion (5). The result is a special case of Theorem 4.1 and Corollary 4.2 from [14], though since the proof is short it is included for completeness.

Theorem 3. *Let $\{\tilde{X}_n\}_{n \ge 1}$ be a sequence with $\tilde{X}_n \in \mathcal{M}_2^{\mathbb{R}^d}$ and $\lim_{n \to \infty} \mathbb{E}[\tilde{X}_n] = \mu$. Let $\{B_n\}_{n \ge 1}$ be a sequence with $B_n \in \mathcal{M}_2^{\mathbb{R}}$ and $\mathbb{E}[B_n] \to 1/2$, and if $I_n \sim$ Uniform($\{1, \dots, n-1\}$), then the sequence $\{B_{I_n}\}_{n \ge 1}$ converges in ℓ_2 to a Uniform([0, 1]) variable.*

For $i = 1, 2$, let I_n^i be independent Uniform($\{1, \dots, n-1\}$) random variables, $\{B_n^i\}$ independent copies of $\{B_n\}$, and $\{\tilde{X}_n^i\}$ independent copies of $\{X_n\}$. Suppose that, for all $n \ge \tau_0$,

$\tilde{X}_n \stackrel{\text{d}}{=} B^1_{I^1_n} \cdot \tilde{X}^1_{I^1_n} + B^2_{I^2_n} \cdot \tilde{X}^2_{I^2_n}$. *Then $\tilde{X}_n$ converges in distribution to a random variable $X \in \mathcal{M}^{\mathbb{R}^d}_2(\mu)$ which satisfies*

$$X \stackrel{\text{d}}{=} U_1 \cdot X^1 + U_2 \cdot X^2, \tag{6}$$

where U_1, U_2 are Uniform([0, 1]) *random variables, $X \stackrel{\text{d}}{=} X^1 \stackrel{\text{d}}{=} X^2$, and U_1, U_2, X^1, and X^2 are independent.*

Proof. The result is proven for $d = 1$, the multidimensional result being analogous. Let X be a random variable satisfying (6). Let $(B_{I^1_n}, U_1)$ and $(B_{I^2_n}, U_2)$ be the optimal couplings, such that $\ell^2_2(B_{I^i_n}, U_i) = \mathbb{E}[(B_{I^i_n} - U_i)^2] \to 0$. Define $d_n = \ell^2_2(\tilde{X}_n, X)$. Then

$$\begin{aligned}
d_n &\leq \mathbb{E}\left[((B^1_{I^1_n}\tilde{X}^1_{I^1_n} + B^2_{I^2_n}\tilde{X}^2_{I^2_n}) - (U_1X^1 + U_2X^2))^2\right] \\
&= 2\,\mathbb{E}\left[(B_{I^1_n}\tilde{X}^1_{I^1_n} - U_1X^1)^2\right] + 2\,\mathbb{E}\left[(B_{I^1_n} \cdot \tilde{X}^1_{I^1_n} - U_1X^1)(B_{I^2_n} \cdot \tilde{X}^2_{I^2_n} - U_2X^2)\right].
\end{aligned} \tag{7}$$

Observe that the assumptions imply, for $I_n \sim$ Uniform($\{1, \ldots, n-1\}$), that

$$\mathbb{E}[B_{I_n}\tilde{X}_{I_n}] = \sum_{j=0}^{n-1} \frac{1}{n}\,\mathbb{E}[B_j]\,\mathbb{E}[\tilde{X}_j] \to \frac{\mu}{2}.$$

This, along with the independence of the variables involved, gives that the second term in (7) converges to 0 as $n \to \infty$. Now, expanding the first term, and disregarding superscripts,

$$\begin{aligned}
\mathbb{E}\left[(B_{I_n}\tilde{X}_{I_n} - UX)^2\right] &= \mathbb{E}\left[\left(B_n(\tilde{X}_{I_n} - X) + X(B_{I_n} - U)\right)^2\right] \\
&= \mathbb{E}\left[(B_{I_n})^2(\tilde{X}_{I_n} - X)^2\right] + \mathbb{E}\left[X^2(B_{I_n} - U)^2\right] \\
&\quad + 2\,\mathbb{E}\left[B_{I_n}X(\tilde{X}_{I_n} - X)(B_{I_n} - U)\right].
\end{aligned} \tag{8}$$

Note that $\mathbb{E}\left[X^2(B_{I_n} - U)^2\right] = \mathbb{E}[X^2] \cdot \mathbb{E}\left[(B_{I_n} - U)^2\right] \to 0$ by construction. Also, since $\mathbb{E}[B_{I_n}(B_{I_n} - U)] \to 0$, the final term in (8) converges to 0 as $n \to \infty$. Thus, the final term to analyze is

$$\begin{aligned}
\mathbb{E}\left[(B_{I_n})^2 \cdot (\tilde{X}_{I_n} - X)^2\right] &= \sum_{j=1}^{n-1} \mathbb{E}\left[(B_{I_n})^2 \cdot (\tilde{X}_{I_n} - X)^2 \mid I_n = j\right] \cdot \mathbb{P}(I_n = j) \\
&= \sum_{j=1}^{n-1} \mathbb{E}\left[\frac{j^2}{n^2}(\tilde{X}_j - X)^2\right] \cdot \frac{1}{n-1} = \sum_{j=1}^{n-1} \frac{j^2}{n^2(n-1)} \cdot d_j.
\end{aligned}$$

This implies that

$$\limsup_{n\to\infty} d_n \leq 2 \limsup_{n\to\infty} \sum_{j=1}^{n-1} \frac{j^2}{n^2(n-1)} \cdot d_j \leq \frac{2}{3} \limsup_{n\to\infty} d_n.$$

This gives a contradiction unless $\limsup_{n\to\infty} d_n = 0$. Since $d_n \geq 0$, this proves that $\lim_{n\to\infty} d_n = 0$, which is the desired convergence. $\square$

Remark 2. Theorem 3 can also be proven by applying [17, Theorem 4.1] with the Zolotarev metric ζ_2. For $\nu, \rho \in \mathcal{M}^{\mathbb{R}}$, the Zolotarev distance ζ_s, $s \geq 0$, is defined as $\zeta_s(X, Y) := \zeta_s(\mathcal{L}(X),$

$\mathcal{L}(Y)) := \sup_{f \in \mathcal{F}_s} |\mathbb{E}[f(X) - f(Y)]|$, where $s = m + \alpha$ with $0 < \alpha \le 1$, $m \in \mathbb{N}_0$, and $\mathcal{F}_s := \{f \in C^m(\mathbb{R}, \mathbb{R}) : |f^{(m)}(x) - f^{(m)}(y)| \le |x - y|^\alpha\}$. The use of Zolotarev metrics in contraction methods was developed in [17] because in some settings the recursive distributional equations do not give a strict contraction in ℓ_p. Most important of these settings is fixed-point equations that occur for normal distributions, which can be handled by ζ_s for $s > 2$.

3.3. Limit laws

Now, Theorem 3 can be applied to prove the result about the $\mathbb{Z}$-urn. The one technicality is that Theorem 3 requires that the random variables in question have the same mean. If X_n was just defined as the nth ball added to the urn, then this would not be true. Thus, it is necessary to rescale by the mean of X_n, and use the result from Section 2 that the mean converges. It is important to remember the quenched setting; all variables are conditional on the *same* underlying urn process.

Proof of Theorems 1 and 2. We again restrict to a one-dimensional process. Let $\{\mu_n\}_{n \ge \tau_0}$ be a realization of the $\mathbb{Z}$-urn process. For $n > \tau_0$, let X_n denote the nth ball added to the urn, conditional on the urn μ_{n-1}. Recall that $\mathbb{E}[X_n \mid \mu_{n-1}] = 2n \cdot A_{n-1}$, where $A_n := \frac{S_n}{n(n+1)}$. Thus, define $\tilde{X}_n := \frac{X_n}{n}$; then $\lim_{n \to \infty} \mathbb{E}[X_n \mid \mu_{n-1}] = 2 \lim_{n \to \infty} A_n = 2A$ exists by Lemma 1. The recursive distributional equation satisfied by $\tilde{X}_n$ is $\tilde{X}_n \stackrel{d}{=} \frac{I_n^1}{n} \cdot \tilde{X}_{I_n^1}^1 + \frac{I_n^2}{n} \cdot \tilde{X}_{I_n^2}^2$ for I_n^i, $i = 1, 2$, i.i.d. Uniform($\{1, \ldots, n-1\}$). As remarked before Theorem 3, the random variables $\frac{I_n^i}{n}$ converge in ℓ_2 to the Uniform([0, 1]) distribution. Thus, taking $B_{I_n}^i = I_n^i/n$, Theorem 3 can then be applied to get that $\tilde{X}_n$ converges in distribution to some X with $\mathbb{E}[X] = 2 \cdot A$ which satisfies (6). Finally, Lemma 4 says that the distribution of X/A must be Gamma with shape 2. Now, recalling that Z_n denotes a draw from the urn μ_n^d, we have the distributional identity $Z_n \stackrel{d}{=} X_{I_n}$, where I_n is a uniformly chosen index from $\{1, \ldots, n\}$. Again scaling $\tilde{Z}_n = \frac{Z_n}{n}$, we get $\tilde{Z}_n \stackrel{d}{=} \frac{I_n}{n} \cdot \tilde{X}_{I_n}$. Since $\tilde{X}_{I_n}/A$ converges in ℓ_2 to $X \sim$ Gamma(2, 1), this proves that $\tilde{Z}_n$ converges in distribution to $U \cdot X$, where $U \sim$ Uniform([0, 1]). By Lemma 3, it follows that the limit point has exponential distribution. $\square$

4. Conclusion and future questions

This paper has introduced a new urn model, extending the product urn model from [19] for finite groups to the integers or real numbers. Among the vast collection of urn models, the $\mathbb{Z}$-urn is interesting because it contains two properties which are challenging to study: drawing more than one ball, and an infinite number of types of balls. It is proven that the rescaled empirical distribution defined by the urn converges to a multiple of an exponential distribution. The multiple, determined by the mean of the limiting urn, is a random variable; the main remaining question is the distribution of this random variable.

Recall the notation $S_n = \sum_{i=1}^n X_n$ for the sum of the labels in the urn at step n, and $A_n = \frac{S_n}{n(n+1)}$. The process $\{A_n\}$ is a martingale and converges almost surely to some limit A. The distribution of A can be explored using simulations. From Section 2, the mean of A is

$$\mathbb{E}[A] = \frac{S_{\tau_0}}{\tau_0(\tau_0 + 1)}.$$

That is, A depends on both the sum of the initial labels in the urn, and the number τ_0 of balls started. Figure 2(a) shows a histogram for the result for A for an urn started with a ball labeled
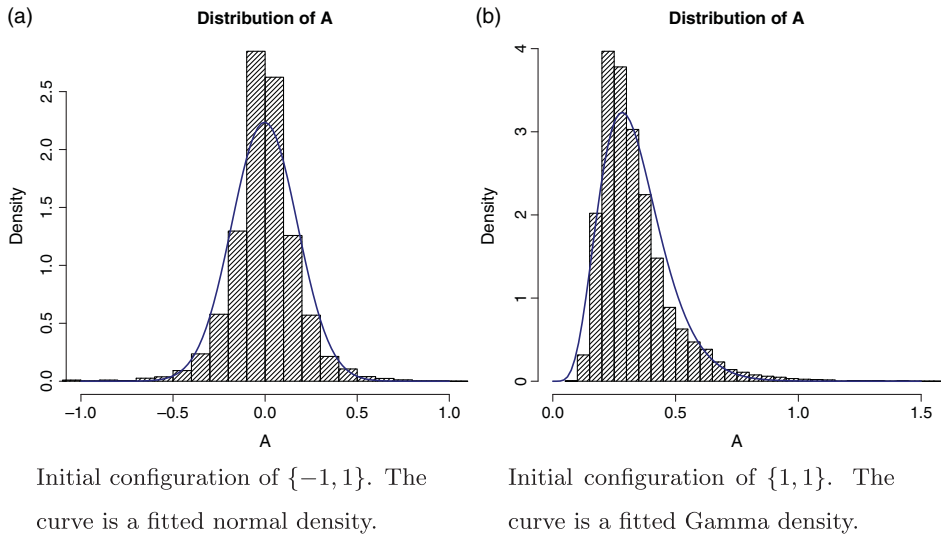
FIGURE 2: Results for *A* from simulations of the $\mathbb{Z}$-urn model from two different initial configurations. An urn realization was run for 5000 rounds and the value of A_n was recorded. A total of 5000 realizations of each urn process were run to display the empirical distribution of *A*.

-1 and a ball labeled 1. The distribution is of course centered around 0, but it is not normal. Figure 2(b) exhibits the results for *A* for urns started with two 1 labels. The mean in this case is $2/(2 \cdot 3) = 1/3$. A χ^2 test for the null hypothesis that a Gamma distribution fits the data resulted in a *p*-value < 0.001, indicating that Gamma is not a good fit. It would be interesting to explore this distribution in future work, though even proving it is continuous seems to be a challenge.

Another interesting problem would be to study the model for different infinite groups, for example $\mathbb{R}$ with the multiplication operation, or the Heisenberg group. The result of an exponential limit law is very specific to the addition process, so the limits for different groups and group operations could be curious.

Acknowledgements

References

[1] ABRAMS, A., LANDAU, H., LANDAU, Z., POMMERSHEIM, J. AND ZASLOW, E. (2007). Random multiplication approaches uniform measure in finite groups. *J. Theoret. Prob.* **20**, 107–118.

[2] BANDYOPADHYAY, A. AND THACKER, D. (2017). Pólya urn schemes with infinitely many colors. *Bernoulli* **23**, 3243–3267.

[3] BICKEL, P. J. AND FREEDMAN, D. A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* **9**, 1196–1217.

[4] BLACKWELL, D. AND MACQUEEN, J. B. (1973). Ferguson distributions via Pólya urn schemes. *Ann. Statist.* **1**, 353–355.

[5] JANSON, S. (2018). A.s. convergence for infinite colour Pólya urns associated with random walks. Preprint, arXiv:1803.04207.

[6] JOHNSON, N. L. AND KOTZ, S. (1977). *Urn Models and Their Application*. John Wiley & Sons, New York.

[7] KNAPE, M. AND NEININGER, R. (2014). Pólya urns via the contraction method. *Combinatorics Prob. Comput.* **23**, 1148–1186.

[8] KOTZ, S. AND STEUTEL, F. W. (1988). Note on a characterization of exponential distributions. *Statist. Prob. Lett.* **6**, 201–203.

[9] KUBA, M., MAHMOUD, H. AND PANHOLZER, A. (2013). Analysis of a generalized Friedman's urn with multiple drawings. *Discrete Appl. Math.* **161**, 2968–2984.

[10] KUBA, M. AND MAHMOUD, H. M. (2017). Two-color balanced affine urn models with multiple drawings. *Adv. Appl. Math.* **90**, 1–26.

[11] LASMAR, N., MAILLER, C. AND SELMI, O. (2018). Multiple drawing multi-colour urns by stochastic approximation. *J. Appl. Prob.* **55**, 254–281.

[12] MCKENZIE ALEXANDER, J., SKYRMS, B. AND ZABELL, S. (2012). Inventing new signals. *Dynamic Games Appl.* **2**, 129–145.

[13] MAHMOUD, H. M. (2009). *Pólya Urn Models*. CRC Press, Boca Raton, FL.

[14] NEININGER, R. (2001). On a multivariate contraction method for random recursive structures with applications to Quicksort. *Random Structures Algorithms* **19**, 498–524.

[15] NEININGER, R. AND RÜSCHENDORF, L. (2004). A general limit theorem for recursive algorithms and combinatorial structures. *Ann. Appl. Prob.* **14**, 378–418.

[16] PEMANTLE, R. (2007). A survey of random processes with reinforcement. *Prob. Surv.* **4**, 1–79.

[17] RACHEV, S. T. AND RÜSCHENDORF, L. (1995). Probability metrics and recursive algorithms. *Adv. Appl. Prob.* **27**, 770–799.

[18] RÖSLER, U. (1991). A limit theorem for "Quicksort." *RAIRO Inform. Théor. Appl.* 25, 85–100.

[19] SIEGMUND, D. AND YAKIR, B. (2005). An urn model of Diaconis. *Ann. Prob.* **33**, 2036–2042.

[20] SKYRMS, B. (2010). *Signals: Evolution, Learning, and Information*. Oxford University Press.

[21] THACKER, D. (2015). Infinite color urn models. PhD thesis. Indian Statistical Institute, Dehli Centre.

[22] WILLIAMS, D. (1991). *Probability with Martingales*. Cambridge University Press.