PÓLYA URNS WITH IMMIGRATION AT RANDOM TIMES

Erol Peköz, Adrian Röllin and Nathan Ross

Boston University, National University of Singapore, University of Melbourne

Abstract

We study the number of white balls in a classical Pólya urn model with the additional feature that, at random times, a black ball is added to the urn. The number of draws between these random times are i.i.d. and, under certain moment conditions on the inter-arrival distribution, we characterize the limiting distribution of the (properly scaled) number of white balls as the number of draws goes to infinity. The possible limiting distributions obtained in this way vary considerably depending on the inter-arrival distribution and are difficult to describe explicitly. However, we show that the limits are fixed points of certain probabilistic distributional transformations, and this fact provides a proof of convergence and leads to properties of the limits. The model can alternatively be viewed as a preferential attachment random graph model where added vertices initially have a random number of edges, and from this perspective, our results describe the limit of the degree of a fixed vertex.

Keywords: Pólya urns; distributional convergence; distributional fixed point equation; preferential attachment random graph.

1 INTRODUCTION AND MAIN RESULTS

Pólya urn schemes form a rich class of fundamental probability models with a long history going back to Eggenberger and Pólya (1923) and extending to present day research. The standard general model is a recursive Markov process which begins with "balls" of different colors in an urn, and at each step a ball is drawn randomly from the urn and returned along with the addition or removal of some prescribed number of balls of each color. The popularity of these models is due to the fact that variations of this basic Pólya urn reinforcement mechanism appear in applications in biology, computer science, statistics, and elsewhere; see Pemantle (2007) and Mahmoud (2009). Here we study the limiting behavior of a new urn model that is a simple variation of the classical Pólya urn and which arises naturally from a certain random graph model. Outside of application, the model is intrinsically interesting since the limiting behavior is subtle and intricately related to our method of proof, and other more standard techniques for analyzing urn models do not naturally apply (a more thorough discussion of existing literature and these other methods of proof can be found in Section 1.3). We now define our model and then state our main results.

Let $\tau_1, \tau_2, \dots$ be i.i.d. non-negative integer valued random variables having distribution $\pi = (\pi_k)_{k\geqslant 0}$, where we assume throughout that $\pi_0 < 1$, and let $T_j = \sum_{i=1}^{j} \tau_i$. It is helpful to think of the τ_i as *inter-arrival* interval lengths in a renewal process, so that the T_j are the arrival times. Consider the following Pólya urn model. Initially, there are b black balls and w white balls. At each step, a ball is drawn and replaced along with another of the same color. Additionally, after draws $T_1, T_2, \dots$, regardless of the outcome of the draw, a single extra black ball is added to the urn. Note that if $\tau_i = 0$, so $T_i = T_{i+1}$, then more than one black ball can be added to the urn between draws.

For example, if $(\tau_1, \dots, \tau_5) = (1, 3, 0, 0, 4)$, then $(T_1, \dots, T_5) = (1, 4, 4, 4, 8)$. At Step 1, a regular Pólya urn step is performed (that is, a ball is drawn and replaced along with a ball of the same color), and then one additional black ball is added since $T_1 = 1$. At Steps 2 and 3, regular Pólya urn steps are performed with no added black ball. Then, at Step 4, a regular Pólya urn step is performed and then three additional black balls are added, since $T_2 = T_3 = T_4 = 4$. Then, four regular Pólya urn steps are performed after which another black ball is added ($T_5 = 8$). Note that given a black ball is added at a particular time step, the total number of balls added at that step has a geometric distribution (support starting at 1) with success probability $1 - \pi_0$. Note also that the number of black balls after Step 0 is not necessarily b; this happens if $\tau_1 = 0$, in which case black balls are added already before the first draw and replacement step is performed.

We study the distribution of the number of white balls in the urn after n steps in this model, denoted by $\mathcal{P}^{\pi}\binom{b}{w};n)$. In particular we show that $\mathcal{P}^{\pi}\binom{b}{w};n)$ properly scaled converges in distribution to a non-standard limit law. The limits for deterministic π are studied in Janson (2006) for $\pi_1 = 1$ and Peköz et al. (2016) for $\pi_k = 1$ with $k > 1$. Before stating the result, we need to describe the limit.

1.1 Urn limit laws

Let $v > 0$, and let $a_1, a_2, \dots$ be a sequence of non-negative numbers so that $a_k > 0$ for at least one $k \geqslant 1$. Let

$$A(x) = \sum_{k \geqslant 1} a_k x^k, \tag{1.1}$$

and assume the radius of convergence $\rho = \sup\{x \geqslant 0 : A(x) < \infty\}$ is either positive or infinite. For such $(a_k)_{k \geqslant 1}$, we define the probability density

$$u(x) = c\, x^{v-1} \exp\Big\{-v \int_0^x \frac{A(t)}{t} dt\Big\}, \qquad 0 < x < \rho, \tag{1.2}$$

where c is an appropriate normalising constant depending on v and $(a_k)_{k \geqslant 1}$, and we denote the corresponding probability distribution by $\mathrm{UL}\big(v; (a_k)_{k \geqslant 1}\big)$. Specific instances of the UL family include many standard non-negative continuous distributions such as the exponential, Rayleigh, absolute normal, gamma, beta, and roots of gamma variables. We first establish that (1.2) is indeed a proper probability density and derive some basic properties of the laws UL.

Lemma 1.1. *Under the assumptions and notation above, the function $u(x)$ defined by (1.2) is a probability density for an appropriate normalisation c. Moreover, $Z \sim \mathrm{UL}\big(v; (a_k)_{k \geqslant 1}\big)$ has finite moments $\mu_k = \mathbb{E}Z^k$ of all orders, which satisfy the relation*

$$\mu_k = \frac{v}{v+k} \sum_{l \geqslant 1} a_l \mu_{k+l}, \qquad \text{for all } k \geqslant 0. \tag{1.3}$$

Furthermore, for $\theta > 0$,

$$\theta Z \sim \mathrm{UL}\big(v; (\theta^{-k} a_k)_{k \geqslant 1}\big). \tag{1.4}$$

Proof. For the first assertion, we show that the density given at (1.2) has finite integral over $(0, \rho)$. Let $k_0 \geqslant 1$ be such that $a_{k_0} > 0$. Observe that

$$\Phi(x) := \int_0^x \frac{A(t)}{t} dt = \sum_{k \geqslant 1} \frac{a_k}{k} x^k$$

so that

$$x^{v-1}\exp\{-v\Phi(x)\} \leqslant x^{v-1}\exp\{-va_{k_0}x^{k_0}/k_0\},$$

which clearly has finite integral. Replacing x^{v-1} by any arbitrary power, we also conclude that all moments are finite. After noting that since the coefficients of (1.1) are all non-negative, Titchmarsh (1958, 7.21) implies that $\lim_{x\to\rho^-} A(x) = \infty$ and hence also $\lim_{x\to\rho^-}\Phi(x) = \infty$, the relation (1.3) is just integration by parts; we have

$$\begin{aligned}\mu_k &= c\int_0^\rho x^{k+v-1}\exp\{-v\Phi(x)\}dx = c\int_0^\rho \frac{x^{k+v}}{k+v}\cdot v\frac{A(x)}{x}\exp\{-v\Phi(x)\}dx\\ &= \frac{v}{k+v}c\int_0^\rho x^{k+v-1}A(x)\exp\{-v\Phi(x)\}dx = \frac{v}{k+v}\sum_{l\geqslant 1} a_l\int_0^\rho x^{k+l}u(x)dx.\end{aligned}$$

Interchange of summation and integration is justified by the monotone convergence theorem, since all coefficients are non-negative. The final assertion (1.4) is straightforward. □

1.2 Limit results for urns with random immigration

To state our first main result, let Beta(α,β), where α and β are positive numbers, denote the law of the beta distribution supported on $(0,1)$ with density proportional to $x^{\alpha-1}(1-x)^{\beta-1}$, and interpret Beta$(\alpha,0)$ as the point mass at 1. We first consider the problem of convergence of moments of the (appropriately scaled) number of white balls in the urn. In what follows, we interpret $\frac{\infty}{\infty+1}$ as 1. Here and below, C is a generic constant that may change from line to line.

Theorem 1.2. *Let b and w be positive integers, let π be a probability distribution on the non-negative integers with mean $0<\mu\leqslant\infty$, and let $\tau\sim\pi$. If k is an integer such that either*

(*a*) *$\mathbb{E}\tau^p<\infty$ for some $p>1$, and $1\leqslant k<(\frac{p}{2}-1)(\mu+1)-1$, or*

(*b*) *there is $\varepsilon>0$ such that $\mathbb{P}[\tau>n]\geqslant Cn^{-(1-\varepsilon)}$ for n large enough, and $k\geqslant 1$,*

then there is a positive constant $m_k(b,w,\pi)$ such that, for $X_n\sim\mathcal{P}^\pi\binom{b}{w};n)$, we have

$$\mathbb{E}\Big\{\Big(\frac{X_n}{n^{\mu/(\mu+1)}}\Big)^k\Big\}\to m_k(b,w,\pi)\qquad \text{as } n\to\infty. \tag{1.5}$$

We now formulate the main distributional convergence result which essentially says that when $b=1$, the scaled urn limits are of the form $\mathrm{UL}\big(w;(a_k)_{k\geqslant1}\big)$ for appropriate choice of $(a_k)_{k\geqslant1}$, and when $b>1$, the limits are in the same family up to multiplication by an independent beta variable.

Theorem 1.3. *Let b and w be positive integers, and let π be a probability distribution on the non-negative integers with mean $0<\mu\leqslant\infty$, and let $\tau\sim\pi$. Assume that either*

(*a*) *$\mathbb{E}\tau^p<\infty$ for all $p\geqslant1$, or*

(*b*) *there is $\varepsilon>0$ such that $\mathbb{P}[\tau>n]\geqslant Cn^{-(1-\varepsilon)}$ for n large enough,*

and let $m_k(1,b+w-1,\pi)$, $k\geqslant1$, be as in Theorem 1.2. Set

$$a_k=\frac{\pi_{k-1}}{m_k(1,b+w-1,\pi)},\qquad \text{for all } k\geqslant 1, \tag{1.6}$$

and let $Z \sim \mathrm{UL}\big(b+w-1; (a_k)_{k\geqslant 1}\big)$. *Then*

$$\mathbb{E}Z^k = m_k(1, b+w-1, \pi), \qquad \textit{for all } k \geqslant 1, \tag{1.7}$$

and, with $X_n \sim \mathcal{P}^\pi\binom{b}{w}; n)$,

$$\mathscr{L}\left(\frac{X_n}{n^{\mu/(\mu+1)}}\right) \to \mathscr{L}(BZ) \qquad \textit{as } n \to \infty, \tag{1.8}$$

where $B \sim \mathrm{Beta}(w, b-1)$ *is independent of* Z.

Our expressions for the moments of (1.7) are not explicit and so then neither are the parameters of the limits, which leads to many intriguing open questions; see Section 3. In the case where π is deterministic, the limiting distributions can be described explicitly in a number of ways, see Janson (2006), Peköz et al. (2013a, 2016), and so it is interesting that adding randomness in this way leads to limiting distributions that are complicated and difficult to describe.

The moment results of Theorem 1.2 follow by first deriving formulas conditional on the partial sums $(T_1, T_2, \ldots)$ of the i.i.d. inter-arrival times $\tau_1, \tau_2, \ldots$, and then using classical moment and concentration inequalities for such quantities. The moment results show that the sequence $n^{-\mu/(\mu+1)}X_n$, $n \geqslant 1$, is tight as long as we have either $\mu < \infty$ and $\mathbb{E}\tau^{6/(\mu+1)+2} < \infty$ or $\mathbb{P}[\tau > n] \geqslant Cn^{-(1-\varepsilon)}$, and so in these cases, a distributional limit follows by showing uniqueness of subsequential limits. For the case $b = 1$, we are able to show that in the two cases just described, any subsequential limit is a fixed point (unique given moments) of a certain distributional transformation which we describe in Section 2 below.

The organization of the remainder of the paper is as follows. We finish this section with a discussion of related literature and then provide a connection between our model and preferential attachment graphs with random number of initial attachments for each vertex. In Section 2 we describe the distributional fixed point equation used to identify the limits appearing in Theorem 1.3. Our study leads to many further questions, especially around descriptions of the limits and moment sequences appearing in Theorem 1.2 and 1.3, and so we discuss some of these in Section 3, where we also list open problems and conjectures. Section 4 contains the proof of Theorem 1.2, Section 5 has the proof of Theorem 1.3, and in Section 6 we derive some basic properties of the UL family.

1.3 Related Literature

The literature around Pólya urn models is too vast for a complete survey, but the the main results and modern techniques are well covered by Chauvin et al. (2011), Chauvin et al. (2015), Chen and Wei (2005), Chen and Kuba (2013), Flajolet et al. (2005), Janson (2004, 2006), Knape and Neininger (2014), Kuba and Mahmoud (2015a,b), Laruelle and Pagès (2013), Pouyanne (2008), and references therein. These papers cover many variations of the standard model, including random replacement rules and drawing multiple balls at a time. Techniques used to study limits include finding appropriate martingales, stochastic approximation, embedding the process into continuous time branching processes, deriving moments or moment generating functions using analytic or algebraic relations derived from the Markovian dynamics of the process, and the contraction method. All of these methods rely on a reasonably nice Markovian dynamics and in general, the model studied here is not Markov in its natural time scale. It is possible to make the model Markov by observing the process at the random times of immigration, but then the dynamics are complicated, and so it is challenging to apply the techniques mentioned above. On the other hand, our distributional fixed point approach is naturally suited to the model and leads to intriguing descriptions of the limiting behavior and to further avenues of study. We leave the question of what can be learned by studying this model with other methods to further work (see Section 3).

1.4 Connection to preferential attachment random graph

In preferential attachment random graph models, vertices are sequentially added and randomly connected to existing vertices such that connections to higher degree nodes are more likely. There are many variations of these popular models; a good reference is van der Hofstad (2016, Chapter 8).

Consider the following sequence $(G(n))_{n\geqslant 0}$ of preferential attachment random graphs. The initial state $G(0)$ is a "seed" graph with s vertices, where the degree or "weight" of vertex $1 \leqslant i \leqslant s$ is $d_i > 0$. We denote the weight of vertex i in $G(n)$ by $d_i(n)$ so note for $1 \leqslant i \leqslant s$, $d_i(0) = d_i$.

Let $\tau_1, \tau_2, \dots$ be i.i.d. distributed according to inter-arrival distribution π. Given the graph $G(n-1)$ having $s+n-1$ vertices, $G(n)$ is formed by adding a vertex labeled $s+n$ and sequentially attaching τ_n edges between it and the vertices of $G(n-1)$ according to the following rules. The first edge attaches to vertex k with probability

$$\frac{d_k(n-1)}{\sum_{i=1}^{s+n-1} d_i(n-1)}, \qquad 1 \leqslant k \leqslant n-1; \tag{1.9}$$

denote by K_1 the vertex which received that first edge. The weight of K_1 is updated immediately, so that the second edge attaches to vertex k with probability

$$\frac{d_k(n-1) + \mathrm{I}[k = K_1]}{1 + \sum_{i=1}^{s+n-1} d_i(n-1)}, \qquad 1 \leqslant k \leqslant n-1.$$

The procedure continues this way, edges attach with probability proportional to weights at that moment, and additional received edges add one to the weight of a vertex, until vertex n has τ_n outgoing edges. Lastly, we set $d_{s+n}(n) = 1$, and let $G(n)$ be the resulting graph. Note that multiple edges between vertices are possible.

This model is a randomized version of the "sequential" model of Berger et al. (2014); also the "N_ℓ" model of Peköz et al. (2014). For related models where the number of edges are random but the updating rule is not sequential (meaning each of the τ_n edges of vertex $s+n$ attach with probability (1.9)) see Deijfen et al. (2009) and a particular choice of parameters in the general model of Cooper and Frieze (2003).

Writing $c_i := \sum_{j=1}^{i} d_j$, the connection between the preferential attachment model above and our urn model is that for $1 \leqslant k < s$,

$$\mathscr{L}\left(\sum_{i=1}^{k} d_i(n)\right) = \mathcal{P}^{\pi}\left(\begin{smallmatrix} c_s - c_k \\ c_k \end{smallmatrix}; \sum_{i=1}^{n} \tau_i\right),$$

where the τ_i's on the right hand side drive the urn process. Thus for π having all positive integer moments finite, we have (in particular) $T_n/n := n^{-1}\sum_{i=1}^{n} \tau_i \to \mu$ almost surely and so Theorem 1.2 implies that for $k = 1, \dots, s$,

$$\mathscr{L}\left(\frac{\sum_{i=1}^{k} d_i(n)}{(\mu n)^{\mu/(\mu+1)}}\right) \to \mathscr{L}(BZ),$$

where, in accord with Theorem 1.3, $Z \sim \mathrm{UL}\big(c_s - 1, \big(\frac{\pi_{k-1}}{m_k}\big)_{k\geqslant 1}\big)$, $m_k = m_k(1, c_s - 1, \pi)$ are the limiting moments given in (1.5) of Theorem 1.2, and $B \sim \mathrm{Beta}(c_k, c_s - c_k - 1)$ is independent of Z.

For later vertices, if $k \geqslant s$, then for $n \geqslant k - s + 1$,

$$\mathscr{L}\Big(\sum_{i=1}^{k} d_i(n)\Big) = \mathcal{P}^{\pi}\Big(\tfrac{1}{c_s+(k-s)+T_{k-s+1}}; \sum_{i=k-s+2}^{n} \tau_i\Big), \tag{1.10}$$

where again the τ_i's on the right hand side drive the urn process. Given $\tau_1, \ldots, \tau_{k-s+1}$, it is still the case that $n^{-1}\sum_{i=k-s+2}^{n} \tau_i \to \mu$ and thus

$$\mathscr{L}\left(\frac{\sum_{i=1}^{k} d_i(n)}{(\mu n)^{\mu/(\mu+1)}} \middle| \big(\tau_1, \ldots, \tau_{k-s+1}\big)\right) \to \mathrm{UL}\big(c_s + (k-s) + T_{k-s+1}, \big(\tfrac{\pi_{k-1}}{m_k}\big)_{k\geqslant 1}\big),$$

where $m_k = m_k(1, c_s + (k-s) + T_{k-s+1}, \pi)$ is the limiting moment sequence (1.5) of Theorem 1.2. Thus the (unconditional) limiting cumulative degree counts are an appropriate mixture of the UL laws.

To our knowledge, these are the first results regarding the degree of fixed vertices in preferential attachment models with random initial degrees. The degree of a randomly chosen node is studied in Deijfen et al. (2009) and Cooper and Frieze (2003).

2 DISTRIBUTIONAL FIXED POINT EQUATION

To describe the distributional fixed point equation used to identify the limits appearing in Theorem 1.3, we first need a preliminary distributional transformation.

Definition 2.1. Let ψ be a probability distribution concentrated on the non-negative integers, and let X be a positive random variable such that $\mathbb{E}X^k < \infty$ for all k for which $\psi_k > 0$. A random variable $X^{(\psi)}$ is said to have the *ψ-power-bias distribution of X* if

$$\mathbb{E}f\big(X^{(\psi)}\big) = \sum_{k\,:\,\psi_k>0} \psi_k \frac{\mathbb{E}\{X^k f(X)\}}{\mathbb{E}X^k} \tag{2.1}$$

for all f for which the expectation on the right hand side exists.

If $\psi_1 = 1$, then the ψ-power-bias distribution is commonly known as the *size-bias distribution*; see for example Arratia et al. (2013) and Brown (2006). If $\psi_k = 1$ for some $k \geqslant 2$, then the ψ-power-bias distribution is sometimes referred to as the *k-power bias distribution*, denoted by $X^{(k)}$. We can realize $X^{(\psi)}$ by first sampling a random index K according to ψ, and conditional on $K = k$, we let $X^{(\psi)}$ have the k-power-bias distribution of X. This description implies that the ψ-power-bias transformation may be amenable to analysis in our setting since constructing constant k-power bias distributions is understood in Pólya urn models Peköz et al. (2013a,b, 2016), Ross (2013), and other discrete probability applications Barbour et al. (1992), Chen et al. (2011), Bartroff and Goldstein (2013).

To establish the distributional transformation for which $\mathrm{UL}\big(w; (a_k)_{k\geqslant 1}\big)$ is a fixed point, note first that if $Z \sim \mathrm{UL}\big(w; (a_k)_{k\geqslant 1}\big)$ and $\mu_k = \mathbb{E}Z^k$, then (1.3) yields, in particular,

$$\sum_{k\geqslant 1} a_k \mu_k = 1,$$

so that

$$\psi_k = a_k \mu_k, \qquad k \geqslant 1 \tag{2.2}$$

defines a probability distribution on the positive integers, which, by (1.4), is invariant to scaling of Z. The next result gives the UL family as fixed point of a distributional transformation; the connection between this transformation and the representation (1.2) was first made in Pakes and Navarro (2007).

Proposition 2.2. *The following holds.*

(*i*) *If* $X \sim \mathrm{UL}\big(w; (a_k)_{k\geqslant 1}\big)$ *and* ψ *is defined as in* (2.2)*, then*

$$\mathscr{L}(X) = \mathscr{L}\big(V_w X^{(\psi)}\big), \tag{2.3}$$

where $V_w \sim \mathrm{Beta}(w, 1)$ *is independent of* $X^{(\psi)}$.

(*ii*) *Let* $w > 0$*, and let* ψ *be a probability distribution on the positive integers. If* X *is a positive random variable such that* $\mathbb{E}X^k < \infty$ *whenever* $\psi_k > 0$ *and* (2.3) *holds, then* $A(x)$*, defined with respect to the sequence* $a_k = \psi_k/\mathbb{E}X^k$ *if* $\psi_k > 0$ *and* $a_k = 0$ *otherwise, has positive or infinite radius of convergence and* $X \sim \mathrm{UL}\big(w; (a_k)_{k\geqslant 1}\big)$.

Proof. To prove (*i*), assume $X \sim \mathrm{UL}(w; (a_n)_{n\geqslant 1})$. Using the formula for the density of products of independent random variables and denoting the density of $X^{(\psi)}$ by $u^{(\psi)}$, we obtain that $V_w X^{(\psi)}$ has density

$$\begin{aligned}
\int_x^\rho w\left(\frac{x}{t}\right)^{w-1} \frac{u^{(\psi)}(t)}{t} dt &= \int_x^\rho w\left(\frac{x}{t}\right)^{w-1} \frac{u(t)}{t} \sum_{n\geqslant 1} \frac{\psi_n}{\mu_n} t^n dt \\
&= \int_x^\rho w\left(\frac{x}{t}\right)^{w-1} \frac{u(t)}{t} \sum_{n\geqslant 1} a_n t^n dt = \int_x^\rho w\left(\frac{x}{t}\right)^{w-1} u(t) \frac{A(t)}{t} dt \\
&= x^{w-1} \int_x^\rho w \frac{A(t)}{t} \frac{u(t)}{t^{w-1}} dt = u(x).
\end{aligned}$$

To prove (*ii*), assume X is a positive random variable such that $\mathbb{E}X^k < \infty$ for all $k \geqslant 1$ with $\psi_k > 0$, and assume (2.3) holds. Since, by Jensen's inequality,

$$\sum_{k\geqslant 1} \frac{\psi_k}{\mathbb{E}X^k} x^k \leqslant \sum_{k\geqslant 1} \frac{1}{(\mathbb{E}X)^k} x^k < \infty$$

whenever $x < \mathbb{E}X$, the radius of convergence of $A(x)$ must be at least $\mathbb{E}X$, which is positive since X is positive. It follows from (2.3) that X has a density, and the representation (1.2) then follows from Pakes and Navarro (2007, Theorem 3.1). □

Remark 2.3. It is important to note that Proposition 2.2 does not answer the question whether, for given $w > 0$ and probability distribution ψ, there is an X satisfying (2.3). It merely says that, if such X exists, then it has to be from the family $\mathrm{UL}(w; (a_k)_{k\geqslant 1})$, where $(a_k)_{k\geqslant 1}$ can be expressed in terms of ψ and the moments of X. Note also that, for given w and ψ, there might a priori be more than one $(a_k)_{k\geqslant 1}$ satisfying $\psi_k = a_k \mathbb{E}X^k$; see the discussion in the next section.

Proposition 2.2 suggests that if a random variable W is such that $\mathscr{L}(W)$ is close in an appropriate sense to $\mathscr{L}(V_w W^{(\psi)})$, then $\mathscr{L}(W)$ is close to $\mathrm{UL}(w, (\psi_k/\mathbb{E}W^k)_{k\geqslant 1})$. We formalize this as a convergence statement in Lemma 5.1 in Section 5. We then apply this result to our urn models, where ψ has the immigration distribution π, but shifted by one, and the limiting moments are those given by Theorem 1.2. That the urn law and its transformation are close is achieved by coupling, in particular that power-biasing our urn models corresponds to adding extra white balls before starting the process (Lemma 5.4), and that multiplying by a beta corresponds to running a classical Pólya urn (Lemma 5.8); see Section 5 for details.

3 OPEN PROBLEMS

We discuss some of the many questions that are not answered by our study.

Question 3.1. Are solutions to the distributional fixed point equation (2.3) unique up to scaling?

This is the most pressing open problem, and a positive answer would have a large impact on our understanding of the relation between limits of our urn model and the family of distributions $\mathrm{UL}\big(w;(a_k)_{k\geqslant 1}\big)$. The main consequence of a positive answer would be the following "inversion" of Theorem 1.3.

Conjecture 3.2. *Fix w and $(a_k)_{k\geqslant 1}$, and let $Z \sim \mathrm{UL}\big(w;(a_k)_{k\geqslant 1}\big)$. Then, with $\pi_k = a_{k+1}\mathbb{E}Z^{k+1}$ for $k \geqslant 0$, (and possibly further conditions on $(a_k)_{k\geqslant 1}$), the sequence $X_n \sim \mathcal{P}^\pi\binom{1}{w};n)$ satisfies*

$$\mathscr{L}\left(\frac{X_n}{n^{\mu/(\mu+1)}}\right) \to \mathscr{L}(\theta Z),$$

where $\theta = m_1(1,w,\pi)/\mathbb{E}Z$ with $m_1(1,w,\pi)$ given by Theorem 1.2.

It is clear that, for any $(a_k)_{k\geqslant 1}$ with positive or infinite radius of convergence, we can define the probability distribution $\pi_k = a_{k+1}\mathbb{E}Z^{k+1}$, $k \geqslant 0$, and consider the limit of the corresponding urn model. But unless the solution to (2.3) is unique up to scaling, our method of proof does not guarantee that the corresponding urn limit is a scaling of the one given by this $(a_k)_{k\geqslant 1}$.

The following question recasts Question 3.1 differently; it must have a positive answer if Question 3.1 has a negative answer.

Question 3.3. Fix $w > 0$. Are there two sequences $(a_k)_{k\geqslant 1}$ and $(\tilde{a}_k)_{k\geqslant 1}$ such that $Z \sim \mathrm{UL}\big(w;(a_k)_{k\geqslant 1}\big)$ and $\tilde{Z} \sim \mathrm{UL}\big(w;(\tilde{a}_k)_{k\geqslant 1}\big)$ are not scaled versions of each other, but such that $a_k\mathbb{E}Z^k = \tilde{a}_k\mathbb{E}\tilde{Z}^k$ for all $k \geqslant 1$?

If Question 3.3 could be answered positively, then we would have a counter example to Conjecture 3.2 — both $(a_k)_{k\geqslant 1}$ and $(\tilde{a}_k)_{k\geqslant 1}$ would give rise to the same immigration distribution, but the corresponding urn model could converge to at most one of them.

One issue with Theorem 1.2 is that the limiting moments m_k are defined rather indirectly, and they are are difficult to calculate explicitly; the same comment applies to moments of the UL family.

Question 3.4. Are there a more explicit formulas for m_k in Theorem 1.2 in terms of w and π; or for the moments of $\mathrm{UL}\big(w;(a_k)_{k\geqslant 1}\big)$ in terms of w and $(a_k)_{k\geqslant 1}$?

There are a few examples where we can make explicit calculations and partially address this last question; see the end of this section.

A natural example that we have struggled to prove anything more specific about than the conclusion of Theorem 1.3, is for π a positive geometric variable. In this case the urn model can be described as follows: at each step, a Pólya urn step is performed and then a p-coin is tossed to determine if an additional black ball is added to the urn. So the process is Markovian, which could make more detailed analyses possible.

Question 3.5. What is a concrete description of the distributional limit of $\mathcal{P}^\pi\binom{1}{w};n)$ (properly scaled) when π is a positive geometric distribution (support starting at 1)?

Question 3.6. There are a large number of ways the model can be generalized: more colors, different replacement rules. What can be said in these cases?

Question 3.7. Can other methods, such as those described in Section 1.3, be applied to strengthen our results? For example, if appropriate martingales can be found, then the convergence can be strengthened to almost sure and in L_p for appropriate p.

We conclude this section with three examples.

3.1 Explicit choices of π

The relationship between the sequences $(a_k)_{k\geqslant 1}$ and π appearing in Theorem 1.3 is rather implicit, and so in this section, we work out some examples where explicit calculations are possible.

Example 3.8 (Deterministic π)**.** If $\pi_k = 1$ for some $k \geqslant 1$, then the scaled limit of $\mathcal{P}^\pi(\frac{1}{w}; n)$ has density proportional to

$$x^{w-1}\exp\{-wx^{k+1}/((k+1)m_{k+1})\}dx,$$

which is the same as an appropriately scaled, standard gamma variable with parameter $w/(k+1)$, raised to the power $1/(k+1)$. For $k = 1$, the urn model is a time homogeneous triangular urn and the limit can be read from Janson (2006). The general case is studied in detail in Peköz et al. (2016), where rates of convergence to the limit are also provided. The limiting moments can be made explicit as well as the constant m_{k+1}.

Example 3.9 (Bernoulli inter-arrival distribution)**.** We study $\mathcal{P}^\pi(\frac{1}{w}; n)$ where $\pi_0 = 1 - \pi_1 \neq 1$. Note that for this choice of π, at each step a Pólya urn step is performed and then a geometric with parameter π_1 (support started at 0) distributed number of black balls are added to the urn. In the spirit of Conjecture 3.2, we start with a positive integer w and positive numbers a_1 and a_2, and then use these to determine π.

First, define the function U for $a > 0$, $z > 0$ and $b \in \mathbb{R}$ by

$$U(a, b, z) = \frac{1}{\Gamma(a)}\int_0^\infty e^{-zt}t^{a-1}(1+t)^{b-a-1}dt.$$

This function is known as Kummer U (also called the confluent hypergeometric function of the second kind; see Abramowitz and Stegun (1964, 13.2.5)). Second, we calculate the normalising constant c in (1.2); one can show that

$$\int_0^\infty x^{w-1}\exp\left\{-w\left(a_1x + a_2x^2/2\right)\right\}dx = \frac{\Gamma(w)U\left(\frac{w}{2}, \frac{1}{2}, \frac{a_1^2 w}{2a_2}\right)}{(2a_2w)^{w/2}},$$

so that

$$u(x) = \frac{(2a_2w)^{w/2}}{\Gamma(w)U\left(\frac{w}{2}, \frac{1}{2}, \frac{a_1^2w}{2a_2}\right)}x^{w-1}\exp\left\{-w\left(a_1x + a_2x^2/2\right)\right\} \qquad \text{for } x > 0. \tag{3.1}$$

Third, we calculate the relevant moments and obtain

$$\mathbb{E}Z = \frac{wa_1}{2a_2}\cdot\frac{U\left(\frac{w}{2}+1, \frac{3}{2}, \frac{wa_1^2}{2a_2}\right)}{U\left(\frac{w}{2}, \frac{1}{2}, \frac{wa_1^2}{2a_2}\right)}, \qquad \mathbb{E}Z^2 = \frac{1+w}{2a_2}\cdot\frac{U\left(\frac{w}{2}+1, \frac{1}{2}, \frac{wa_1^2}{2a_2}\right)}{U\left(\frac{w}{2}, \frac{1}{2}, \frac{wa_1^2}{2a_2}\right)}, \tag{3.2}$$

Putting this together we obtain

$$\pi_0 = 1 - \pi_1 = \frac{wa_1^2}{2a_2}\cdot\frac{U\left(\frac{w}{2}+1, \frac{3}{2}, \frac{wa_1^2}{2a_2}\right)}{U\left(\frac{w}{2}, \frac{1}{2}, \frac{wa_1^2}{2a_2}\right)} = \frac{U\left(\frac{w+1}{2}, \frac{1}{2}, \frac{wa_1^2}{2a_2}\right)}{U\left(\frac{w+1}{2}, \frac{3}{2}, \frac{wa_1^2}{2a_2}\right)}. \tag{3.3}$$

The second equality of (3.3) follows by applying the identity $U(a,b,z) = z^{1-b}U(1+a-b, 2-b, z)$ (see Abramowitz and Stegun (1964, 13.1.29)) to both the numerator and the denominator of the middle expression of (3.3) with $z = wa_1^2/(2a_2)$ and $a = w/2+1$ and $b = 3/2$, respectively, $a = w/2$ and $b = 1/2$. As a check on (3.2), we can see directly that

$$\pi_0 + \pi_1 = a_1 \mathbb{E}Z + a_2 \mathbb{E}Z^2 = 1$$

from Abramowitz and Stegun (1964, 13.4.18) with $a = w/2+1$, $b = 1/2$, and $z = (wa_1^2)/(2a_2)$.

Theorems 1.2 and 1.3 give moment and distributional convergence results for the urn model with inter-arrival distribution (π_0, π_1). Furthermore, it is possible to show directly that for fixed w, the function on positive pairs of numbers

$$(a_1, a_2) \mapsto \pi_0$$

is surjective on $(0,1)$; hence, every inter-arrival distribution concentrated on $\{0,1\}$ can be generated by starting with an appropriate a_1 and a_2. Finally, we note that Conjecture 3.2 is verified in this case since if

$$\tilde{a}_1 \mathbb{E}\tilde{Z} = a_1 \mathbb{E}Z, \qquad \tilde{a}_2 \mathbb{E}\tilde{Z}^2 = a_2 \mathbb{E}Z^2,$$

then (3.2) implies that $\tilde{a}_1^2/\tilde{a}_2 = a_1^2/a_2$, which implies $\tilde{a}_1^2/a_1^2 = \tilde{a}_2/a_2 =: \theta^2$ (which is the same as the conjecture, noting (1.4)).

Conjectural Example 3.10 (Power law inter-arrival distribution)**.** Let α and β be positive numbers, and set $a = (\beta\alpha^{-1}, \beta\alpha^{-2}, \beta\alpha^{-3}, \ldots)$. Then for $0 < x < \alpha$,

$$\sum_{k \geqslant 1} a_k x^k/k = -\beta \log(1 - x/\alpha).$$

Thus, if $Z \sim \mathrm{UL}(w; (a_k)_{k\geqslant 1})$ has density given by (1.2) with $v = w$, we find $\mathscr{L}(\alpha^{-1}Z) = \mathrm{Beta}(w, w\beta + 1)$ and that

$$\mathbb{E}Z^j = \alpha^j \frac{\Gamma(w(\beta+1)+1)\Gamma(w+j)}{\Gamma(w)\Gamma(w(\beta+1)+j+1)}.$$

Following the blueprint of Conjecture 3.2, define for $j = 0, 1, \ldots,$

$$\pi_j = a_{j+1}\mathbb{E}Z^{j+1} = \beta \frac{\Gamma(w(\beta+1)+1)\Gamma(w+j+1)}{\Gamma(w)\Gamma(w(\beta+1)+j+2)}. \tag{3.4}$$

These calculations suggest that if α, β and w are positive numbers, and π has distribution given by (3.4), then there is a constant $\theta > 0$ such that, for $X_n \sim \mathcal{P}^{\pi}(\frac{1}{w}; n)$ and as $n \to \infty$,

$$\mathscr{L}\big(n^{-\frac{\mu}{\mu+1}} X_n\big) \to \mathscr{L}(\theta Z),$$

where $\mathscr{L}(\alpha^{-1}Z) = \mathrm{Beta}(w, w\beta+1)$ and μ denotes the mean of π given by

$$\mu = \begin{cases} (w+1)\,(\beta w - 1)^{-1} & \text{if } w\beta > 1, \\ \infty & \text{if } w\beta \leqslant 1. \end{cases}$$

The previous statement is conjectural for two reasons. In the case that $w\beta > 1$, π has finite mean but not all moments finite, so even convergence in this case is not covered by Theorem 1.3. For $w\beta < 1$, Theorem 1.3 applies and says that $\mathscr{L}\big(n^{-\frac{\mu}{\mu+1}} X_n\big)$ converges in distribution to $\mathrm{UL}(w, (\pi_k/m_k(1,w,\pi))_{k\geqslant 1})$, but without a result like Conjecture 3.2, we cannot conclude that $\pi_k/m_k(1,w,\pi) = \theta^k\beta\alpha^{-k}$ for some $\theta > 0$.

Also note that $\beta \to 0$ roughly corresponds to $\pi_k = 0$ for all k, and so $\tau = \infty$, which should behave as a classical Pólya urn, and indeed the conjectured limit tends to the anticipated $\mathrm{Beta}(w, 1)$. In general the π distribution in this case is heavy-tailed and the conjecture suggests that extra balls are not added with enough frequency to get too far away from the classical Pólya urn.

4 PROOF OF THEOREM 1.2

We will show the following result, which is the analogue of Theorem 1.2, but for factorial moments. In what follows, we interpret products $\prod_{j=a}^{b}$ as 1 whenever $b < a$.

Proposition 4.1. *Let b and w be positive integers, let π be a probability distribution on the non-negative integers with mean $0 < \mu \leqslant \infty$, and let $\tau \sim \pi$. Let $X_n \sim \mathcal{P}^\pi(\binom{b}{w}; n)$, and set*

$$D_{k,n} = \prod_{j=0}^{k-1} (X_n + j)\,, \qquad k \geqslant 1,\ n \geqslant 0.$$

If either

(a) $\mathbb{E}\tau^p < \infty$ *for some* $p > 1$, *and* $1 \leqslant k < (\frac{p}{2} - 1)(\mu + 1) - 1$, *or*

(b) *there is* $\varepsilon > 0$ *such that* $\mathbb{P}[\tau > n] \geqslant Cn^{-(1-\varepsilon)}$ *for n large enough, and* $k \geqslant 1$,

then there is a positive constant $m_k(b, w, \pi)$ such that

$$\mathbb{E}\left\{\frac{D_{k,n}}{n^{k\mu/(\mu+1)}}\right\} \to m_k(b, w, \pi) \qquad \text{as } n \to \infty. \tag{4.1}$$

To prove the proposition we need some lemmas. We first establish a moment formula for $D_{k,n}$, conditional on the immigration times $T_1, T_2, \ldots$.

Lemma 4.2. *Let $D_{k,n}$ be as in Proposition 4.1, let $T = (T_1, T_2, \ldots)$ be the sequence of immigration times of the process, and for $n \geqslant 0$, let $N_n = \#\{i \geqslant 1 : T_i \leqslant n\}$, the number of immigrations up to and including draw n. Then, for any $k \geqslant 1$ and $n \geqslant 1$,*

$$\mathbb{E}(D_{k,n}|T) = \frac{\Gamma(w+k)}{\Gamma(w)} \prod_{j=0}^{n-1} \frac{b+w+k+j+N_j}{b+w+j+N_j} \tag{4.2}$$

$$= \frac{\Gamma(w+k)\Gamma(b+w)}{\Gamma(w)\Gamma(b+w+k)} \frac{\Gamma(b+w+N_{n-1}+n+k)}{\Gamma(b+w+N_{n-1}+n)} \prod_{j=1}^{N_{n-1}} \frac{b+w+j-1+T_j}{b+w+j-1+T_j+k}. \tag{4.3}$$

Proof. Let $X_n \sim \mathcal{P}^\pi(\binom{b}{w}; n)$. To shorten the formulas, let $c = b+w$. Since the total number of balls in the urn after draw $n-1$ is $c + N_{n-1} + n - 1$, we have

$$\mathbb{P}\big[X_n = X_{n-1} + 1 \,\big|\, T, X_{n-1}\big] = \frac{X_{n-1}}{c + N_{n-1} + n - 1},$$

and we easily find

$$\mathbb{E}(D_{k,n}\,|\,T, X_{n-1}) = D_{k,n-1}\frac{c+k+N_{n-1}+n-1}{c+N_{n-1}+n-1}.$$

Iterating yields

$$\mathbb{E}(D_{k,n}|T) = \mathbb{E}(D_{k,0}|T) \prod_{j=0}^{n-1} \frac{c+k+j+N_j}{c+j+N_j},$$

which is easily seen to be (4.2). Now, set $T_0 = 0$ and note that for $i \geqslant 1$, if $T_{i-1} < T_i$, then $N_{T_{i-1}} = \cdots = N_{T_i - 1} = i-1$, so we can rewrite this last expression as

$$\begin{aligned}
\frac{\mathbb{E}(D_{k,n}|T)}{\mathbb{E}(D_{k,0}|T)} &= \left(\prod_{i=1}^{N_{n-1}} \prod_{j=T_{i-1}}^{T_i - 1} \frac{c+k+j+i-1}{c+j+i-1} \right) \prod_{j=T_{N_{n-1}}}^{n-1} \frac{c+k+j+N_{n-1}}{c+j+N_{n-1}} \\
&= \left(\prod_{i=1}^{N_{n-1}} \frac{\Gamma(c+k+i-1+T_i)\Gamma(c+i-1+T_{i-1})}{\Gamma(c+k+i-1+T_{i-1})\Gamma(c+i-1+T_i)} \right) \\
&\qquad \times \frac{\Gamma(c+k+N_{n-1}+n)\Gamma(c+N_{n-1}+T_{N_{n-1}})}{\Gamma(c+k+N_{n-1}+T_{N_{n-1}})\Gamma(c+N_{n-1}+n)} \\
&= \left(\prod_{i=1}^{N_{n-1}} \frac{\Gamma(c+k+i-1+T_i)}{\Gamma(c+i-1+T_i)} \right) \left(\prod_{i=0}^{N_{n-1}-1} \frac{\Gamma(c+i+T_i)}{\Gamma(c+k+i+T_i)} \right) \\
&\qquad \times \frac{\Gamma(c+k+N_{n-1}+n)\Gamma(c+N_{n-1}+T_{N_{n-1}})}{\Gamma(c+k+N_{n-1}+T_{N_{n-1}})\Gamma(c+N_{n-1}+n)} \\
&= \left(\prod_{i=1}^{N_{n-1}} \frac{\Gamma(c+k+i-1+T_i)\Gamma(c+i+T_i)}{\Gamma(c+k+i+T_i)\Gamma(c+i-1+T_i)} \right) \\
&\qquad \times \frac{\Gamma(c+k+N_{n-1}+n)\Gamma(c)}{\Gamma(c+k)\Gamma(c+N_{n-1}+n)},
\end{aligned}$$

which, using that $x\Gamma(x) = \Gamma(x+1)$, easily simplifies to (4.3). □

We use Lemma 4.2 to establish the almost sure behavior of $\mathbb{E}(D_{k,n}|T)$.

Lemma 4.3. *Let π be a probability distribution on the non-negative integers with mean $0 < \mu \leqslant \infty$, and let $\tau, \tau_1, \tau_2, \ldots$ be a sequence of independent and identically distributed random variables with distribution π. For $i \geqslant 1$, let $T_i = \sum_{j=1}^{i} \tau_i$, and for $n \geqslant 0$, let $N_n = \#\{i \geqslant 1 : T_i \leqslant n\}$. If either*

(a) there is $\varepsilon > 0$ such that $\mathbb{E}\tau^{1+\varepsilon} < \infty$, or

(b) there is $\varepsilon > 0$ such that $\mathbb{P}[\tau > n] \geqslant Cn^{-(1-\varepsilon)}$ for n large enough,

then, for any $\alpha > 0$ and $\beta > 0$, there exists a (possibly random) positive number $\chi(\alpha, \beta, \pi)$ such that, almost surely,

$$n^{-\frac{\alpha\mu}{1+\mu}} \prod_{i=1}^{n} \Big(1 + \frac{\alpha}{\beta + i + N_i}\Big) \to \chi(\alpha, \beta, \pi), \qquad \text{as } n \to \infty. \tag{4.4}$$

Proof. **Case *(a)*.** Taking logarithm in (4.4), it is enough to show that

$$\sum_{j=1}^{n} \log\Big(1 + \frac{\alpha}{\beta + j + N_j}\Big) - \frac{\alpha\mu}{1+\mu} \log n$$

converges almost surely to a (possibly random) real number. Since both

$$\log(n) - \sum_{j=1}^{n} \frac{1}{j} \qquad \text{and} \qquad \sum_{j=1}^{n} \log(1 + x_j) - \sum_{j=1}^{n} x_j \tag{4.5}$$

converge, provided

$$\sum_{j=1}^{\infty} x_j^2 < \infty, \tag{4.6}$$

it is enough to consider convergence of

$$\begin{aligned} &\sum_{j=1}^{n} \Big(\frac{1}{\beta + j + N_j} - \frac{\mu}{(1+\mu)j} \Big) \\ &\qquad = \sum_{j=1}^{n} \Big(\frac{1}{\beta + j + N_j} - \frac{1}{j + N_j} \Big) + \sum_{j=1}^{n} \Big(\frac{1}{j + N_j} - \frac{\mu}{(1+\mu)j} \Big). \end{aligned} \tag{4.7}$$

Now, to prove (4.6) for $x_j = 1/(\beta + j + N_j)$, which justifies the second approximation in (4.5) and also convergence of the first sum on the right hand side of (4.7), we observe that, almost surely,

$$\sum_{j=1}^{\infty} \Big(\frac{1}{\beta + j + N_j} \Big)^2 \leqslant \sum_{j=1}^{\infty} \frac{1}{j^2} < \infty. \tag{4.8}$$

In order to prove convergence of the second sum on the right hand side of (4.7), we need a refined estimate for the renewal law of large numbers. Assume without loss of generality that $\varepsilon < 1$. Let $e_n^+ = n\big(\mu^{-1} + n^{-\varepsilon/2}\big)$ and $E_n^+ = \lceil e_n^+ \rceil$, and observe that

$$\begin{aligned} \Big\{ \frac{N_n}{n} - \frac{1}{\mu} \geqslant n^{-\varepsilon/2} \Big\} &= \big\{ N_n \geqslant e_n^+ \big\} \\ &= \big\{ T_{E_n^+} \leqslant n \big\} \\ &= \Big\{ \frac{T_{E_n^+} - \mu E_n^+}{E_n^+} \leqslant \frac{n - \mu E_n^+}{E_n^+} \Big\}. \end{aligned}$$

Likewise, with $e_n^- = n\big(\mu^{-1} - n^{-\varepsilon/2}\big)$, $E_n^- = \lfloor e_n^- \rfloor + 1$, and n large enough to ensure $E_n^- > 0$,

$$\begin{aligned} \Big\{ \frac{N_n}{n} - \frac{1}{\mu} \leqslant -n^{-\varepsilon/2} \Big\} &= \big\{ N_n \leqslant e_n^- \big\} \\ &= \big\{ T_{E_n^-} > n \big\} \\ &= \Big\{ \frac{T_{E_n^-} - \mu E_n^-}{E_n^-} > \frac{n - \mu E_n^-}{E_n^-} \Big\}. \end{aligned}$$

From this and the fact that $|n - \mu E_n^{\pm}|/E_n^{\pm} = \Theta(n^{-\varepsilon/2})$, it is not difficult to see that there is a constant $C > 0$ such that

$$\begin{aligned} \limsup_{n \geqslant 1} \Big\{ \Big| \frac{N_n}{n} - \frac{1}{\mu} \Big| \geqslant n^{-\varepsilon/2} \Big\} &\subset \limsup_{n \geqslant 1} \Big\{ \Big| \frac{T_n}{n} - \mu \Big| \geqslant C n^{-\varepsilon/2} \Big\} \\ &= \limsup_{n \geqslant 1} \Big\{ \frac{1}{n^{1-\varepsilon/2}} \Big| \sum_{i=1}^{n} (\tau_i - \mu) \Big| \geqslant C \Big\}. \end{aligned} \tag{4.9}$$

It follows from Petrov (1975, Theorem 17, p. 274, with $a_n = n^{1-\varepsilon/2}$) and the fact that $\mathbb{E}\tau^{1+\varepsilon} < \infty$ that the last event in (4.9) has probability zero (alternatively use the Marcinkiewicz-Zygmund strong law of large numbers). Therefore,

$$\limsup_{n\to\infty} n^{\varepsilon/2}\Big|\frac{N_n}{n} - \frac{1}{\mu}\Big| < \infty \tag{4.10}$$

almost surely. Since

$$\begin{aligned}\sum_{j=1}^{n}\Big|\frac{1}{j+N_j} - \frac{\mu}{(1+\mu)j}\Big| &= \sum_{j=1}^{n}\Big|\frac{1}{(1+N_j/j)j} - \frac{1}{(1+1/\mu)j}\Big| \\ &= \sum_{j=1}^{n}\frac{|N_j/j - 1/\mu|}{(1+N_j/j)(1+1/\mu)j} \\ &\leqslant \sum_{j=1}^{n}\frac{j^{\varepsilon/2}|N_j/j - 1/\mu|}{j^{1+\varepsilon/2}}.\end{aligned}$$

we conclude that, using (4.10) and the fact that $\sum_{j\geqslant 1} j^{-(1+\varepsilon/2)} < \infty$, the last sum converges almost surely as $n \to \infty$.

Case (*b*). Following the proof of the μ finite case (interpreting $\infty/(1+\infty)$ as 1) up to and including (4.8), it is sufficient to establish, as in (4.10), that $\limsup_{n\to\infty} n^{\varepsilon'}\frac{N_n}{n} < \infty$ almost surely for some $\varepsilon' > 0$. Observe that

$$\begin{aligned}\mathbb{P}[T_n \leqslant n^{1+\varepsilon}] &\leqslant \mathbb{P}[\max\{\tau_1,\ldots,\tau_n\} \leqslant n^{1+\varepsilon}] = \big(\mathbb{P}[\tau_1 \leqslant n^{1+\varepsilon}]\big)^n = \big(1 - \mathbb{P}[\tau_1 > n^{1+\varepsilon}]\big)^n \\ &\leqslant \Big(1 - \frac{C}{n^{(1-\varepsilon)(1+\varepsilon)}}\Big)^n \\ &= \Big(1 - \frac{C}{n^{1-\varepsilon^2}}\Big)^n \leqslant \exp\big(-Cn^{\varepsilon^2}\big).\end{aligned}$$

By Borel-Cantelli,

$$\mathbb{P}\Big[\limsup_{n\to\infty}\Big\{\frac{1}{n^{1+\varepsilon/2}}T_n \leqslant n^{\varepsilon/2}\Big\}\Big] = 0,$$

so that $\frac{1}{n^{1+\varepsilon/2}}T_n \to \infty$ almost surely. By the usual relation between T_n and N_n, this implies that, for $\varepsilon' := 1 - \frac{1}{1+\varepsilon/2} > 0$,

$$\limsup_{n\to\infty} n^{\varepsilon'}\frac{N_n}{n} < \infty$$

almost surely. □

The next result provides moment bounds for applying dominated convergence to strengthen the convergence of Lemma 4.3.

Lemma 4.4. *Under the assumptions of Lemma 4.3, if either*

(*a*) $\mathbb{E}\tau^p < \infty$ *for some* $p > 1$, *and* $1 \leqslant k < (\frac{p}{2} - 1)(\mu + 1)$, *or*

(*b*) $\mathbb{E}\tau = \infty$ *and* $k \geqslant 1$,

then, with $D_{k,n}$ *be as in Proposition 4.1,*

$$\limsup_{n\to\infty}\frac{\mathbb{E}D_{k,n}}{n^{k\mu/(1+\mu)}} < \infty.$$

Proof. **Case (a).** Using representation given by (4.3), and the fact that $\Gamma(x+k) \leqslant (x+k)^k\Gamma(x)$, we conclude that there is a constant $C = C(b,w,k)$ such that

$$\mathbb{E}(D_{k,n}|T) \leqslant C(c+k+N_{n-1}+n)^k \prod_{j=1}^{N_{n-1}} \Big(1 - \frac{k}{c+j-1+T_j+k}\Big), \tag{4.11}$$

where we write $c = b+w$ to shorten formulas. Let A_n be the event that $|N_{n-1}-(n-1)/\mu| \leqslant u_n := (n-1)/(\mu+1)-1$. We have

$$\mathbb{E}D_{k,n} = \mathbb{E}(D_{k,n}\mathrm{I}[A_n]) + \mathbb{E}(D_{k,n}\mathrm{I}[A_n^c]), \tag{4.12}$$

and we show that both terms on the right hand side of (4.12) are $\mathrm{O}(n^{k\mu/(\mu+1)})$. Below it is important to notice that A_n is in the sigma-algebra generated by T. Now, for the first term of (4.12), use the expression (4.11) to find that, under the event A_n,

$$\mathbb{E}(D_{k,n}|T)\mathrm{I}[A_n] \leqslant C(c+k+n/\mu+u_n+n)^k \prod_{j=1}^{\varphi_n} \left(1 - \frac{k}{c+j-1+T_j+k}\right),$$

where $\varphi_n = \lfloor \frac{n-1}{\mu} - u_n \rfloor$, and note that $1 \leqslant \varphi_n = \Theta(n)$ by our definition of u_n. Since

$$(c+k+n/\mu+u_n+n)^k = \mathrm{O}(n^k),$$

it is sufficient to show that

$$\mathbb{E}\prod_{j=1}^{\varphi_n} \left(1 - \frac{k}{c+j-1+T_j+k}\right) = \mathrm{O}\big(n^{-k/(1+\mu)}\big).$$

Let $1/2 < \alpha < 1$ and set $U_\alpha := \sup\{j \geqslant 1 : T_j > j\mu + j^\alpha\}$ to be the last time that the centered random walk $(T_j - j\mu)_{j\geqslant 0}$ is larger than j^α; note that U_α is almost surely finite by the law of the iterated logarithm. Defining the empty product to be one, we have

$$\begin{aligned}
&\mathbb{E}\prod_{j=1}^{\varphi_n} \left(1 - \frac{k}{c+j-1+T_j+k}\right) \\
&\quad \leqslant \mathbb{E}\prod_{j=U_\alpha}^{\varphi_n} \left(1 - \frac{k}{c+j-1+T_j+k}\right) \leqslant \mathbb{E}\prod_{j=U_\alpha}^{\varphi_n} \left(1 - \frac{k}{c+j-1+j\mu+j^\alpha+k}\right) \\
&\quad \leqslant \prod_{j=1}^{\varphi_n} \left(1 - \frac{k}{c+j-1+j\mu+j^\alpha+k}\right) \mathbb{E}\prod_{j=1}^{U_\alpha} \left(1 + \frac{k}{j+j\mu+j^\alpha}\right).
\end{aligned} \tag{4.13}$$

We show the first product of (4.13) is $\mathrm{O}(n^{-k/(1+\mu)})$, and the second is bounded. Taking logarithm in the first product, we claim

$$\sum_{j=1}^{\varphi_n} \log\left(1 - \frac{k}{c+j-1+j\mu+j^\alpha+k}\right) + \frac{k}{1+\mu}\log(n)$$

converges as $n \to \infty$. Since $\log(\varphi_n) - \log(n)$ converges, we can replace $\log(n)$ with $\log(\varphi_n)$. Now, similar to the proof of Lemma 4.3, since

$$\sum_{j=1}^{n} \log(1+x_j) - \sum_{j=1}^{n} x_j$$

converges provided $\sum_{j=1}^{\infty} x_j^2 < \infty$, it is enough to consider the convergence of

$$\begin{aligned}
-\sum_{j=1}^{\varphi_n} & \frac{1}{c+j-1+j\mu+j^\alpha+k} + \frac{1}{1+\mu}\log(\varphi_n) \\
&= \sum_{j=1}^{\varphi_n} \left(\frac{1}{j(1+\mu)} - \frac{1}{c+j(1+\mu)+j^\alpha+k-1} \right) + \mathrm{O}(1) \\
&= \sum_{j=1}^{\varphi_n} \frac{c+j^\alpha+k-1}{(1+\mu)j(c+j(1+\mu)+j^\alpha+k-1)} + \mathrm{O}(1) \\
&\leqslant C\sum_{j=1}^{\varphi_n} j^{-2+\alpha} + \mathrm{O}(1);
\end{aligned}$$

here C is some constant and the sum is convergent since $\alpha < 1$. For the second product of (4.13), easy variations of the arguments above (or in the proof of Lemma 4.3) show that there is a constant C (depending on k and μ) such that for any $t \geqslant 1$,

$$\prod_{j=1}^{t}\left(1+\frac{k}{j+j\mu+j^\alpha}\right) \leqslant \prod_{j=1}^{t}\left(1+\frac{k}{j+j\mu}\right) \leqslant Ct^{k/(\mu+1)}.$$

Substituting $t = U_\alpha$, it is enough to show that $\mathbb{E}U_\alpha^{k/(\mu+1)} < \infty$. We choose $\alpha < 1$ close enough to one to ensure $p > \left(\frac{k}{\mu+1}+1\right)/\left(\alpha-1/2\right)$ and find

$$\begin{aligned}
\mathbb{P}[U_\alpha \geqslant x] &= \mathbb{P}\left[\cup_{j\geqslant x}\{T_j > j\mu + j^\alpha\}\right] \\
&\leqslant \sum_{j\geqslant x} \mathbb{P}\left[|T_j - j\mu| > j^\alpha\right] \\
&\leqslant \sum_{j\geqslant x} \frac{\mathbb{E}|T_j - j\mu|^p}{j^{p\alpha}} \qquad (4.14) \\
&\leqslant C_p \mathbb{E}|\tau-\mu|^p \sum_{j\geqslant x} j^{-p(\alpha-1/2)},
\end{aligned}$$

where the last inequality follows from Lemma 4.5 below. Using (4.14), we obtain

$$\begin{aligned}
\mathbb{E}U_\alpha^{k/(\mu+1)} &= \frac{k}{\mu+1}\int_0^\infty x^{k/(\mu+1)-1}\mathbb{P}[U_\alpha > x]dx \\
&= \frac{k}{\mu+1}\int_0^\infty x^{k/(\mu+1)-1}\mathbb{P}[U_\alpha \geqslant \lfloor x\rfloor + 1]dx \\
&\leqslant C_p\frac{k\mathbb{E}|\tau-\mu|^p}{\mu+1}\int_0^\infty x^{k/(\mu+1)-1}\sum_{k\geqslant\lfloor x\rfloor+1} k^{-p(\alpha-1/2)}dx \\
&\leqslant \frac{kC_p\mathbb{E}|\tau-\mu|^p}{(\mu+1)(p(\alpha-1/2)-1)}\int_0^\infty x^{k/(\mu+1)-1}\lfloor x\rfloor^{-p(\alpha-1/2)+1}dx \\
&\leqslant \frac{kC_p\mathbb{E}|\tau-\mu|^p}{(\mu+1)(p(\alpha-1/2)-1)}\int_0^\infty x^{k/(\mu+1)-p(\alpha-1/2)}dx < \infty;
\end{aligned}$$

the finiteness is by the assumption that $p > \left(\frac{k}{\mu+1}+1\right)/\left(\alpha-1/2\right)$.

For the second term of (4.12), first note that the term (4.2) is decreasing in the N_i, so the conditional expectation (without the indicator) is almost surely bounded

$$\mathbb{E}(D_{k,n}|T) \leqslant \frac{\Gamma(w+k)}{\Gamma(w)}\prod_{j=0}^{n-1}\frac{c+k+j}{c+j} \leqslant \frac{\Gamma(w+k)}{\Gamma(w)}\cdot\frac{\Gamma(c)}{\Gamma(c+k)}\cdot\frac{\Gamma(c+k+n)}{\Gamma(c+n)} = \mathrm{O}(n^k). \qquad (4.15)$$

Now noting that A_n^c is in the sigma-algebra generated by T, it is enough to show that $\mathbb{P}[A_n^c] = \mathrm{O}\big(n^{-k/(\mu+1)}\big)$. Denoting $\omega_n := \lceil (n-1)/\mu + u_n \rceil$ and using the moment bound of Lemma 4.5 below, we have

$$\begin{aligned}
\mathbb{P}[A_n^c] &= \mathbb{P}\big[|N_{n-1} - (n-1)/\mu| > u_n\big] \\
&\leqslant \mathbb{P}[T_{\omega_n} < n-1] + \mathbb{P}[T_{\varphi_n} \geqslant n-1] \\
&\leqslant \mathbb{P}[T_{\omega_n} - \mu\omega_n < n - 1 - \mu\omega_n] + \mathbb{P}[T_{\varphi_n} - \mu\varphi_n \geqslant n - 1 - \mu\varphi_n] \\
&\leqslant C_{2k/(\mu+1)} \mathbb{E}|\tau - \mu|^{2k/(\mu+1)} \left(\frac{\omega_n^{k/(\mu+1)}}{(\mu\omega_n - n - 1)^{2k/(\mu+1)}} + \frac{\varphi_n^{k/(\mu+1)}}{(n - 1 - \mu\varphi_n)^{2k/(\mu+1)}} \right).
\end{aligned}$$

But since $\varphi_n = (n-1)/\mu - u_n - \varepsilon_1$ and $\omega_n = (n-1)/\mu + u_n + \varepsilon_2$ for some $\varepsilon_1, \varepsilon_2 \in [0,1)$, the last expression is $\mathrm{O}(n^{-k/(\mu+1)})$, as desired.

Case (*b*). Since $X_n \leqslant n + w$, then $D_{k,n} \leqslant (n + w + k)^k$, and so $\mathbb{E}D_{k,n} = \mathrm{O}(n^k)$, as required. □

We now give the proof of Proposition 4.1 and then note that Theorem 1.2 easily follows from that result.

Proof of Proposition 4.1. **Case (*a*).** Lemma 4.3(a) applied to (4.2) of Lemma 4.2 implies that $n^{-k\mu/(\mu+1)}\mathbb{E}[D_{k,n}|T]$ converges almost surely to a positive random variable $\chi(k, b+w, \pi)$. Using Lemma 4.4 and dominated convergence (see, for example, Durrett (2010, Exercise 3.2.5)), it follows that $n^{-k\mu/(\mu+1)}\mathbb{E}D_{k,n} \to \mathbb{E}\chi(k, b+w, \pi)$.

Case (*b*). Analogous to Case (*a*). □

Proof of Theorem 1.2. In both cases (*a*) and (*b*), Lemma 4.3 implies that X_n converges to infinity almost surely. Hence, regular and factorial moments are asymptotically equivalent, and Theorem 1.2 follows directly from Proposition 4.1 with the same constants $m_k(b, w, \pi)$. □

The following lemma is given in Petrov (1975, 16, Page 60) where it is attributed to Dharmadhikari and Jogdeo (1969).

Lemma 4.5. *Let $Y_1, \ldots, Y_n$ be independent random variables such that for $i = 1, \ldots, n$, $\mathbb{E}Y_i = 0$ and $\mathbb{E}|Y_1|^p < \infty$, and let $S_n = \sum_{i=1}^n Y_i$. Then*

$$\mathbb{E}|S_n|^p \leqslant C_p n^{p/2-1} \sum_{i=1}^n \mathbb{E}|Y_i|^p,$$

where

$$C_p = \frac{1}{2} p(p-1) \max(1, 2^{p-3}) \left(1 + \frac{2}{p} K_{2m}^{(p-2)/2m}\right),$$

and the integer m satisfies $2m \leqslant p < 2m+2$, and

$$K_{2m} = \sum_{r=1}^m \frac{r^{2m-1}}{(r-1)!}.$$

We prove the convergence first for $b = 1$, and then the general case follows easily from an auxiliary Pólya urn argument.

Recall that $W_n = n^{-\mu/(\mu+1)} X_n$ and we want to derive the distributional limit of the sequence W_n. The method of proof is to show tightness of the sequence $\mathscr{L}(W_n)$, and then use the characterizing properties of the limit given by Proposition 2.2 to prove convergence. To simplify notation, for a probability distribution $\pi = (\pi_k)_{k\geqslant 0}$, let $\pi^* = (\pi^*_k)_{k\geqslant 1}$ be the distribution defined as $\pi^*_k = \pi_{k-1}$ for $k \geqslant 1$. Moreover, let $\mathcal{S}(\pi^*) = \{k \geqslant 1 : \pi^*_k > 0\}$ be the support of π^*.

Lemma 5.1. *Let $(W_n)_{n\geqslant 0}$ be a sequence of non-negative random variables. Let $w > 0$, let $\pi = (\pi_k)_{k\geqslant 0}$ be a probability distribution, and let $(m_k)_{k\in\mathcal{S}(\pi^*)}$ be positive numbers. If*

(i) for each $k \in \mathcal{S}(\pi^*)$ there is $\varepsilon > 0$ such that $\limsup_{n\to\infty} \mathbb{E}W_n^{k+\varepsilon} < \infty$,

(ii) $\lim_{n\to\infty} \mathbb{E}W_n^k = m_k$ for all $k \in \mathcal{S}(\pi^*)$, and

(iii) for each n there is a coupling $\big(W_n, B_n W_n^{(\pi^*)}\big)$, where $W_n^{(\pi^*)}$ has the π^*-power-bias distribution of W_n defined through (2.1), where $B_n \sim \mathrm{Beta}(w, 1)$ is independent of $W_n^{(\pi^*)}$, and such that as $n \to \infty$,

$$\mathscr{L}\big(W_n - B_n W_n^{(\pi^*)}\big) \to 0,$$

then $\mathscr{L}(W_n) \to \mathrm{UL}\big(w; (a_k)_{k\geqslant 1}\big)$ as $n \to \infty$, where $a_k = \pi^*_k/m_k$ for $k \in \mathcal{S}(\pi^*)$ and $a_k = 0$ for $k \notin \mathcal{S}(\pi^*)$.

Proof. From (i) we conclude that $\limsup_{n\to\infty} \mathbb{E}W_n < \infty$, so that the sequence $(\mathscr{L}(W_n))_{n\geqslant 1}$ is tight. Thus, we assume that $\mathscr{L}(W_n) \to \mathscr{L}(W)$ and show that this implies $W \sim \mathrm{UL}\big(w; (a_k)_{k\geqslant 1}\big)$. As per Proposition 2.2, it is enough to show that

$$(a)\ \mathbb{E}W^k = m_k, \quad k \in \mathcal{S}(\pi^*), \qquad \text{and} \qquad (b)\ \mathscr{L}(W) = \mathscr{L}\big(V_w W^{(\pi^*)}\big), \tag{5.1}$$

where $V_w \sim \mathrm{Beta}(w, 1)$ is independent of $W^{(\pi^*)}$. Now, (i), (ii) and dominated convergence (see, for example, Durrett (2010, Exercise 3.2.5)) imply that $\mathbb{E}W^k = m_k$ for $k \in \mathcal{S}(\pi^*)$, which is (a). Using (iii) and Slutsky's theorem, we conclude that $\mathscr{L}\big(B_n W_n^{(\pi^*)}\big) \to \mathscr{L}(W)$. But we also have that $\mathscr{L}\big(B_n W_n^{(\pi^*)}\big) \to \mathscr{L}\big(V_w W^{(\pi^*)}\big)$. Indeed, first show $\mathscr{L}(W_n^{(\pi^*)}) \to \mathscr{L}(W^{(\pi^*)})$: for bounded and continuous f,

$$\mathbb{E}f\big(W_n^{(\pi^*)}\big) = \sum_{k\geqslant 1} \pi^*_k \frac{\mathbb{E}\big(W_n^k f(W_n)\big)}{\mathbb{E}W_n^k} \leqslant \|f\|_\infty, \tag{5.2}$$

and by (i) and dominated convergence, $\mathbb{E}\big(W_n^k f(W_n)\big) \to \mathbb{E}\big(W^k f(W)\big)$. So by bounded convergence applied to the sum in (5.2), as $n \to \infty$,

$$\mathbb{E}f\big(W_n^{(\pi^*)}\big) = \sum_{k\geqslant 1} \pi^*_k \frac{\mathbb{E}\big(W_n^k f(W_n)\big)}{\mathbb{E}W_n^k} \longrightarrow \sum_{k\geqslant 1} \pi^*_k \frac{\mathbb{E}\big(W^k f(W)\big)}{\mathbb{E}W^k} = \mathbb{E}f\big(W^{(\pi^*)}\big).$$

Moreover, it's obvious that $\mathscr{L}(B_n) \to \mathscr{L}(V_w)$, and, using independence of the relevant pairs of variables, $\mathscr{L}\big((B_n, W_n^{(\pi^*)})\big) \to \mathscr{L}\big((V_w, W^{(\pi^*)})\big)$. Now the continuous mapping theorem implies $\mathscr{L}\big(B_n W_n^{(\pi^*)}\big) \to \mathscr{L}\big(V_w W^{(\pi^*)}\big)$, as desired. Combining these facts, we find that (b) also holds, and it follows from Proposition 2.2 that $W \sim \mathrm{UL}\big(w; (a_k)_{k\geqslant 1}\big)$. □

Our strategy to proof Theorem 1.3 is to apply Lemma 5.1 to

$$W_n = \frac{X_n}{n^{\mu/(1+\mu)}}.$$

Assuming that π has all positive moments finite or infinite mean, and then choosing m_k as in (1.5), we conclude that (i) and (ii) of Lemma 5.1 are satisfied. Thus, it is sufficient to show (iii) of Lemma 5.1. We develop the coupling of W_n to a variable distributed as $V_{w,n}W_n^{(\pi^*)}$ over a series of lemmas, working first on $W_n^{(\pi^*)}$. Denote the rising factorial

$$x^{\overline{k}} := x(x+1)\dots(x+k-1).$$

Definition 5.2. Let ψ be a probability distribution concentrated on the positive integers, and let W be a positive random variable such that $\mathbb{E}W^k < \infty$, for all k in $\mathcal{S}(\psi)$. A random variable $W^{[\psi]}$ is said to have the *ψ-rising-factorial-bias distribution of W* if

$$\mathbb{E}f\big(W^{[\psi]}\big) = \sum_{k\in\mathcal{S}(\psi)} \psi_k \frac{\mathbb{E}\big(W^{\overline{k}}f(W)\big)}{\mathbb{E}W^{\overline{k}}} \tag{5.3}$$

for all f for which the expectation on the right hand side exists. If $\psi_k = 1$ for some $k \geqslant 1$, then we simply write $W^{[k]}$ to denote $W^{[\psi]}$.

The next lemma relates the π^*-rising-factorial-bias distribution of W_n to its π^*-power-bias distribution.

Lemma 5.3. *Let b and w be positive integers, let π be a distribution on the non-negative integers, let $\tau \sim \pi$, and assume that either*

- (a) $\mathbb{E}\tau^p < \infty$ *for all* $p \geqslant 1$, *or*
- (b) *there is* $\varepsilon > 0$ *such that* $\mathbb{P}[\tau > n] \geqslant Cn^{-(1-\varepsilon)}$ *for n large enough.*

Let $X_n \sim \mathcal{P}^\pi\binom{b}{w};n)$, and let $X_n^{(\pi^*)}$, respectively $X_n^{[\pi^*]}$, have the π^*-power-bias, respectively the π^*-rising-factorial-bias distribution of X_n. Then

$$d_{\mathrm{TV}}\Big(\mathscr{L}\big(X_n^{(\pi^*)}\big), \mathscr{L}\big(X_n^{[\pi^*]}\big)\Big) \to 0 \qquad \text{as } n\to\infty.$$

Proof. We show that for each fixed $k \geqslant 1$,

$$d_{\mathrm{TV}}\Big(\mathscr{L}\big(X_n^{(k)}\big), \mathscr{L}\big(X_n^{[k]}\big)\Big) \to 0, \tag{5.4}$$

from which the lemma follows by bounded convergence and the fact that in general, for random variables (X, Y, U) defined on the same probability space,

$$d_{\mathrm{TV}}(\mathscr{L}(X), \mathscr{L}(Y)) \leqslant \mathbb{E}\, d_{\mathrm{TV}}\big(\mathscr{L}(X|U), \mathscr{L}(Y|U)\big).$$

Both $X_n^{(k)}$ and $X_n^{[k]}$ have densities with respect to X_n, and so

$$2\, d_{\mathrm{TV}}\Big(\mathscr{L}\big(X_n^{(k)}\big), \mathscr{L}\big(X_n^{[k]}\big)\Big)$$
$$= \sum_{j\geqslant 0} \mathbb{P}(X_n = j)\left|\frac{j^{\overline{k}}}{\mathbb{E}D_{k,n}} - \frac{j^k}{\mathbb{E}X_n^k}\right|$$

$$= \sum_{j\geqslant 0} \frac{\mathbb{P}(X_n = j)}{\mathbb{E}D_{k,n}} \left| j^k \left(1 - \frac{\mathbb{E}D_{k,n}}{\mathbb{E}X_n^k}\right) + \sum_{i=0}^{k-1} \begin{bmatrix} k \\ i \end{bmatrix} j^i \right|$$

$$\leqslant \frac{\mathbb{E}X_n^k}{\mathbb{E}D_{k,n}} \left| 1 - \frac{\mathbb{E}D_{k,n}}{\mathbb{E}X_n^k} \right| + \frac{1}{\mathbb{E}D_{k,n}} \sum_{i=0}^{k-1} \begin{bmatrix} k \\ i \end{bmatrix} \mathbb{E}X_n^i, \tag{5.5}$$

where the $\left[\begin{smallmatrix} k \\ i \end{smallmatrix}\right]$ are unsigned Stirling numbers of the first kind. But due to the moment or tail assumptions on π, Proposition 4.1 implies that $\mathbb{E}D_{i,n} = \Theta(n^{i\mu/(\mu+1)})$ for all $i = 1, \ldots$. Therefore, $\mathbb{E}X_n^i$ must be of the same order, and moreover, $\mathbb{E}D_{k,n}/\mathbb{E}X_n^k \to 1$ as $n \to \infty$. Applying these facts with (5.5) implies the lemma. □

We use the rising factorial bias distribution because it can be connected back to our (unbiased) urn models.

Lemma 5.4. *Let b and w be positive integers, let π be a distribution on the non-negative integers, let $\tau \sim \pi$, and assume that either (a) or (b) from Proposition 4.1 holds. Let $X_n \sim \mathcal{P}^\pi\big(\begin{smallmatrix} b \\ w \end{smallmatrix}; n\big)$, and let $\big(Y_n(k)\big)_{k\in\mathcal{S}(\pi^*),n\geqslant 0}$ be a family of random variable such that $Y_n(k) + k \sim \mathcal{P}^\pi\big(\begin{smallmatrix} b \\ w+k \end{smallmatrix}; n\big)$. If $X_n^{[\pi^*]}$ has the π^*-rising-factorial-bias distribution of X_n, then*

$$d_{\mathrm{TV}}\Big(\mathscr{L}\big(X_n^{[\pi^*]}\big), \mathscr{L}\big(Y_n(\tau^*)\big)\Big) \to 0 \qquad \text{as } n \to \infty,$$

where $\tau^* \sim \pi^*$ is independent of $\big(Y_n(k)\big)_{k\in\mathcal{S}(\pi^*),n\geqslant 0}$.

Proof. As in the start of the proof of Lemma 5.3, it is sufficient to show that for each $k \geqslant 1$,

$$d_{\mathrm{TV}}\Big(\mathscr{L}\big(X_n^{[k]}\big), \mathscr{L}\big(Y_n(k)\big)\Big) \to 0, \tag{5.6}$$

For the remainder of the proof, we keep $k \geqslant 1$ fixed and, thus, drop it from out notation. We define three urn process, coupled together through the immigration times in the following way.

First, let $X = (X_0, X_1, X_2, \ldots)$ be a realisation of the immigration urn model starting with b black and w white balls, and with immigration distribution π; let $T = (T_1, T_2, \ldots)$ be the corresponding sequence of arrival times of immigrating black balls. Second, let $\tilde{X} = (\tilde{X}_1, \tilde{X}_2, \ldots)$ be a sequence of random variables such that, given T,

$$\mathbb{P}[\tilde{X}_n = j|T] = \frac{j^{\overline{k}}\mathbb{P}[X_n = j|T]}{\mathbb{E}(D_{k,n}|T)}; \tag{5.7}$$

that is, $\tilde{X}_n$ has the k-rising-factorial-bias distribution of X_n conditional on T. Third, let $Y_1, Y_2, \ldots$ be a realisation of the urn model starting with b black and $w+k$ white balls, where the immigration times are also T. We note that, given T, the joint distribution of the three processes is not going to be relevant.

Applying representation (4.2) from Lemma 4.2 to $Y_n + k \sim \mathcal{P}^\pi\big(\begin{smallmatrix} b \\ w+k \end{smallmatrix}; n\big)$, we obtain that, for any $l \geqslant 1$,

$$\mathbb{E}\left\{ \prod_{j=0}^{l-1} (Y_n + k + j) \,\middle|\, T \right\}$$

$$= \frac{\Gamma(w+k+l)}{\Gamma(w+k)} \prod_{j=0}^{n-1} \frac{b+w+k+l+j+N_j}{b+w+k+j+N_j}$$

$$
\begin{aligned}
&= \frac{\Gamma(w)}{\Gamma(w+k)} \prod_{j=0}^{n-1} \frac{b+w+j+N_j}{b+w+k+j+N_j} \times \frac{\Gamma(w+k+l)}{\Gamma(w)} \prod_{j=0}^{n-1} \frac{b+w+k+l+j+N_j}{b+w+j+N_j} \\
&= \frac{1}{\mathbb{E}\big(D_{k,n}|T\big)} \times \mathbb{E}\bigg\{ \prod_{j=0}^{k+l-1} (X_n+j) \,\bigg|\, T \bigg\} \\
&= \frac{1}{\mathbb{E}\big(D_{k,n}|T\big)} \times \mathbb{E}\bigg\{ \prod_{j=0}^{k-1} (X_n+j) \times \prod_{j=0}^{l-1} (X_n+k+j) \,\bigg|\, T \bigg\} \\
&= \mathbb{E}\bigg\{ \prod_{j=0}^{l-1} (\tilde{X}_n+k+j) \,\bigg|\, T \bigg\}.
\end{aligned}
$$

Taking expectations on both sides of the previous display and using the method of moments, we deduce that, in fact, $\mathscr{L}(Y_n) = \mathscr{L}(\tilde{X}_n)$ for all $n \geqslant 0$. Thus, we have reduced the problem to showing that, as $n \to \infty$,

$$
d_{\mathrm{TV}}\big(\mathscr{L}\big(X_n^{[k]}\big), \mathscr{L}\big(\tilde{X}_n\big)\big) \to 0. \tag{5.8}
$$

Using (5.3) and (5.7), we find

$$
\begin{aligned}
2\, d_{\mathrm{TV}}\big(\mathscr{L}\big(X_n^{[k]}\big), \mathscr{L}\big(\tilde{X}_n\big)\big) &= \sum_{j\geqslant 0} \bigg| \mathbb{E}\bigg\{ \frac{j^{\overline{k}}\mathbb{P}[X_n=j|T]}{\mathbb{E}(D_{k,n}|T)} - \frac{j^{\overline{k}}\mathbb{P}[X_n=j|T]}{\mathbb{E}D_{k,n}} \bigg\} \bigg| \\
&\leqslant \mathbb{E}\bigg\{ \bigg| 1 - \frac{\mathbb{E}(D_{k,n}|T)}{\mathbb{E}D_{k,n}} \bigg| \sum_{j\geqslant 0} \frac{j^{\overline{k}}\mathbb{P}[X_n=j|T]}{\mathbb{E}(D_{k,n}|T)} \bigg\} \\
&= \mathbb{E}\bigg| 1 - \frac{\mathbb{E}(D_{k,n}|T)}{\mathbb{E}D_{k,n}} \bigg|.
\end{aligned} \tag{5.9}
$$

Now, by Lemma 4.3 and Proposition 4.1, we have that, almost surely,

$$
\frac{\mathbb{E}(D_{k,n}|T)}{\mathbb{E}D_{k,n}} \to 1 \qquad \text{as } n \to \infty.
$$

Moreover, Jensen's inequality implies $\mathbb{E}(\mathbb{E}(D_{k,n}|T)^2) \leqslant \mathbb{E}D_{k,n}^2 \leqslant \mathbb{E}D_{2k,n}$, and thus, again by Proposition 4.1,

$$
\sup_{n\geqslant 1} \frac{\mathbb{E}D_{2k,n}}{\big(\mathbb{E}D_{k,n}\big)^2} < \infty.
$$

Hence, by dominated convergence, the right hand side of (5.9) tends to zero, which concludes the proof. □

The next two lemmas move us from $Y_n(\tau^*)$ defined in Lemma 5.4 to a variable that will be used as a surrogate for $W_n^{(\pi^*)}$.

Lemma 5.5. *Let w be a positive integer, let π be a probability distribution on the non-negative integers. Let $(Y_n(k))_{k\in\mathcal{S}(\pi^*)}$ be a family of random variables such that $Y_n(k) + k \sim \mathcal{P}^\pi\big({1 \atop w+k}; n\big)$ for $k \in \mathcal{S}(\pi^*)$. Moreover, let $\tilde{X} = (\tilde{X}_0, \tilde{X}_1, \dots)$ be a realisation of an immigration urn process with immigration distribution π, starting with zero black balls and $w+1$ white balls, so that $\tilde{X}_n \sim \mathcal{P}^\pi\big({0 \atop w+1}; n\big)$. Let $\tilde{\tau}$ be time of the first arrival in the urn process $\tilde{X}$, and let $\tilde{Y}_n = \tilde{X}_{n+\tilde{\tau}} - \tilde{\tau} - 1$. Then*

$$
\mathscr{L}\big(Y_n(\tau^*)\big) = \mathscr{L}\big(\tilde{Y}_n\big),
$$

where $\tau^* \sim \pi^*$ is independent of $(Y_n(k))_{k\geqslant 1}$.

Proof. Consider the urn process $\tilde{X}$. Since there are no black balls in the urn initially, the first $\tilde{\tau}$ draws all come up white and so $\tilde{\tau}$ white balls are added to the urn, $\tilde{\tau}$ steps elapse, and one black ball is added. At this point there are $w + \tilde{\tau} + 1$ white balls in the urn and 1 black ball. Thus we find that $\mathcal{P}^{\pi}\binom{0}{w+1}; n + \tilde{\tau}) = \mathcal{P}^{\pi}\binom{1}{w+\tilde{\tau}+1}; n)$, which is exactly the statement of the lemma. □

Lemma 5.6. *Let w be a positive integer, let π be a probability distribution on the non-negative integers, let $\tilde{X} = (\tilde{X}_0, \tilde{X}_1, \dots)$, $\tilde{\tau}$ and $(\tilde{Y}_n)_{n\geqslant 1}$ be defined as in Lemma 5.5. Then,*

$$\frac{1}{n^{\mu/(\mu+1)}}\big(\tilde{Y}_n - (\tilde{X}_n - w - 1)\big) \xrightarrow{P} 0.$$

Proof. The only difference between the two variables is the number of steps the process is run, and the shifts $\tilde{\tau}$ and $w - 1$. Since, at each time step, the number of white balls in the urn increase by at most one, we have

$$\big|\tilde{Y}_n - (\tilde{X}_n - w - 1)\big| = \big|\tilde{X}_{n+\tilde{\tau}} - \tilde{\tau} - 1 - (\tilde{X}_n - w - 1)\big| \leqslant 2\tilde{\tau} + w.$$

Divided by the scaling $n^{\mu/(\mu+1)}$, the right hand side tends to zero in probability. □

The previous lemmas imply we can use $n^{-\mu/(\mu+1)}(\tilde{X}_n - w - 1)$ as a surrogate for $W_n^{(\pi^*)}$, and the next result shows how to relate this variable back to the original W_n using a classical Pólya urn.

Lemma 5.7. *Let w be a positive integer and let π be a probability distribution on the non-negative integers. Let $\tilde{X} = (\tilde{X}_1, \tilde{X}_2, \dots)$ be as in Lemma 5.5, and let $\big(Q_w(n)\big)_{n\geqslant 0}$ be the number of white balls in a classical Pólya urn sequence started with 1 black ball and w white balls. Then*

$$Q_w\big(\tilde{X}_n - w - 1\big) \sim \mathcal{P}^{\pi}\tbinom{1}{w}; n).$$

Proof. Start with an urn having w white balls, 1 gray ball, and 0 black balls. The urn follows the rules of a classical Pólya urn with three colors, but at the arrival times $T_1, T_2, \dots$, driven by π, a black ball is added to the urn. It is clear that $\tilde{X}_n - w - 1$ equals the number of times a gray or white ball is drawn after n steps in this urn process, and each time a gray or white ball is drawn, the chance of it being white is proportional to the number of white balls in the urn at that moment, just as in a classical Pólya urn. So $Q_w\big(\tilde{X}_n - w - 1\big)$ is distributed as the number of white balls in the described urn after n steps, but this distribution is exactly $\mathcal{P}^{\pi}\binom{1}{w}; n)$ since the 1 gray ball can now be viewed as a "black" ball. □

To get to the beta variable V_w in the coupling (and to transfer to the general $b > 1$ case), we need the result of Peköz et al. (2014, Lemma 2.3), which provides a close coupling of a classical Pólya urn to its beta limit. Denote by $\mathcal{P}\binom{b}{w}; n)$ the law of the number of white balls in a classical Pólya urn started with b black and w white balls after n draws and replacements.

Lemma 5.8. *Let β, ω and k be positive integers. There is a coupling $\big(Q_{\beta,\omega}(n), V_{\beta,\omega}\big)$ with $Q_{\beta,\omega}(n) \sim \mathcal{P}\binom{\beta}{\omega}; n)$ and $V_{\beta,\omega} \sim \mathrm{Beta}(\omega, \beta)$, such that, almost surely,*

$$|Q_{\beta,\omega}(n) - nV_{\beta,\omega}| < \beta(4\omega + \beta + 1).$$

We are now in position to complete the proof of Theorem 1.3.

Proof of Theorem 1.3. We first show the result for $b = 1$. Let $W_n = n^{-\mu/(\mu+1)} X_n$. By (1.5) and because either (a) or (b) is satisfied, there is a sequence $m = (m_1, m_2, \dots)$ such that $\mathbb{E}W_n^k \to m_k$. We want to show that $\mathscr{L}(W_n) \to \mathrm{UL}\big(w; (a_k)_{k\geqslant 1}\big)$, and we do so by showing (i), (ii) and (iii) of Lemma 5.1. By (1.5), (i), (ii) easily follow. To show (iii), Lemmas 5.3–5.6 imply that we can couple variables $\big(W_n^{(\pi^*)}, n^{-\mu/(\mu+1)}\hat{X}_n\big)$, where $W_n^{(\pi^*)}$ has the π^*-power-bias distribution of W_n and $\hat{X}_n = \tilde{X}_n - w - 1$, such that

$$\big(W_n^{(\pi^*)} - n^{-\mu/(\mu+1)}\hat{X}_n\big) \xrightarrow{P} 0 \qquad \text{as } n \to \infty. \tag{5.10}$$

Moreover, Lemma 5.7 implies

$$\hat{W}_n := \frac{Q_w\big(\hat{X}_n\big)}{n^{\mu/(\mu+1)}} \overset{\mathscr{D}}{=} W_n,$$

where $Q_w(n)$ is defined as in Lemma 5.7. By Lemma 5.8, there is a coupling $(Q_w\big(\hat{X}_n\big), V_w\hat{X}_n)$ with $V_w \sim \mathrm{Beta}(w, 1)$ independent of $\hat{X}_n$ and such that

$$\big|Q_w\big(\hat{X}_n\big) - V_w\hat{X}_n\big| < w + 1$$

almost surely. From these last two displays, we have a coupling $\big(V_w W_n^{(\pi^*)}, \hat{W}_n\big)$ with the appropriate marginals satisfying

$$\begin{aligned}\big|V_w W_n^{(\pi^*)} - \hat{W}_n\big| &\leqslant \big|V_w W_n^{(\pi^*)} - n^{-\mu/(\mu+1)}V_w\hat{X}_n\big| + \big|n^{-\mu/(\mu+1)}V_w\hat{X}_n - \hat{W}_n\big| \\ &< \big|W_n^{(\pi^*)} - n^{-\mu/(\mu+1)}\hat{X}_n\big| + n^{-\mu/(\mu+1)}(w+1),\end{aligned}$$

which according to (5.10) tends to zero in probability, as desired. Finally, the convergence of the moments of W_n to those of its limit follows since $\mathbb{E}W_n^k \to m_k < \infty$ for all $k \geqslant 1$; this implies (1.7).

For the general case $b > 1$, let $X_n \sim \mathcal{P}^\pi\big(\substack{b\\w}; n\big)$ and $X_n' \sim \mathcal{P}^\pi\big(\substack{1\\w+b-1}; n\big)$. We show that

$$\mathscr{L}(X_n) = \mathcal{P}\big(\substack{b-1\\w}; X_n' - (b+w-1)\big), \tag{5.11}$$

and then the result follows easily from Lemma 5.8. To establish (5.11), consider an urn that at step zero has w white balls, $b-1$ gray balls, and 1 black ball. The urn follows the rules of a classical Pólya urn but at the arrival times $T_1, T_2, \dots$, driven by π, a black ball is added to the urn. It is clear that $X_n' - (b+w-1)$ is distributed as the number of times a gray or white ball is drawn after n steps in this urn process, and each time a gray or white ball is drawn, the chance it is white is proportional to the number of white balls in the urn at that moment, just as in a classical Pólya urn. So $\mathcal{P}\big(\substack{b-1\\w}; X_n' - (b+w-1)\big)$ is the distribution of the number of white balls in the urn process after n steps, and this is exactly $\mathcal{P}^\pi\big(\substack{b\\w}; n\big)$ if we now view the $b-1$ gray balls as black. □

6 SOME PROPERTIES OF THE **UL** FAMILY

In this section we derive some basic properties of the UL family. First we record some moment and tail bounds.

Proposition 6.1 (Moment Bounds). *Fix w and $(a_k)_{k\geqslant 1}$, let $Z \sim \mathrm{UL}\big(w; (a_k)_{k\geqslant 1}\big)$, and let c be the normalising constant from* (1.2)*, depending only on w and $(a_k)_{k\geqslant 1}$. Then for any positive integer m,*

$$\mathbb{E}Z^m \leqslant \inf_{\{\ell : a_\ell > 0\}} \frac{c}{\ell}\left(\frac{wa_\ell}{\ell}\right)^{-(w+m)/\ell}\Gamma\left(\frac{w+m}{\ell}\right).$$

Moreover, $\mathrm{UL}\big(w; (a_k)_{k\geqslant 1}\big)$ *is uniquely determined by its moments.*

Proof. If ℓ is such that $a_\ell > 0$, we have

$$\begin{aligned}\mathbb{E}Z^m &= c\int_0^\rho x^{w+m-1}e^{-w\sum_{k\geqslant 1}\frac{a_k}{k}x^k}dx \\ &\leqslant c\int_0^\rho x^{w+m-1}e^{-w\frac{a_\ell}{\ell}x^\ell}dx \\ &\leqslant c\int_0^\infty x^{w+m-1}e^{-w\frac{a_\ell}{\ell}x^\ell}dx \\ &= \frac{c}{\ell}\left(\frac{wa_\ell}{\ell}\right)^{-(w+m)/\ell}\Gamma\left(\frac{w+m}{\ell}\right),\end{aligned}$$

which proves the first assertion. For the second, the bound above and Stirling's approximation shows that

$$\limsup_{m\to\infty}\frac{\left(\mathbb{E}Z^m\right)^{1/2m}}{m} < \infty,$$

and so in particular, Carleman's condition for the Stieljes moment problem is satisfied. □

Proposition 6.2 (Mills Ratio Tail Bound)**.** *Fix w and $(a_k)_{k\geqslant 1}$, and let $Z \sim \mathrm{UL}\big(w;(a_k)_{k\geqslant 1}\big)$. For each $\alpha > 0$, there is a constant C_α such that for $x > \alpha$,*

$$P(Z \geqslant x) \leqslant C_\alpha u(x).$$

Proof. We show that $\frac{P(Z\geqslant x)}{u(x)}$ is non-increasing in x, from which the proposition follows with $C_\alpha := \frac{P(Z\geqslant\alpha)}{u(\alpha)}$. Note that $u(x) = ce^{-B(x)}$, where we define

$$B(x) := -(w-1)\log(x) + \sum_{k\geqslant 1}\frac{a_k}{k}x^k.$$

Note that $B''(x) \geqslant 0$ so B' is non-decreasing. Then

$$\frac{d}{dx}\left(\frac{P(Z\geqslant x)}{u(x)}\right) = B'(x)e^{B(x)}\int_x^\rho e^{-B(y)}dy - 1 \leqslant e^{B(x)}\int_x^\rho B'(y)e^{-B(y)}dy - 1 = 0. \quad \square$$

In Theorem 1.3 we showed that if $b > 1$, then the limiting distribution of our urn model can be expressed as $\mathrm{UL}\big(w;(a_k)_{k\geqslant 1}\big)$ multiplied by a beta random variable. It is natural to ask if such distributions are again in the UL family. Our next examples show that this is not true in general, not even for the limits appearing in Theorem 1.3.

Example 6.3. If $U \sim \mathrm{Beta}(1,1)$ and $X \sim \mathrm{Exp}(1)$, then $\mathscr{L}(UX)$ is not in the UL family. Indeed, the density of UX for $x > 0$ is $\int_x^\infty e^{-t}/t\,dt$, which goes to infinity like $-\log(x)$ as $x \to 0$, and hence is not in the UL class.

Example 6.4. Let $\pi = \delta_1$ be the point mass at 1; as discussed in Section 3.1, the scaled limit of $\mathcal{P}^{\delta_1}(\frac{1}{w};n)$ has density proportional to

$$x^{w-1}\exp\{-Cx^2\}dx \tag{6.1}$$

for some constant C. By Theorem 1.3, the scaled limit of $\mathcal{P}^{\delta_1}(\frac{2}{1};n)$ has distribution $\mathscr{L}(BZ)$, where Z has density proportional to (6.1) with $w = 2$, and where $B \sim \mathrm{Beta}(1,1)$ is independent of Z. Using the density formula for products of independent random variables, we obtain that BZ has density proportional to

$$\int_x^\infty e^{-Cy^2}dy$$

which, up to scaling and multiplicative constants, is known as the *complementary error function* erfc(x). If $BZ \sim \mathrm{UL}(v, (a'_k)_{k\geqslant 1})$ for some positive integer v and positive sequence $(a'_k)_{k\geqslant 1}$, then since $\lim_{x\to 0}\int_x^\infty e^{-Cy^2}dy > 0$, we must have $v = 1$. In this case, a'_k are just the coefficients in the Taylor series expansion about zero of $-\log(\int_x^\infty e^{-Cy^2}dy)$, but since

$$-\frac{\partial^4}{\partial x^4}\log \mathrm{erfc}(x)\bigg|_{x=0} = \frac{32(3-\pi)}{\pi^2} < 0,$$

we would have $a'_4 = \frac{32(3-\pi)}{4!\pi^2} < 0$ in representation (1.2), so that BZ cannot be in the UL family.

7 ACKNOWLEDGEMENTS

AR is supported by NUS Research Grant R-155-000-167-112. EP, AR, and NR are supported by ARC grant DP150101459. This work was done partially while the authors were visiting the Institute for Mathematical Sciences, National University of Singapore in 2015 and 2016, supported in part by the Institute. We thank a referee for their careful reading and constructive comments.

REFERENCES

Abramowitz, M. and Stegun, I. A. (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series.* For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C.

Arratia, R., Goldstein, L., and Kochman, F. (2013). Size bias for one and all. Preprint `http://arxiv.org/abs/1308.2729`.

Barbour, A. D., Holst, L., and Janson, S. (1992). *Poisson approximation*, volume 2 of *Oxford Studies in Probability.* The Clarendon Press Oxford University Press, New York. Oxford Science Publications.

Bartroff, J. and Goldstein, L. (2013). A Berry-Esseen bound for the uniform multinomial occupancy model. *Electron. J. Probab.*, 18:no. 27, 29.

Berger, N., Borgs, C., Chayes, J. T., and Saberi, A. (2014). Asymptotic behavior and distributional limits of preferential attachment graphs. *Ann. Probab.*, 42(1):1–40.

Brown, M. (2006). Exploiting the waiting time paradox: applications of the size-biasing transformation. *Probab. Engrg. Inform. Sci.*, 20(2):195–230.

Chauvin, B., Mailler, C., and Pouyanne, N. (2015). Smoothing equations for large Pólya urns. *J. Theoret. Probab.*, 28(3):923–957.

Chauvin, B., Pouyanne, N., and Sahnoun, R. (2011). Limit distributions for large Pólya urns. *Ann. Appl. Probab.*, 21(1):1–32.

Chen, L. H. Y., Goldstein, L., and Shao, Q.-M. (2011). *Normal approximation by Stein's method.* Probability and its Applications (New York). Springer, Heidelberg.

Chen, M.-R. and Kuba, M. (2013). On generalized Pólya urn models. *J. Appl. Probab.*, 50(4):1169–1186.

Chen, M.-R. and Wei, C.-Z. (2005). A new urn model. *J. Appl. Probab.*, 42(4):964–976.

Cooper, C. and Frieze, A. (2003). A general model of web graphs. *Random Structures Algorithms*, 22(3):311–335.

Deijfen, M., van den Esker, H., van der Hofstad, R., and Hooghiemstra, G. (2009). A preferential attachment model with random initial degrees. *Ark. Mat.*, 47(1):41–72.

Dharmadhikari, S. W. and Jogdeo, K. (1969). Bounds on moments of certain random variables. *Ann. Math. Statist.*, 40:1506–1509.

Durrett, R. (2010). *Probability: theory and examples.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, fourth edition.

Eggenberger, F. and Pólya, G. (1923). Über die statistik verketteter vorgänge. *Z. angew. Math Mech.*, 3:279–289.

Flajolet, P., Gabarró, J., and Pekari, H. (2005). Analytic urns. *Ann. Probab.*, 33(3):1200–1233.

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

Janson, S. (2006). Limit theorems for triangular urn schemes. *Probab. Theory Related Fields*, 134(3):417–452.

Knape, M. and Neininger, R. (2014). Pólya urns via the contraction method. *Combin. Probab. Comput.*, 23(6):1148–1186.

Kuba, M. and Mahmoud, H. M. (2015a). Two-color balanced affine urn models with multiple drawings II: large-index and triangular urns. Preprint https://arxiv.org/abs/1509.09053.

Kuba, M. and Mahmoud, H. M. (2015b). Two-colour balanced affine urn models with multiple drawings I: central limit theorems. Preprint https://arxiv.org/abs/1503.09069.

Laruelle, S. and Pagès, G. (2013). Randomized urn models revisited using stochastic approximation. *Ann. Appl. Probab.*, 23(4):1409–1436.

Mahmoud, H. M. (2009). *Pólya urn models*. Texts in Statistical Science Series. CRC Press, Boca Raton, FL.

Pakes, A. G. and Navarro, J. (2007). Distributional characterizations through scaling relations. *Aust. N. Z. J. Stat.*, 49(2):115–135.

Peköz, E. A., Röllin, A., and Ross, N. (2013a). Degree asymptotics with rates for preferential attachment random graphs. *Ann. Appl. Probab.*, 23(3):1188–1218.

Peköz, E. A., Röllin, A., and Ross, N. (2013b). Total variation error bounds for geometric approximation. *Bernoulli*, 19(2):610–632.

Peköz, E. A., Röllin, A., and Ross, N. (2014). Joint degree distributions of preferential attachment random graphs. Preprint http://arxiv.org/abs/1402.4686. To appear in *Adv. Appl. Probab.*

Peköz, E. A., Röllin, A., and Ross, N. (2016). Generalized gamma approximation with rates for urns, walks and trees. *Ann. Probab.*, 44(3):1776–1816.

Pemantle, R. (2007). A survey of random processes with reinforcement. *Probab. Surv.*, 4:1–79.

Petrov, V. V. (1975). *Sums of independent random variables*. Springer-Verlag, New York-Heidelberg. Translated from the Russian by A. A. Brown, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82.

Pouyanne, N. (2008). An algebraic approach to Pólya processes. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44(2):293–323.

Ross, N. (2013). Power laws in preferential attachment graphs and Stein's method for the negative binomial distribution. *Adv. in Appl. Probab.*, 45(3):876–893.

Titchmarsh, E. C. (1958). *The theory of functions*. Oxford University Press, Oxford. Reprint of the second (1939) edition.

van der Hofstad, R. (2016). Random graphs and complex networks. Final Book Draft of 1 May, 2016 http://www.win.tue.nl/~rhofstad/NotesRGCN.html.