# LIMIT THEOREMS FOR TRIANGULAR URN SCHEMES

SVANTE JANSON

Abstract. We study a generalized Pólya urn with balls of two colours and a triangular replacement matrix; the urn is not required to be balanced. We prove limit theorems describing the asymptotic distribution of the composition of the urn after a long time. Several different types of asymptotics appear, depending on the ratio of the diagonal elements in the replacement matrix; the limit laws include normal, stable and Mittag-Leffler distributions as well as some less familiar ones. The results are in some cases similar to, but in other cases strikingly different from, the results for irreducible replacement matrices.

## 1. Introduction

A generalized Pólya urn may contain balls of several different types (or colours), say $1, \dots, q$. The urn evolves according to a Markov process as follows. At each time $n \geq 1$, one of the balls in the urn is drawn at random. The colour of the drawn ball is inspected and a set of balls, depending on the drawn colour, is added to the urn (with or without replacement of the drawn ball).

We will in this paper for simplicity assume that there are only two colours, black and white. The composition of the urn after $n$ draws may thus be represented as a vector $(X_n, Y_n)$, where $X_n$ [$Y_n$] is the number of black [white] balls in the urn. The urn starts with a given vector $(X_0, Y_0) = (x_0, y_0)$, which we assume is non-random. To fix the notation, we assume that if a black ball is drawn, it is replaced together with $\alpha$ additional black balls and $\beta$ white balls; if a white ball is drawn, it is replaced together with $\gamma$ black and $\delta$ white balls. These numbers may conveniently be collected in the *replacement matrix*

$$A := \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}. \tag{1.1}$$

(We may here allow $\alpha = -1$ or $\delta = -1$; this means that the drawn ball is not replaced. Other negative entries are possible too in some cases, see Remark 1.11 below.)

In Pólya's original urn [7, 22], the added balls are always of the same colour as the drawn ball, i.e. $\beta = \gamma = 0$; further, $\alpha = \delta$. The generalized case described here has been studied by many authors, including [5, 12, 16]. An important case, which includes many applications, is when the replacement

*Date*: March 24, 2004; revised March 18, 2005.

matrix is irreducible; in our setting with two colours this means that $\beta, \gamma > 0$. Limit theorems for the irreducible case have been given by many authors, for example [2, 3, 4]; see also [14] and the further references given there. In contrast, we will here study the case of a triangular replacement matrix, i.e. for two colours $\beta = 0$ or $\gamma = 0$; we will throughout assume that $\beta = 0$. (If $\gamma = 0$, we interchange the two colours.) It has been known for a long time that other phenomena arise in this case [4, 24, 20, 19]; the purpose of this paper is to give a complete description of the asymptotic distribution of $(X_n, Y_n)$ in the triangular case.

**Remark 1.1.** More generally, the replacement vectors $(\alpha, \beta)$ and $(\gamma, \delta)$ may be random. It is possible that our methods can be extended to that case, but we have not attempted it.

We will use the embedding method of Athreya and Karlin [2, 3]: Let $(\mathcal{X}(t), \mathcal{Y}(t))$, $t \geq 0$ be a continuous time Markov branching process with particles of two types (black and white); each particle lives a random time with an exponential distribution Exp(1), and on its death a black [white] particle is replaced by $1 + \alpha$ black and $\beta$ white [$\gamma$ black and $1 + \delta$ white] new particles. It is then easy to see that $(\mathcal{X}(t), \mathcal{Y}(t))$ observed at the (a.s. distinct) times of deaths gives the urn process above.

Our main results are Theorems 1.3 and 1.4 below. Some degenerate cases are studied in Section 2.

Note first that $X_n$ and $Y_n$ are linearly dependent. Indeed, if $\delta = 0$, then $Y_n = y_0$ is non-random, and otherwise $Y_n$ determines $X_n$ by the following lemma. Hence we will mainly state results for $Y_n$; the reader may easily translate these into results for $X_n$.

**Lemma 1.2.** *Suppose $\delta \neq 0$. Then, for every $n$,*

$$X_n - x_0 + \frac{\alpha - \gamma}{\delta}(Y_n - y_0) = n\alpha$$

*Proof.* Each draw increases the left hand side by $\alpha$. □

We consider first the case $\delta, \gamma > 0$. The limits below are as $n \to \infty$.

**Theorem 1.3.** *Consider a generalized Pólya urn with two colours and a triangular replacement matrix $\left(\begin{smallmatrix} \alpha & 0 \\ \gamma & \delta \end{smallmatrix}\right)$, i.e. $\beta = 0$. Suppose further that $\delta > 0$, $\gamma > 0$ and $y_0 > 0$.*

(i) *If $\alpha < \delta/2$, then*

$$n^{-1/2}\Big(Y_n - \delta\frac{\delta - \alpha}{\delta - \alpha + \gamma}n\Big) \xrightarrow{\mathrm{d}} N(0, \sigma^2), \tag{1.2}$$

*where*

$$\sigma^2 = \frac{\gamma\delta^3(\delta - \alpha)}{(\delta - 2\alpha)(\delta - \alpha + \gamma)^2}. \tag{1.3}$$

(ii) *If* $\alpha = \delta/2$, *then*

$$(n \ln n)^{-1/2}\Big(Y_n - \delta\frac{\delta-\alpha}{\delta-\alpha+\gamma}n\Big) \xrightarrow{\mathrm{d}} N(0,\sigma^2), \tag{1.4}$$

*where*

$$\sigma^2 = \frac{\alpha\gamma\delta^2}{(\alpha+\gamma)^2}. \tag{1.5}$$

(iii) *If* $\delta/2 < \alpha < \delta$, *then*

$$n^{-\alpha/\delta}\Big(Y_n - \delta\frac{\delta-\alpha}{\delta-\alpha+\gamma}n\Big) \xrightarrow{\mathrm{d}} W := -\frac{\delta(\delta-\alpha)^{1+\alpha/\delta}}{\alpha(\delta-\alpha+\gamma)^{1+\alpha/\delta}}Z, \tag{1.6}$$

*where* $Z$ *is a random variable with the characteristic function given by* (7.4) *and* (7.5) *below.*

(iv) *If* $\alpha = \delta$, *then*

$$\frac{\ln^2 n}{n}\Big(Y_n - \frac{\alpha^2}{\gamma}\frac{n}{\ln n} - \frac{\alpha^2}{\gamma}\frac{n\ln\ln n}{\ln^2 n}\Big) \xrightarrow{\mathrm{d}} W := \frac{\alpha^2}{\gamma^2}\Big(\gamma\ln\frac{\gamma}{\alpha} + \gamma - \alpha - Z\Big), \tag{1.7}$$

*where* $Z$ *is a random variable with the characteristic function given by* (7.4) *and* (7.6) *below.*

(v) *If* $\alpha > \delta$, *then*

$$n^{-\delta/\alpha}Y_n \xrightarrow{\mathrm{d}} W := \delta\alpha^{\delta/\alpha}Z^{-\delta/\alpha}, \tag{1.8}$$

*where* $Z$ *is a random variable with the characteristic function given by* (7.4) *and* (7.5) *or* (7.7) *below.*

In cases (iii) and (iv), the limit variable $W$ is a linear function of $Z$, and thus the characteristic function of $W$ is immediately obtained from the characteristic function of $Z$. Our formula (7.4) is so complicated, however, that we refrain from stating the result explicitly; the main fact is that a non-trivial limit exists. In case (v), we cannot give the characteristic function of $W$, but its moments are calculated in Theorem 1.6 below. In some special cases we have simpler results, see Theorems 1.7 and 1.8 and Section 8. In particular, if $x_0 = \gamma$ and $y_0 = \delta$, then $Z$ (and thus $W$ in cases (iii) and (iv)) has a $\delta/\alpha$-stable distribution, see Theorem 8.6. The general case remains rather elusive, however, see also Section 3.

Case (iv) of Theorem 1.3 has earlier been studied by Pemantle and Volkov [20, Theorem 2.3]; their result easily implies ours, except that the limit $W$ is not identified.

In the diagonal case $\gamma = \beta = 0$, we have a companion result. Note that the case $\alpha = \delta$ [16] is the original Pólya urn [7, 22], and that the case $\alpha > \delta$ is as in Theorem 1.3, unlike the cases $\alpha \le \delta$. (Theorem 1.4 is much simpler than Theorem 1.3, and has probably been observed before. It is included here mainly for completeness and comparison. Except for the identification of $W$, the case $\alpha \ne \delta$ follows from [20, Theorem 2.2].)

**Theorem 1.4.** *Consider a generalized Pólya urn with two colours and a diagonal replacement matrix* $\left(\begin{smallmatrix}\alpha & 0\\ 0 & \delta\end{smallmatrix}\right)$, *i.e.* $\beta = \gamma = 0$. *Suppose further that* $\alpha > 0$, $\delta > 0$, $x_0 > 0$ *and* $y_0 > 0$. *Let* $U \sim \Gamma(x_0/\alpha, 1)$ *and* $V \sim \Gamma(y_0/\delta, 1)$ *be two independent Gamma distributed random variables.*

(i) *If* $\alpha < \delta$, *then*

$$n^{-\alpha/\delta}(n\delta - Y_n) \xrightarrow{\mathrm{d}} W := \delta U V^{-\alpha/\delta}. \tag{1.9}$$

(ii) *If* $\alpha = \delta$, *then*

$$n^{-1} Y_n \xrightarrow{\mathrm{d}} W := \delta \frac{V}{U+V}; \tag{1.10}$$

*thus* $W/\delta$ *has a* $\mathrm{Beta}(y_0/\delta, x_0/\alpha)$ *distribution.*

(iii) *If* $\alpha > \delta$, *then*

$$n^{-\delta/\alpha} Y_n \xrightarrow{\mathrm{d}} W := \delta U^{-\delta/\alpha} V. \tag{1.11}$$

Another exceptional case is when $\delta = 0$. Then the number of white balls is constant $y_0$. The number $\mathcal{X}(t)$ of black balls in the branching process is a generalized Yule process (see Section 5) with immigration, see e.g. [13, Chapter 7.1]. We have the following limit results.

**Theorem 1.5.** *Consider a generalized Pólya urn with two colours and a replacement matrix* $\left(\begin{smallmatrix}\alpha & 0\\ \gamma & 0\end{smallmatrix}\right)$, *i.e.* $\delta = \beta = 0$.

(i) *If* $\alpha < 0$, $\gamma > 0$, *and* $y_0 > 0$, *then* $(X_n)_{n\ge 0}$ *is a persistent irreducible Markov chain with period* $\gamma/|\alpha| + 1$. *Assume, for convenience,* $\alpha = -1$, *and let* $W$ *have the compound Poisson distribution with probability generating function*

$$\mathbb{E}\, z^W = \sum_{k=0}^{\infty} q_k z^k = \exp\Bigl( y_0 \sum_{j=1}^{\gamma} \frac{z^j - 1}{j} \Bigr), \tag{1.12}$$

*where thus* $q_k = \mathbb{P}(W = k)$. *Then* $X_n \equiv x_0 - n \pmod{\gamma + 1}$, *and for every* $k \ge 0$,

$$\mathbb{P}(X_n = k) - \mathbf{1}[k \equiv x_0 - n \pmod{\gamma+1}] \sum_{j=k-\gamma}^{k} q_j \to 0,$$

*with* $q_j = 0$ *for* $j < 0$. *In other words, if* $W_i := (\gamma+1)\lceil (W-i)/(\gamma+1)\rceil + i$ *is* $W$ *rounded upwards to the nearest integer* $\equiv i \pmod{\gamma+1}$, *then* $X_{(\gamma+1)n+j} \xrightarrow{\mathrm{d}} W_{x_0 - j}$.

(ii) *If* $\alpha = 0$, *then*

$$n^{-1/4}\bigl(X_n - \sqrt{2\gamma y_0}\, n^{1/2}\bigr) \xrightarrow{\mathrm{d}} N\bigl(0, \sqrt{2\gamma^3 y_0/3}\bigr). \tag{1.13}$$

(iii) *If* $\alpha > 0$ *and either* $x_0 > 0$ *or* $\gamma > 0$, *then*

$$(\ln n)^{-1/2}\Bigl(X_n - \alpha n - \frac{\gamma - \alpha}{\alpha} y_0 \ln n\Bigr) \xrightarrow{\mathrm{d}} N\Bigl(0, \frac{(\gamma-\alpha)^2}{\alpha} y_0\Bigr). \tag{1.14}$$

Note that if $\alpha = -1$ and $\gamma = 1$, then $W \sim \mathrm{Po}(y_0)$ in (i).

The limit distributions above are non-degenerate except in some trivial cases, see Section 2. We add some results on the limit $W$ in Theorem 1.3(v).

**Theorem 1.6.** *In Theorem 1.3(v) ($\alpha > \delta > 0$ and $\gamma > 0$), the limit $W$ has moments of all orders given by, for $r > 0$,*

$$\mathbb{E}\, W^r = \delta^r \frac{\Gamma(y_0/\delta + r)}{\Gamma(y_0/\delta)\Gamma(r\delta/\alpha)} \int_0^\infty (1+x)^{-x_0/\alpha} \cdot \Bigl(1 + \frac{\delta}{\alpha} x^{\delta/\alpha} \int_0^x \bigl(1-(1+u)^{-\gamma/\alpha}\bigr) u^{-\delta/\alpha-1}\, du\Bigr)^{-y_0/\delta - r} x^{r\delta/\alpha - 1}\, dx.$$

In the special case $\alpha = \gamma + \delta$, when the same number of balls are added each time, this leads to a simple formula, first found by Puyhaubert [23].

**Theorem 1.7.** *If $\alpha = \gamma + \delta$ and $\gamma, \delta > 0$, the moments of $W$ in* (1.8) *are*

$$\mathbb{E}\, W^r = \delta^r \frac{\Gamma\bigl((x_0+y_0)/\alpha\bigr)\Gamma(y_0/\delta + r)}{\Gamma(y_0/\delta)\Gamma\bigl((x_0+y_0+r\delta)/\alpha\bigr)}, \qquad r \geq 0. \tag{1.15}$$

Note that if we start with $(x_0, y_0) = (\gamma, \delta)$, then this formula simplifies to $\mathbb{E}(W/\delta)^r = r!/\Gamma(1 + r\delta/\alpha)$, and thus $W$ is a Mittag-Leffler distribution with parameter $\delta/\alpha$, i.e. $Z$ has a $\delta/\alpha$-stable distribution [21], [9, XIII.8(b)]; this also follows from Theorem 8.6. We can generalize this as follows. (See also the special case in [10, Corollary 12], [11].)

**Theorem 1.8.** *Suppose that $\alpha = \gamma + \delta$ with $\gamma, \delta > 0$ Suppose further $y_0 > 0$ and either $x_0 = 0$ or $x_0 = \gamma$. Then $W$ in* (1.8) *has a density function $cx^{y_0/\delta - x_0/\gamma} f(x/\delta)$, $x > 0$, where $c > 0$ is a normalizing constant and $f$ is the density function of a Mittag-Leffler distribution with parameter $\delta/\alpha$ having moments $\Gamma(1+r)/\Gamma(1+r\delta/\alpha)$.*

Note that the density function $f(x) = (\alpha/\delta)x^{-\alpha/\delta - 1} g(x^{-\alpha/\delta})$, where $g$ is the density function of a $\delta/\alpha$-stable distribution with Laplace transform $e^{-\lambda^{\delta/\alpha}}$; thus $f(x) = (\alpha/\delta\pi)\sum_{k=1}^\infty \Gamma(k\delta/\alpha + 1)\sin(k\pi\delta/\alpha)(-x)^{k-1}/k!$, $x > 0$, see [9, XIII.8(b) and XVII.6].

Theorem 1.8 covers only two values of $x_0$. (These are clearly equivalent, since starting with $(0, y_0)$, we necessarily first draw a white ball, and thus the process is the same as starting with $(\gamma, y_0 + \delta)$, with $n$ shifted one step.) It is possible to use Theorems 1.6 and 1.8 to find expressions for the density of $W$ also in some cases with other $x_0$, for example if $x_0 = \alpha$ or $\alpha + \gamma$. We give an example illustrating this in Example 3.1.

For comparison, we give also the moments of the limit in the diagonal case; these follow immediately from the standard formulas for the moments of Gamma and Beta distributions. Note that in cases (i) and (iii), the moments are infinite outside the indicated ranges.

**Theorem 1.9.** *The moments of $W$ in Theorem 1.4 ($\gamma = 0$) are given by:*

(i) *If* $\alpha < \delta$ *then*

$$\mathbb{E}\, W^r = \delta^r \frac{\Gamma(x_0/\alpha + r)\Gamma(y_0/\delta - r\alpha/\delta)}{\Gamma(x_0/\alpha)\Gamma(y_0/\delta)}, \qquad -x_0/\alpha < r < y_0/\alpha.$$

(ii) *If* $\alpha = \delta$*, then* (1.15) *holds.*

(iii) *If* $\alpha > \delta$ *then*

$$\mathbb{E}\, W^r = \delta^r \frac{\Gamma(x_0/\alpha - r\delta/\alpha)\Gamma(y_0/\delta + r)}{\Gamma(x_0/\alpha)\Gamma(y_0/\delta)}, \qquad -y_0/\delta < r < x_0/\delta.$$

**Remark 1.10.** In the case of an irreducible replacement matrix, it is well-known that the type of the asymptotics depends on the relation between the eigenvalues of the replacement matrix, and in particular on whether the largest eigenvalue is at least twice the real part of any other eigenvalue (in which case we have asymptotic normality), see [2] and [14]. For two colours we have two eigenvalues $\lambda_1 \geq \lambda_2$, both real, and the three cases are $\lambda_2 < \frac{1}{2}\lambda_1$, $\lambda_2 = \frac{1}{2}\lambda_1$, and $\lambda_2 > \frac{1}{2}\lambda_1$; in the two first cases we have asymptotic normality (with a log factor in the asymptotic variance in the second case), but not in the third.

In the triangular case, the eigenvalues are simply the diagonal elements $\alpha$ and $\delta$, so $\lambda_1 = \max(\alpha, \delta)$ and $\lambda_2 = \min(\alpha, \delta)$. We thus see that for $\alpha < \delta$ and $\gamma > 0$ (Theorem 1.3(i)(ii)(iii)), we have the same behaviour as in the irreducible case, while there are several differences when $\alpha > \delta$ or $\gamma = 0$; in particular, there is no normality when $\delta < \frac{1}{2}\alpha$.

Indeed, the setting in [14] is somewhat more general than the irreducible case (we really need only that the largest eigenvalue is simple and has a strictly positive left eigenvector), and it is easily verified that a triangular urn with $\gamma, \delta > 0$ and $\alpha < \delta$ satisfies the conditions in [14]. In particular, Theorems 3.22–3.24 in [14] imply our Theorem 1.3(i)(ii)(iii), except for the explicit (but complicated) description of the limit in (iii); the variances in (1.3) and (1.5) can easily be computed as in Example 7.2 in [14], with minor modifications because now $\beta = 0$.

As discussed in [14], an explanation for the asymptotic normality in the irreducible case is that when $\lambda_2 \leq \frac{1}{2}\lambda_1$, the initial value and the results of the first draws have a negligible effect in the long run. The composition is thus effectively determined by the outcome of the large number of later draws, each having a negligible effect, which is a typical situation for asymptotic normality. On the other hand, if $\lambda_2 > \frac{1}{2}\lambda_1$, then imbalances caused by the first random draws magnify at a sufficient rate to remain important also for large $n$; indeed, the composition even after a long time is essentially determined by the results of the first few draws.

In a triangular urn, the same mechanism works if $\delta > \alpha$, so that the largest eigenvalue is $\delta$, which is connected to the white balls. On the other hand, if $\alpha > \delta$, the largest eigenvalue is connected to the black balls. Since $\beta = 0$, they do not communicate with the white balls and there is no similar smooting effect on the number of white balls caused by the large number of black draws. Hence the mechanism fails, which explains why there is no normal regime for $\delta < \frac{1}{2}\alpha$. Similarly, there is no such smoothing mechanism at all in the diagonal case.

In this connection, note that in the cases Theorem 1.3(i)(ii) (where we have asymptotic normality), the limit distribution does not depend on the initial composition $(x_0, y_0)$, while it does depend on it in the other cases, see (7.4) and Theorems 8.5 and 1.7. In fact, even the existence of mean and variance may depend on $x_0$ and $y_0$, see Theorem 8.2.

**Remark 1.11.** In the definition of the generalized Pólya urn and the 2-type branching process above, it is implicit that $\alpha, \beta, \gamma, \delta$ are integers, since we interpret them as numbers. However, this restriction is really not necessary; we may let $\alpha, \beta, \gamma, \delta$ be *real* numbers [15]. The urn process then can be defined as a Markov process with state space $\{(x, y) \in \mathbb{R}^2 : x, y \geq 0 \text{ and } x + y > 0\}$; the transitions are from $(x, y)$ to $(x + \alpha, y + \beta)$ with probability $x/(x + y)$ and from $(x, y)$ to $(x + \gamma, y + \delta)$ with probability $y/(x + y)$. The branching process is defined with the same state space; if it reaches (or starts in) a state $(x, y)$, it waits a random time with an $\mathrm{Exp}(1/(x+y))$ distribution and then jumps according to the same transition probabilities. (In other words, the two possible transitions from a state $(x, y)$ have intensities $x$ and $y$, respectively.) Note that this "branching proceess" has a property corresponding to the independence of the family histories of the descendants of different individuals in a true branching process: if $(x, y) = (x_1, y_1) + (x_2, y_2)$, then the process started at $(x, y)$ has the same distribution as the sum of two independent processes started at $(x_1, y_1)$ and $(x_2, y_2)$.

In order for these definitions to be legitimate, we have to assume that $\alpha, \beta, \gamma, \delta$ are such that we never will leave the state space defined above (which would mean that we are required to remove a ball from the urn that does not exist). Such urns are called *tenable*. There is no problem if $\alpha, \beta, \gamma, \delta \geq 0$, but we may also allow negative numbers under certain conditions, for example $\alpha < 0$, $\beta, \gamma, \delta \geq 0$ such that $\gamma$ and $x_0$ are multiples of $|\alpha|$, and $y_0 > 0$ (for example in Theorem 1.5(i)).

All our results and proofs (with proper interpretations) are valid in this generality, although we for simplicity sometimes phrase the proofs as if we counted balls.

One advantage of the extension to real $\alpha, \beta, \gamma, \delta$ is that a homogeneity becomes evident: For any $\lambda > 0$, the generalized Pólya urn with parameters $\lambda\alpha, \lambda\beta, \lambda\gamma, \lambda\delta$ and initial value $(\lambda x_0, \lambda y_0)$ is $\lambda(X_n, Y_n)$; this explains why the parameters appear as ratios such as $\alpha/\delta$ in our results. (One might use this and assume e.g. $\delta = 1$ in some proofs below in order to simplify the notation.)

**Remark 1.12.** Another extension that is useful in some applications is to let the different colours have different *weights* (or *activities*), say $\lambda$ and $\mu$: we now assume that balls are drawn with probabilities proportional to their weights. (This makes sense for any positive real weights. In the branching process, this means that particles of different types have different expected life times, see [2] and [14].) We can reduce this to the extension in Remark 1.11, since $(\lambda X_n, \mu Y_n)$ is a process of that type with replacement matrix $\begin{pmatrix} \lambda\alpha & \mu\beta \\ \lambda\gamma & \mu\delta \end{pmatrix}$. Hence all our results apply to the weighted case too, with straightforward modifications.

**Remark 1.13.** Many authors have studied generalized Pólya urns under the simplifying assumption that a fixed number of balls is added each time, i.e. that the replacement matrix has constant row sums. (In our triangular 2-colour case, this is $\alpha = \gamma + \delta$.) Such urns are called *balanced*. This means that the total number of balls is deterministic, which is very useful in some methods, but not needed in ours (except in some special results such as Theorem 1.7). Note that in our case, constant row sums can appear only in Theorem 1.3(v) and Theorem 1.4(ii).

Kotz, Mahmoud and Robert [19] give exact formulas and some asymptotics for 2-type urns. They study two triangular examples in detail, and comment that the case $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$ gives asymptotics "of an essentially different character" than cases with constant row sums. From our point of view, as expressed in Theorem 1.3, this difference is due to the relation between the eigenvalues (i.e., the diagonal elements $\alpha$ and $\delta$) rather than having equal row sums or not.

**Remark 1.14.** In [14] too, limit result for urns were proved using the embedding in a multi-type branching process. However, the method there is quite different: a functional limit theorem is proved for the branching process and the result studied at the (random) time of the $n$:th death gives limit results for the urn. In this paper, we instead use the embedding to derive an exact formula for a generating function (not for $X_n$ or $Y_n$, but for a related quantity; see Theorem 7.1), from which our results follow by traditional methods. We do not prove corresponding limit results for the branching process, although they undoubtedly exist. (When $\alpha < \delta$, such results hold by [14], see Remark 1.10.)

By [14, Theorem 3.31], the normal limit results in Theorem 1.3(i)(ii) can be extended to functional limit results, describing how $Y_n$ varies with $n$ for large $n$; we omit the details. Similarly, [14, Theorem 3.24] shows that in Theorem 1.3(iii), we have convergence a.s. to some $W$ (necessarily with the distribution described above). The same is true in Theorem 1.3(iv) by [20, Theorem 2.3], and in the diagonal case Theorem 1.4, as seen by the proof in Section 11 or (when $\alpha \neq \delta$) by [20, Theorem 2.2]. We conjecture that the same is true in Theorem 1.3(v), but this seems to remain an open problem.

Our results suggest two other problems:

**Problem 1.15.** *Find better descriptions of the limits $W$ and $Z$.*

**Problem 1.16.** *Extend the results of this paper to three or more colours!*

It is possible that the methods of this paper can be used, perhaps together with an induction on the number of colours, but certainly a non-trival amount of further work would be needed.

The proofs of the results above (and some related results) are given in Sections 4–12. In Section 13, we show how a version of our method gives some exact results in a diagonal example.

Note that Flajolet, Gabarró and Pekari [10], see also [11], recently have given a detailed study of two-colour urns, including triangular cases, by different methods; the reader is encouraged to compare the papers. See in particular [10, Section 5.5] and [11], which treat the case $\alpha = \gamma + \delta$, $\gamma = 1$, and find a more refined version (a local limit theorem) of (1.8) and Theorem 1.8 (when $x_0 = 0$, $y_0 = 1$) for that case.

A completely different set of limit theorems for branching processes (and urns) with triangular replacement matrices are given by Drmota and Vatutin [6]; they study a different problem (the number of particles of different colours ever born for a process that dies out, conditioned on the sum over all colours) and obtain several different asymptotic distributions in different cases.

**Acknowledgements.** Part of this research was done during a visit to Université de Versailles Saint-Quentin, Versailles, France. I thank Philippe Flajolet, Nicolas Pouyanne and Vincent Puyhaubert for interesting discussions and for sharing unpublished results with me, which have inspired some of the results above. In particular, I would never have even attempted to find the moments in Theorem 1.7, had I not known the result from Puyhaubert's work. I further thank a referee for helpful remarks.

## 2. Degenerate cases

For completeness, we discuss briefly some degenerate cases.

If $y_0 = 0$, i.e. we start with black balls only, then we never will have any white balls, and the urn is deterministic: $(X_n, Y_n) = (x_0 + \alpha n, 0)$.

The case $x_0 = \gamma = 0$ is similar, with only white balls.

If $\delta < 0$, the white balls will eventually disappear. This is obvious for the branching process, and thus it is a.s. true for the urn model too. Hence, for all $n$ larger than some random value, using Lemma 1.2, $(X_n, Y_n) = (\alpha n + x_0 + (\gamma - \alpha)y_0/|\delta|, 0)$.

The case $\alpha < 0$, $\gamma = 0$ is similar.

If $\delta = 0$, then $Y_n = y_0$ is constant. If further $\alpha = \gamma = 0$, then nothing at all is ever added and the urn is utterly trivial.

If $\alpha = \gamma$, then $X_n = x_0 + \alpha n$ is deterministic.

If $\alpha = \gamma = 0$ and $\delta > 0$, $y_0 > 0$, we may interchange the colours and obtain a normal limit for $Y_n$ from Theorem 1.5(iii).

In all other cases with $\beta = 0$, Theorems 1.3–1.5 yield non-degenerate limits.

## 3. Examples

We give three examples that illustrate what we can and cannot say about the limit distributions, hoping that they may serve as inspiration for further research.

**Example 3.1.** The simplest non-normal case is $\left(\begin{smallmatrix}2&0\\1&1\end{smallmatrix}\right)$, i.e. $\alpha=2$, $\gamma=\delta=1$. We assume $y_0>0$, and Theorem 1.3(v) yields $n^{-1/2}Y_n \xrightarrow{\mathrm{d}} W$.

If $x_0=0$ or 1, Theorem 1.8 applies. In this case ($\delta/\alpha=1/2$), the Mittag-Leffler density $f(x)=\pi^{-1/2}e^{-x^2/4}$, and thus $W$ has density $cx^{y_0-x_0}e^{-x^2/4}$, $x>0$. (In this example $c$ stands for unspecified positive constants.) (Hence, $W/\sqrt{2}$ has the Chi distribution $\chi(y_0+1-x_0)$, i.e. $W^2\sim\Gamma((y_0+1-x_0)/2,4)$.)

In particular, if $x_0=0$ and $y_0=1$, $W$ has a Rayleigh distribution [10, Corollary 12], [11]. If $x_0=y_0=1$, $W\overset{\mathrm{d}}{=}|U|$ with $U\sim N(0,2)$.

If $x_0>1$, the description of the limit $W$ is less simple. If $x_0=\alpha=2$, the moment in Theorem 1.7 differs from the case $x_0=0$ by a factor $y_0/(y_0+r)$, which is the $r$:th moment of a random variable $V$ with density $g(x):=y_0x^{y_0-1}$ on $(0,1)$, i.e. $V\sim\mathrm{Beta}(y_0,1)$. Consequently $W\overset{\mathrm{d}}{=}W_0V$, where $W_0$ is as in the case $x_0=0$, i.e. has a density $h_0(x):=cx^{y_0}e^{-x^2/4}$, $x>0$, and $V$ has the density $g(x)$, with $U$ and $V$ independent. It follows that $W$ has density $\int_x^\infty h_0(y)g(x/y)y^{-1}\,dy=cx^{y_0-1}\int_x^\infty e^{-y^2/4}\,dy$, $x>0$.

In the same way, we can find the density for every integer $x_0$, but the formulas become more complicated for larger $x_0$.

**Example 3.2.** A related, non-balanced, case is $\left(\begin{smallmatrix}2&0\\3&1\end{smallmatrix}\right)$, i.e. $\alpha=2$, $\gamma=3$, $\delta=1$. Assume for simplicity $x_0=0$ and $y_0=1$. Again Theorem 1.3(v) yields $n^{-1/2}Y_n\xrightarrow{\mathrm{d}}W$, but now we have no simple description of $W$.

The moments can be calculated from Theorem 1.6 in this case too. The inner integral in Theorem 1.6 is $2(1+2x)x^{-1/2}(1+x)^{-1/2}-2x^{-1/2}$, and thus

$$\mathbb{E}\,W^r=\frac{r!}{\Gamma(r/2)}\int_0^\infty(1+x)^{(r+1)/2}(1+2x)^{-r-1}x^{r/2-1}\,dx. \tag{3.1}$$

This is easily calculated by `Maple` for small $r$; we find $\mathbb{E}\,W=\frac{3}{8}\sqrt{2\pi}$, $\mathbb{E}\,W^2=\frac{7}{8}+\frac{3}{16}\sqrt{2}\ln(\sqrt{2}+1)$, $\mathbb{E}\,W^3=\frac{39}{64}\sqrt{2\pi}$, $\mathbb{E}\,W^4=\frac{219}{128}+\frac{135}{256}\sqrt{2}\ln(\sqrt{2}+1)$, $\mathbb{E}\,W^5=\frac{825}{512}\sqrt{2\pi}$, …, but no simple formula is evident.

It is easily seen that all odd moments are rational multiples of $\sqrt{2\pi}$; if $r=2k-1$, expand $(1+x)^k=((1+2x)-x)^k$ in (3.1) by the binomial theorem, let $y=2x$ and apply (4.2) below. This gives an explicit formula for the odd moments, but we do not find any simple form for them, and leave further investigations to the reader.

**Example 3.3.** Consider $\left(\begin{smallmatrix}1&0\\1&1\end{smallmatrix}\right)$, i.e. $\alpha=\gamma=\delta=1$, a case studied in [19]. By Theorem 1.3(iv),

$$\frac{\ln^2 n}{n}Y_n-\ln n-\ln\ln n\xrightarrow{\mathrm{d}}W=-Z.$$

It follows easily from e.g. (8.5) and (8.8) below that

$$\mathbb{E}\, e^{-\tau Z} = \frac{1}{\Gamma(y_0)} \int_0^\infty (\tau + x)^{\tau - x_0} x^{x_0 + y_0 - 1} e^{-x}\, dx, \qquad \tau \geq 0.$$

Moments of $Z$ and $W$ can be calculated from this formula. For example, cf. Theorem 8.2,

$$- \mathbb{E}\, W = \mathbb{E}\, Z = \frac{x_0}{y_0 - 1} - \frac{\Gamma'(y_0)}{\Gamma(y_0)}$$

if $x_0 = 0$ or $y_0 > 1$ (and $\infty$ otherwise).

## 4. Preliminaries

**Infinitely divisible distributions.** For the general theory of infinitely divisible distributions, we refer to a general probability book such as [9] or [17]. We will here only need a special case.

Suppose that $\nu$ is a measure on $\mathbb{R} \setminus \{0\}$ such that $\int x^2\, d\nu(x) < \infty$ and that $\lambda \in \mathbb{R}$. We say that a random variable $Y$ has an *infinitely divisible distribution with Lévy measure $\nu$ and mean $\lambda$* if its characteristic function is given by

$$\mathbb{E}\, e^{itY} = e^{it\lambda + \int (e^{itx} - 1 - itx)\, d\nu(x)}. \tag{4.1}$$

In this case, $\mathbb{E}\, Y = \lambda$ and $\operatorname{Var} Y = \int x^2\, d\nu(x)$.

In our cases, the Lévy measure is supported on $\{x > 0\}$; such infinitely divisible distribution are called *spectrally positive.* It is easily seen that in this case, $\mathbb{E}\, e^{-tY} < \infty$ for every $t > 0$, and that (4.1) extends to complex $t$ with $\operatorname{Im} t \geq 0$ (thus giving a formula for $\mathbb{E}\, e^{-tY}$ for $\operatorname{Re} t \geq 0$).

A well-known example is the Gamma distribution $\Gamma(p, a)$ (where $a, p > 0$). The characteristic function is $(1 - iat)^{-p}$, which is of the form (4.1) with $\lambda = ap$ and $d\nu(y) = py^{-1} e^{-a^{-1}y}\, dy$, $y > 0$ [9, XVII.3(d)].

**Some Beta integrals.** Recall that the Beta function is defined by $B(z, w) = \Gamma(z)\Gamma(w)/\Gamma(z + w)$. We have the well-known and classical formula [1, 6.2.1]

$$\int_0^\infty (1 + x)^{-a} x^{b-1}\, dx = B(a - b, b) = \frac{\Gamma(a - b)\Gamma(b)}{\Gamma(a)}, \qquad a > b > 0. \tag{4.2}$$

For $b < 0$, we have the following related results, which perhaps are less well-known, although they too undoubtedly have been known for a long time. (The formulas extend to complex $a$ and $b$, with real parts satisfying the corresponding inequalities, but we only need real parameters in this paper.)

**Lemma 4.1.** (i) *If $-1 < b < 0$ and $a > b$, then*

$$\int_0^\infty \big((1 + x)^{-a} - 1\big) x^{b-1}\, dx = B(a - b, b) = \Gamma(a - b)\Gamma(b)/\Gamma(a).$$

(ii) *If $-2 < b < -1$ and $a > b$, then*

$$\int_0^\infty \big((1 + x)^{-a} - 1 + ax\big) x^{b-1}\, dx = B(a - b, b) = \Gamma(a - b)\Gamma(b)/\Gamma(a).$$

(iii) *If $a > -1$, then, with $\psi(z) = \Gamma'(z)/\Gamma(z)$,*

$$\int_0^1 \bigl((1+x)^{-a}-1+ax\bigr)x^{-2}\,dx+\int_1^\infty \bigl((1+x)^{-a}-1\bigr)x^{-2}\,dx = a\bigl(\psi(a+1)-\psi(2)\bigr).$$

*Proof.* (i) and (ii) follow from (4.2) by one and two integrations by parts, respectively; we omit the details. For (iii), note that if $0 < \varepsilon < 1$, then, using (i),

$$\begin{aligned}
\int_0^1 &\bigl((1+x)^{-a}-1+ax\bigr)x^{\varepsilon-2}\,dx+\int_1^\infty \bigl((1+x)^{-a}-1\bigr)x^{\varepsilon-2}\,dx \\
&= \int_0^\infty \bigl((1+x)^{-a}-1\bigr)x^{\varepsilon-2}\,dx + \int_0^1 ax^{\varepsilon-1}\,dx \\
&= \frac{\Gamma(a-\varepsilon+1)\Gamma(\varepsilon-1)}{\Gamma(a)} + \frac{a}{\varepsilon} = \frac{\Gamma(a-\varepsilon+1)\Gamma(\varepsilon+1)+(\varepsilon-1)a\Gamma(a)}{\varepsilon(\varepsilon-1)\Gamma(a)}.
\end{aligned}$$

As $\varepsilon \to 0$, this converges, by l'Hôpital's rule, to

$$\frac{-\Gamma'(a+1)+\Gamma(a+1)\Gamma'(1)+a\Gamma(a)}{-\Gamma(a)} = a\bigl(\psi(a+1)-\psi(1)-1\bigr),$$

and (iii) follows because $\psi(2) = \psi(1)+1$. □

## 5. The case of one colour

As a preparation, we start by studying the one-colour case with replacement matrix $(\alpha)$. The urn process is trivial; we add $\alpha$ new balls each time, so $X_n = x_0 + \alpha n$. The corresponding continuous time branching process is more interesting; it is a pure birth process where each particle has an Exp(1) lifetime and then splits into $\alpha+1$ particles. For $\alpha = 1$, this is known as the *Yule process*. The generating function is easily determined, see e.g. [3, Remark III.5.1].

**Lemma 5.1.** *Let $\mathcal{X}(t)$ be a generalized Yule process, where each particle splits into $\alpha+1$ after an* Exp(1) *lifetime, and suppose $\mathcal{X}(0) = x_0$. Then, for each $t \ge 0$ and $0 < |z| \le 1$, $\mathbb{E}\, z^{\mathcal{X}(t)} = \Phi(z;t)^{x_0}$, where*

$$\Phi(z;t) := z\bigl((1-z^\alpha)e^{\alpha t}+z^\alpha\bigr)^{-1/\alpha} = \bigl(1+(z^{-\alpha}-1)e^{\alpha t}\bigr)^{-1/\alpha} \tag{5.1}$$

*if $\alpha \ne 0$, while $\Phi(z;t) := z$ if $\alpha = 0$.*

**Remark 5.2.** In accordance with Remark 1.11, $\alpha$ and $x_0$ may here be real numbers, assuming $x_0 \ge 0$ and either $\alpha \ge 0$ or $\alpha < 0$ with $x_0$ a multiple of $|\alpha|$.

**Remark 5.3.** For, say, $0 < z \le 1$, there is no problem to interpret the fractional powers in (5.1). Later, however, we want to take $z$ complex, and we then have to choose the right branches. There is no serious problem; we will take $z = e^{is}$ with $s$ real, and then there is a unique branch of (5.1) that is a continuous function of $s$ with value 1 for $s = 0$. We will use this interpretation of $\Phi(e^{is};t)$ without further comment, and Lemma 5.1 then gives the characteristic function of $\mathcal{X}(t)$. The same applies to $G_m(z)$ and $\Psi(z;t)$ in Theorem 6.1 below.

*Proof.* As said above, this is well-known, at least when $\alpha$ is an integer. For completeness we give a simple proof valid for all $\alpha \neq 0$ and $x_0 > 0$ (the cases $\alpha = 0$ or $x_0 = 0$ are trivial).

Define $p_k(t) := \mathbb{P}(\mathcal{X}(t) = k\alpha + x_0)$. Then

$$p_k'(t) = \big((k-1)\alpha + x_0\big)p_{k-1}(t) - (k\alpha + x_0)p_k(t),$$

with $p_{-1}(t) = 0$ and $p_k(0) = \delta_{k0}$. This system is solved by

$$p_k(t) = \frac{\Gamma(k + x_0/\alpha)}{\Gamma(x_0/\alpha)k!}e^{-x_0t}\big(1 - e^{-\alpha t}\big)^k, \tag{5.2}$$

and the result follows by

$$\begin{aligned}\mathbb{E}\, z^{\mathcal{X}(t)} &= \sum_{k=0}^{\infty} p_k(t) z^{k\alpha + x_0} = z^{x_0}e^{-x_0t}\sum_{k=0}^{\infty}\binom{-x_0/\alpha}{k}\big(-z^{\alpha}(1-e^{-\alpha t})\big)^k \\ &= z^{x_0}e^{-x_0t}\big(1 - z^{\alpha}(1-e^{-\alpha t})\big)^{-x_0/\alpha}.\end{aligned}$$

□

## 6. Stopping after $m$ white draws

Let $X^{(m)}$ be the number of black balls when $m$ white balls have been drawn, i.e. (assuming $\delta \neq 0$) $X_n$ for the smallest $n$ such that the number $Y_n$ of white balls is $y_0 + m\delta$, and let $G_m(z) := \mathbb{E}\, z^{X^{(m)}}$ be its probability generating function. We will give an explicit formula for $G_m(z)$. This will in the next section be used to derive a limit theorem for $X^{(m)}$.

**Theorem 6.1.** *Consider a generalized Pólya urn with triangular replacement matrix $\left(\begin{smallmatrix}\alpha & 0\\ \gamma & \delta\end{smallmatrix}\right)$ and initial composition $(x_0, y_0)$, with $\delta > 0$ and $y_0 > 0$. Then, for $0 < |z| \leq 1$,*

$$G_m(z) = \delta\frac{\Gamma(y_0/\delta + m)}{\Gamma(y_0/\delta)\Gamma(m)}z^{\gamma}\int_0^{\infty} e^{-y_0t}\Phi(z;t)^{x_0}\Psi(z;t)^{m-1}\,dt$$

*for any $m \geq 1$, where $\Phi(z;t)$ is given by* (5.1) *(and $\Phi(z;t) := z$ if $\alpha = 0$), and*

$$\Psi(z;t) := \delta\int_0^t e^{-\delta u}\Phi(z;u)^{\gamma}\,du. \tag{6.1}$$

**Remark 6.2.** Recall that non-integer parameters are possible, see Remark 1.11. See also Remark 5.3.

*Proof.* Let $P_{x,y}(t) := \mathbb{P}\big(\mathcal{X}(t) = x$ and $\mathcal{Y}(t) = y\big)$ and $\tilde{P}_{x,y} := \mathbb{P}\big(\mathcal{X}(t) = x$ and $\mathcal{Y}(t) = y$ for some $t \geq 0\big)$.

**Lemma 6.3.** *Assume $\delta > 0$, $\alpha \neq 0$ and $x, y \geq 0$ with $x + y > 0$. Then*

$$\tilde{P}_{x,y} = (x+y)\int_0^{\infty} P_{x,y}(t)\,dt.$$

*Proof.* If the composition $(x, y)$ occurs during the evolution, it remains for a period of time with an $\mathrm{Exp}(1/(x+y))$ distribution, and then never occurs again. Thus, the expected time the composition equals $(x, y)$ is $\tilde{P}_{x,y} \cdot 1/(x+y)$. Finally, by Fubini's theorem, this expected time equals

$$\mathbb{E}\int_0^\infty \mathbf{1}[\mathcal{X}(t) = x \text{ and } \mathcal{Y}(t) = y]\, dt = \int_0^\infty P_{x,y}(t)\, dt. \qquad \square$$

Now consider $X^{(m+1)}$. It equals $x$ if some white draw (necessarily the $(m+1)$:th) leads to $(x, y_0 + (m+1)\delta)$, i.e. if at some time we reach $(x-\gamma, y_0 + m\delta)$ and the next draw is white. Consequently, using Lemma 6.3 (with a minor modification if $\alpha = 0$),

$$\begin{aligned}\mathbb{P}(X^{(m+1)} = x) &= \tilde{P}_{x-\gamma, y_0+m\delta} \cdot \frac{y_0 + m\delta}{x - \gamma + y_0 + m\delta} \\ &= (y_0 + m\delta)\int_0^\infty P_{x-\gamma,y_0+m\delta}(t)\, dt.\end{aligned}$$

Thus, if we define

$$H_m(z;t) := \mathbb{E}\big(z^{\mathcal{X}(t)}; \mathcal{Y}(t) = y_0 + m\delta\big) = \sum_x z^x P_{x,y_0+m\delta}(t),$$

we have, for every $m \geq 0$,

$$\begin{aligned}G_{m+1}(x) &= \sum_x \mathbb{P}(X^{(m+1)} = x) z^x \\ &= (y_0 + m\delta) z^\gamma \int_0^\infty \sum_x P_{x-\gamma,y_0+m\delta}(t) z^{x-\gamma}\, dt \qquad (6.2)\\ &= (y_0 + m\delta) z^\gamma \int_0^\infty H_m(z;t)\, dt.\end{aligned}$$

**Lemma 6.4.** *Assume $\delta > 0$ and $y_0 > 0$. Then, for every $m \geq 0$,*

$$H_m(z;t) = \frac{\Gamma(y_0/\delta + m)}{\Gamma(y_0/\delta)\, m!} e^{-y_0 t} \Phi(z;t)^{x_0} \Psi(z;t)^m.$$

*Proof.* The idea is to look at the black and white balls separately, as two different (but dependent) processes. Suppose that $m$ white balls are drawn before time $t$, at times $t_1 < \cdots < t_m < t$. The $\gamma$ black balls added at $t_i$ increase by drawings of these balls and their offspring; the total size of their families at time $t$ has, by Lemma 5.1, the probability generating function $\Phi(z; t-t_i)^\gamma$. Considering also the offspring of the original $x_0$ black balls, and noting that different groups evolve independently (by the additive property of the branching process), we see that the probability generating function of the number of black balls at time $t$, conditioned on the number $m$ and times $t_1, \dots, t_m$ of white draws, is

$$\Phi(z;t)^{x_0} \prod_{i=1}^m \Phi(z; t - t_i)^\gamma. \qquad (6.3)$$

Let us now ignore the black balls. The white balls form a pure birth process, and the joint density of the first $m$ jumps is

$$y_0 e^{-y_0 t_1}(y_0+\delta)e^{-(y_0+\delta)(t_2-t_1)}\cdots(y_0+(m-1)\delta)e^{-(y_0+(m-1)\delta)(t_m-t_{m-1})}$$
$$= \delta^m \frac{\Gamma(y_0/\delta+m)}{\Gamma(y_0/\delta)} e^{\delta t_1+\cdots+\delta t_{m-1}-(y_0+(m-1)\delta)t_m}$$

on $0 < t_1 < \cdots < t_m$.

Restricting to $t_m < t$ and multiplying by $e^{-(y_0+m\delta)(t-t_m)}$, the probability of no further white draw before $t$, we obtain the density of white draws at $t_1, \dots, t_m$, and only then before $t$. Multiplying by (6.3) and integrating yields

$$H_m(z,t) = \int\cdots\int_{0<t_1<\cdots<t_m<t} \delta^m \frac{\Gamma(y_0/\delta+m)}{\Gamma(y_0/\delta)} e^{\delta(t_1+\cdots+t_m)-(y_0+m\delta)t}$$
$$\cdot \Phi(z,t)^{x_0} \prod_{i=1}^m \Phi(z;t-t_i)^\gamma \, dt_1\cdots dt_m.$$

Since the integrand is symmetric in $t_1, \dots, t_m$, we may integrate over $[0,t]^m$ instead and divide by $m!$; this integral factors and we obtain

$$H_m(z,t) = \frac{\Gamma(y_0/\delta+m)}{m!\,\Gamma(y_0/\delta)} e^{-y_0 t}\Phi(z,t)^{x_0}\Bigl(\delta\int_0^t e^{\delta(t_1-t)}\Phi(z;t-t_1)^\gamma \, dt_1\Bigr)^m.$$

The lemma follows by (6.1) and the change of variables $u = t_1 - t$. ☐

Theorem 6.1 now follows from (6.2) and Lemma 6.4, if we replace $m$ by $m-1$. ☐

## 7. A LIMIT THEOREM FOR $X^{(m)}$

We use the exact formula in Theorem 6.1 to obtain the asymptotic distribution of $X^{(m)}$ as $m \to \infty$. We have to distinguish between the same five cases as in Theorem 1.3. In two cases we find a normal limit distribution. In the other cases, we give an explicit but complicated formula for the characteristic function of the limit. In Section 8 we will derive other expressions for the characteristic function and moment generating function, and we will compute the mean and variance of the limit. In some special cases, we find much simpler results (in particular, a stable distribution in Theorem 8.6); in general all our expressions are complicated and perhaps mainly useful to show the existence of a limit. (It is quite possible that there might exist a much simpler formula than the ones we have found, or some other simple description of the limit distribution that we have failed to see.)

Define for convenience, (using the principal branch)

$$R_1(z) = (1+z)^{-\gamma/\alpha} - 1, \tag{7.1}$$
$$R_2(z) = (1+z)^{-\gamma/\alpha} - 1 + (\gamma/\alpha)z = R_1(z) + (\gamma/\alpha)z. \tag{7.2}$$

**Theorem 7.1.** *Consider a generalized Pólya urn with triangular replacement matrix* $\left(\begin{smallmatrix}\alpha & 0\\ \gamma & \delta\end{smallmatrix}\right)$ *and initial composition* $(x_0, y_0)$, *with* $\delta > 0$ *and* $y_0 > 0$, *and let* $m \to \infty$.

(i) *If* $\alpha < \delta/2$, *then*

$$m^{-1/2}\Big(X^{(m)} - \frac{\gamma}{1-\alpha/\delta}m\Big) \xrightarrow{\mathrm{d}} N(0, \sigma_1^2),$$

*where*

$$\sigma_1^2 = \frac{\alpha^2\gamma\delta(\delta+\gamma-\alpha)}{(\delta-2\alpha)(\delta-\alpha)^2}. \tag{7.3}$$

(ii) *If* $\alpha = \delta/2$, *then*

$$(m\ln m)^{-1/2}\Big(X^{(m)} - \frac{\gamma}{1-\alpha/\delta}m\Big) \xrightarrow{\mathrm{d}} N\big(0, \gamma(\gamma+\alpha)\big).$$

(iii) *If* $\delta/2 < \alpha < \delta$, *then*

$$m^{-\alpha/\delta}\Big(X^{(m)} - \frac{\gamma}{1-\alpha/\delta}m\Big) \xrightarrow{\mathrm{d}} Z,$$

*where* $Z$ *is a random variable with the characteristic function*

$$\mathbb{E}\, e^{itZ} = \frac{1}{\Gamma(y_0/\delta)}\int_0^\infty (1 - i\alpha t x^{-\alpha/\delta})^{-x_0/\alpha} e^{g(t,x)} x^{y_0/\delta-1}\,dx, \tag{7.4}$$

*with*

$$g(t,x) := \int_x^\infty R_2(-i\alpha t u^{-\alpha/\delta})\,du - i\frac{\gamma}{1-\alpha/\delta}t x^{1-\alpha/\delta} - x. \tag{7.5}$$

(iv) *If* $\alpha = \delta$, *then*

$$m^{-1}\Big(X^{(m)} - \gamma m \ln m\Big) \xrightarrow{\mathrm{d}} Z,$$

*where* $Z$ *is a random variable with the characteristic function* (7.4), *where now*

$$g(t,x) := \int_x^\infty R_2(-i\alpha t u^{-1})\,du - i\gamma t \ln x - x. \tag{7.6}$$

(v) *If* $\alpha > \delta$, *then*

$$m^{-\alpha/\delta}X^{(m)} \xrightarrow{\mathrm{d}} Z,$$

*where* $Z$ *is a random variable with the characteristic function* (7.4), *with* $g$ *given by* (7.5) *or, equivalently,*

$$g(t,x) := \int_x^\infty R_1(-i\alpha t u^{-\alpha/\delta})\,du - x. \tag{7.7}$$

*Proof.* In order to prove the five parts together, we begin with a lemma. Let $\Phi$ and $\Psi$ be as above, see (5.1), (6.1) and Remark 5.3.

**Lemma 7.2.** *Let $a_m$, $b_m$ and $\rho$ be real numbers with $a_m \to 0$ and $b_m/m^{1/2} \to \rho$ as $m \to \infty$. Define, for $0 < x < m$, $t_m(x) := (\ln m - \ln x)/\delta = \delta^{-1}\ln(m/x)$ and $\Phi^*_m(s,x) := \Phi\bigl(e^{isa_m}; t_m(x)\bigr)$, $\Psi^*_m(s,x) := \Psi\bigl(e^{isa_m}; t_m(x)\bigr)$. Suppose that for some functions $f, g : \mathbb{R}\times\mathbb{R}_+ \to \mathbb{R}$ and every $s \in \mathbb{R}$ and $x > 0$, as $m \to \infty$,*

$$\Phi^*_m(s,x) = f(s,x) + o(1), \tag{7.8}$$

$$\Psi^*_m(s,x) = 1 + \frac{isb_m}{m} + \frac{g(s,x)}{m} + o\Bigl(\frac{1}{m}\Bigr). \tag{7.9}$$

*Suppose further that $f(s,x)$ and $g(s,x)$ are continuous in $s$ for every fixed $x$. Then*

$$a_m X^{(m)} - b_m \xrightarrow{\mathrm{d}} Z$$

*for a random variable $Z$ with the characteristic function*

$$\mathbb{E}\, e^{isZ} = \frac{1}{\Gamma(y_0/\delta)} \int_0^\infty f(s,x)^{x_0} e^{g(s,x)+\rho^2 s^2/2} x^{y_0/\delta - 1}\, dx. \tag{7.10}$$

*Proof.* The characteristic function of $a_m X^{(m)} - b_m$ is by Theorem 6.1 and the change of variable $t = t_m(x)$,

$$\begin{aligned}\mathbb{E}\, e^{is(a_m X^{(m)} - b_m)} &= e^{-isb_m} G_m(e^{isa_m}) \\ &= \frac{\Gamma(y_0/\delta + m)}{\Gamma(y_0/\delta)\Gamma(m)} m^{-y_0/\delta} e^{isa_m\gamma} \int_0^m e^{-isb_m} x^{y_0/\delta} \Phi^*_m(s,x)^{x_0} \Psi^*_m(s,x)^{m-1} \frac{dx}{x}.\end{aligned} \tag{7.11}$$

As $m \to \infty$, the factor in front of the integral tends to $1/\Gamma(y_0/\delta)$, and, by assumption, $\Phi^*_m(s,x) \to f(s,x)$, while $\Psi^*_m(s,x) - 1 = O(m^{-1/2})$ and thus

$$\begin{aligned}(m-1)\ln \Psi^*_m(s,x) &= m(\Psi^*_m - 1) - m\tfrac{1}{2}(\Psi^*_m - 1)^2 + o(1) \\ &= isb_m + g(s,x) + s^2\rho^2/2 + o(1)\end{aligned}$$

and

$$e^{-isb_m} \Psi^*_m(s,x)^{m-1} \to e^{g(s,x) + s^2\rho^2/2}.$$

Consequently, the integrand in (7.11) tends to the integrand in (7.10). Moreover, $\Phi^*_m(s,x)$ is a characteristic function by Lemma 5.1, so it is bounded by 1. This also implies, by (6.1),

$$|\Psi(e^{is}; t)| \le \delta \int_0^t e^{-\delta u} |\Phi(e^{is}; u)|^\gamma\, du \le \delta \int_0^t e^{-\delta u}\, du = 1 - e^{-\delta t},$$

so for $m \ge 2$,

$$|\Psi^*_m(s,x)|^{m-1} \le (1 - e^{-\delta t_m(x)})^{m-1} = \Bigl(1 - \frac{x}{m}\Bigr)^{m-1} \le e^{-(m-1)x/m} \le e^{-x/2}.$$

Consequently, the integrand in (7.11) is bounded by $x^{y_0/\delta - 1} e^{-x/2}$; hence also its limit, the integrand in (7.10), is bounded by the same function. We can thus apply dominated convergence to (7.11), which shows that $\mathbb{E}\, e^{is(a_m X^{(m)} - b_m)}$ converges as $m \to \infty$, for every fixed real $s$, to the right hand side of (7.10). Further, dominated convergence again shows that this right hand side is a continuous function of $s$. The lemma follows by the continuity theorem for characteristic functions [17, Theorem 5.22]. $\square$

Note that, by (5.1) and (6.1), setting $y = e^{-\delta u}$, when $\alpha \neq 0$,

$$\Phi_m^*(s,x) = \Phi\bigl(e^{isa_m}; t_m(x)\bigr) = \bigl(1 + (e^{-i\alpha s a_m} - 1)(m/x)^{\alpha/\delta}\bigr)^{-1/\alpha}, \tag{7.12}$$

$$\begin{aligned}\Psi_m^*(s,x) &= \Psi\bigl(e^{isa_m}; t_m(x)\bigr) = \delta \int_0^{t_m(x)} e^{-\delta u} \Phi(e^{isa_m}; u)^\gamma \, du \\ &= \int_{x/m}^1 \Phi\Bigl(e^{isa_m}; -\frac{\ln y}{\delta}\Bigr)^\gamma dy = \int_{x/m}^1 \bigl(1 + (e^{-i\alpha s a_m} - 1) y^{-\alpha/\delta}\bigr)^{-\gamma/\alpha} dy;\end{aligned} \tag{7.13}$$

when $\alpha = 0$ we easily obtain $\Phi_m^*(s,x) = e^{isa_m}$ and $\Psi_m^*(s,x) = e^{i\gamma s a_m}(1 - x/m)$.

We will apply Lemma 7.2, with the following choices of $a_m$ and $b_m$ in the five different cases:

| | | | |
|---|---|---|---|
| (i) | $\alpha < \delta/2$: | $a_m = m^{-1/2}$, | $b_m = \gamma(1-\alpha/\delta)^{-1} m^{1/2}$; |
| (ii) | $\alpha = \delta/2$: | $a_m = (m \ln m)^{-1/2}$, | $b_m = 2\gamma (m/\ln m)^{1/2}$; |
| (iii) | $\delta/2 < \alpha < \delta$: | $a_m = m^{-\alpha/\delta}$, | $b_m = \gamma(1-\alpha/\delta)^{-1} m^{1-\alpha/\delta}$; |
| (iv) | $\alpha = \delta$: | $a_m = m^{-\alpha/\delta} = m^{-1}$, | $b_m = \gamma \ln m$; |
| (v) | $\alpha > \delta$: | $a_m = m^{-\alpha/\delta}$, | $b_m = 0$. |

Note that $\rho = \gamma(1-\alpha/\delta)^{-1}$ in case (i), and $\rho = 0$ in cases (ii)–(v).

In cases (i) and (ii),

$$(e^{-i\alpha s a_m} - 1)(m/x)^{\alpha/\delta} = O(a_m m^{\alpha/\delta}) \to 0,$$

and thus (7.12) shows that (7.8) holds with $f(s,x) = 1$.

In cases (iii), (iv), (v),

$$(e^{-i\alpha s a_m} - 1)(m/x)^{\alpha/\delta} \to -i\alpha s x^{-\alpha/\delta},$$

and hence, by (7.12), (7.8) holds with

$$f(s,x) = (1 - i\alpha s x^{-\alpha/\delta})^{-1/\alpha}. \tag{7.14}$$

Next, in cases (i) and (ii), for $x/m \le y \le 1$,

$$(e^{-i\alpha s a_m} - 1) y^{-\alpha/\delta} = O\bigl(a_m(1 + m^{\alpha/\delta})\bigr) \to 0.$$

If $\alpha \neq 0$, we make a Taylor expansion in (7.13) and obtain, as $m \to \infty$,

$$\begin{aligned}
\Psi_m^*(s,x) &= \int_{x/m}^1 \Bigl(1 - \frac{\gamma}{\alpha}(e^{-i\alpha s a_m} - 1)y^{-\alpha/\delta} \\
&\qquad + \frac{1}{2}\frac{\gamma}{\alpha}\bigl(\frac{\gamma}{\alpha}+1\bigr)(e^{-i\alpha s a_m} - 1)^2 y^{-2\alpha/\delta} + O(a_m^3 y^{-3\alpha/\delta})\Bigr)\,dy \\
&= 1 - \frac{x}{m} - \frac{\gamma}{\alpha}(e^{-i\alpha s a_m} - 1)\frac{1}{1-\alpha/\delta}\Bigl(1 - \Bigl(\frac{x}{m}\Bigr)^{1-\alpha/\delta}\Bigr) \\
&\qquad + \frac{1}{2}\frac{\gamma(\gamma+\alpha)}{\alpha^2}(e^{-i\alpha s a_m} - 1)^2 \int_{x/m}^1 y^{-2\alpha/\delta}\,dy + o(m^{-1}) \\
&= 1 - \frac{x}{m} + \frac{\gamma}{1-\alpha/\delta}\Bigl(isa_m + \alpha\frac{s^2 a_m^2}{2}\Bigr) \\
&\qquad - \gamma(\gamma+\alpha)\frac{s^2 a_m^2}{2}\int_{x/m}^1 y^{-2\alpha/\delta}\,dy + o(m^{-1}).
\end{aligned} \tag{7.15}$$

Hence, for case (i) ($\alpha \neq 0$),

$$m\bigl(\Psi_m^*(s,x) - 1\bigr) = -x + \frac{i\gamma s}{1-\alpha/\delta}m^{1/2} + \frac{\alpha\gamma}{1-\alpha/\delta}\frac{s^2}{2} - \frac{\gamma(\gamma+\alpha)}{1-2\alpha/\delta}\frac{s^2}{2} + o(1).$$

It is easily checked that this holds for $\alpha = 0$ too.

In case (i), thus (7.9) holds with

$$g(s,x) = -x + \Bigl(\frac{\alpha\gamma}{1-\alpha/\delta} - \frac{\gamma(\gamma+\alpha)}{1-2\alpha/\delta}\Bigr)\frac{s^2}{2}.$$

We obtain from (7.10), recalling $f(s,x) = 1$ in this case,

$$\mathbb{E}\, e^{isZ} = \frac{1}{\Gamma(y_0/\delta)}\int_0^\infty e^{-\sigma_1^2 s^2/2 - x} x^{y_0/\delta - 1}\,dx = e^{-\sigma_1^2 s^2/2} \tag{7.16}$$

with

$$\sigma_1^2 = \frac{\gamma(\gamma+\alpha)}{1-2\alpha/\delta} - \frac{\alpha\gamma}{1-\alpha/\delta} - \rho^2 = \frac{\gamma(\gamma+\alpha)}{1-2\alpha/\delta} - \frac{\alpha\gamma}{1-\alpha/\delta} - \frac{\gamma^2}{(1-\alpha/\delta)^2},$$

which is equivalent to (7.3). Hence $Z \sim N(0,\sigma_1^2)$. This proves case (i).

For case (ii), (7.15) yields

$$\Psi_m^*(s,x) = 1 - \frac{x}{m} + \frac{i\gamma s}{1-\alpha/\delta}a_m - \gamma(\gamma+\alpha)\frac{s^2}{2m\ln m}\int_{x/m}^1 y^{-1}\,dy + o(m^{-1})$$

and thus

$$m\bigl(\Psi_m^*(s,x) - 1\bigr) = -x + 2i\gamma s\Bigl(\frac{m}{\ln m}\Bigr)^{1/2} - \gamma(\gamma+\alpha)\frac{s^2}{2} + o(1).$$

Hence, (7.9) holds with $g(s,x) = -x - \gamma(\gamma+\alpha)s^2/2$. Thus (7.16) holds with $\sigma_1^2 = \gamma(\gamma+\alpha)$, so $Z \sim N(0,\gamma(\gamma+\alpha))$ and Theorem 7.1(ii) follows.

In cases (iii), (iv), (v), use (7.1) to rewrite (7.13) as

$$\begin{aligned}\Psi^*_m(s,x) &= \int_{x/m}^1 \Bigl(1 + R_1\bigl((e^{-i\alpha s a_m}-1)y^{-\alpha/\delta}\bigr)\Bigr)\,dy \\ &= 1 - \frac{x}{m} + \frac{1}{m}\int_x^m R_1\bigl((e^{-i\alpha s a_m}-1)m^{\alpha/\delta}t^{-\alpha/\delta}\bigr)\,dt,\end{aligned} \tag{7.17}$$

where $(e^{-i\alpha s a_m}-1)m^{\alpha/\delta}t^{-\alpha/\delta} \to -i\alpha s t^{-\alpha/\delta}$ as $m\to\infty$.

Now, $|R_1(z)| = O(z)$ when $z = O(1)$ and $\operatorname{Re} z > -\frac12$, say. Hence, the last integrand is $O(t^{-\alpha/\delta})$. In case (v), we thus can use dominated convergence, and find

$$m\bigl(\Psi^*_m(s,x)-1\bigr) \to -x + \int_x^\infty R_1(-i\alpha s t^{-\alpha/\delta})\,dt,$$

so (7.9) holds with

$$g(s,x) = -x + \int_x^\infty R_1(-i\alpha s t^{-\alpha/\delta})\,dt. \tag{7.18}$$

Lemma 7.2 now gives part (v) of Theorem 7.1; (7.4) follows from (7.10) and (7.14), and (7.18) gives (7.7), which (because $\alpha/\delta > 1$) implies (7.5).

In cases (iii) and (iv), rewrite (7.17) using (7.2) as

$$\begin{aligned}m(\Psi^*_m(s,x)-1) = -x &+ \int_x^m R_2\bigl((e^{-i\alpha s a_m}-1)m^{\alpha/\delta}t^{-\alpha/\delta}\bigr)\,dt \\ &- \frac{\gamma}{\alpha}\int_x^m \bigl(e^{-i\alpha s a_m}-1\bigr)m^{\alpha/\delta}t^{-\alpha/\delta}\,dt.\end{aligned}$$

Since $R_2(z) = O(|z|^2)$ for $z = O(1)$ and $\operatorname{Re} z > -1/2$, we can use dominated convergence in the first integral. Further, $(e^{-i\alpha s a_m}-1)m^{\alpha/\delta} = -i\alpha s + O(m^{-\alpha/\delta})$, and we obtain

$$m(\Psi^*_m(s,x)-1) = -x + \int_x^\infty R_2\bigl(-i\alpha s t^{-\alpha/\delta}\bigr)\,dt + i\gamma s\int_x^m t^{-\alpha/\delta}\,dt + o(1).$$

In case (iii), $\int_x^m t^{-\alpha/\delta}\,dt = (1-\alpha/\delta)^{-1}\bigl(m^{1-\alpha/\delta} - x^{1-\alpha/\delta}\bigr)$ and (7.9) holds with

$$g(s,x) = -x + \int_x^\infty R_2\bigl(-i\alpha s t^{-\alpha/\delta}\bigr)\,dt - \frac{i\gamma s}{1-\alpha/\delta}x^{1-\alpha/\delta}.$$

In case (iv), $\int_x^m t^{-\alpha/\delta}\,dt = \ln m - \ln x$, and (7.9) holds with

$$g(s,x) = -x + \int_x^\infty R_2\bigl(-i\alpha s t^{-\alpha/\delta}\bigr)\,dt - i\gamma s\ln x.$$

In both cases the result follows by Lemma 7.2. ∎

## 8. More on the limit $Z$

In this section we derive some properties and alternative descriptions of the distribution of the limit $Z$ in Theorem 7.1(iii)(iv)(v), including simpler formulas in special cases. Note that the limit distribution depends not only on the parameters $\alpha, \gamma, \delta$, but also on the initial values $x_0$ and $y_0$. We assume throughout the section that we are in cases (iii)–(v); thus $\alpha > \delta/2$, further $\delta > 0$ and $y_0 > 0$. We allow $\gamma = 0$ or $x_0 = 0$, but the case $\gamma = x_0 = 0$ is trivial, since then the urn never contains any black ball, so $X^{(m)} = 0$ and $Z = 0$.

The characteristic function (7.4) in Theorem 7.1(iii)(iv)(v) can be written

$$\mathbb{E}\, e^{itZ} = \Gamma(y_0/\delta)^{-1} \int_0^\infty \phi_x(t) x^{y_0/\delta - 1} e^{-x}\, dx,$$

where

$$\phi_x(t) := (1 - i\alpha t x^{-\alpha/\delta})^{-x_0/\alpha} e^{g(t,x)+x}. \tag{8.1}$$

**Theorem 8.1.** *Assume $\alpha > \delta/2 > 0$. For each $x > 0$, $\phi_x$ is the characteristic function of a spectrally positive infinitely divisible distribution $\mu_x$ with mean*

$$\lambda(x) = x_0 x^{-\alpha/\delta} - \gamma(1 - \alpha/\delta)^{-1} x^{1-\alpha/\delta} \tag{8.2}$$

*(replaced by $x_0 x^{-\alpha/\delta} - \gamma \ln x$ if $\alpha = \delta$), and Lévy measure $\nu_x$ with density, on $y > 0$,*

$$\frac{d\nu_x(y)}{dy} = \frac{x_0}{\alpha} e^{-\alpha^{-1} x^{\alpha/\delta} y} y^{-1} + \frac{\alpha^{-\gamma/\alpha}}{\Gamma(\gamma/\alpha)} y^{\gamma/\alpha - 1} \int_x^\infty u^{\gamma/\delta} e^{-\alpha^{-1} u^{\alpha/\delta} y}\, du. \tag{8.3}$$

*Hence, the distribution of $Z$ is the mixture $\Gamma(y_0/\delta)^{-1} \int_0^\infty x^{y_0/\delta - 1} e^{-x} \mu_x\, dx$ of infinitely divisible distributions. (For $\gamma = 0$ we interpret $1/\Gamma(0) = 0$, so the second term in* (8.3) *vanishes.)*

*Proof.* The first factor in (8.1) is the characteristic function of a Gamma distribution $\Gamma(x_0/\alpha, \alpha x^{-\alpha/\delta})$, which by Section 4 has mean $x_0 x^{-\alpha/\delta}$ and Lévy measure $\frac{x_0}{\alpha} y^{-1} e^{-\alpha^{-1} x^{\alpha/\delta} y}\, dy$.

For the second factor, note first that if $\gamma = 0$, then $g(t,x) = -x$ and thus the second factor equals 1, which trivially is an infinitely divisible characteristic function with mean 0 and Lévy measure 0.

Thus assume $\gamma > 0$. Then, for $v > 0$,

$$(1 - ivt)^{-\gamma/\alpha} = \Gamma(\gamma/\alpha)^{-1} \int_0^\infty e^{ivty} y^{\gamma/\alpha - 1} e^{-y}\, dy$$

and thus, writing $h(t)=e^{it}-1-it$,

$$\begin{aligned}
R_1(-ivt)&=\Gamma(\gamma/\alpha)^{-1}\int_0^\infty \bigl(e^{ivty}-1\bigr)y^{\gamma/\alpha-1}e^{-y}\,dy\\
R_2(-ivt)&=\Gamma(\gamma/\alpha)^{-1}\int_0^\infty h(vty)y^{\gamma/\alpha-1}e^{-y}\,dy\\
&=\frac{v^{-\gamma/\alpha}}{\Gamma(\gamma/\alpha)}\int_0^\infty h(tz)z^{\gamma/\alpha-1}e^{-v^{-1}z}\,dz.
\end{aligned}$$

Hence, by Fubini's theorem,

$$\begin{aligned}
\int_x^\infty R_2(-i\alpha t u^{-\alpha/\delta})\,du&=\frac{\alpha^{-\gamma/\alpha}}{\Gamma(\gamma/\alpha)}\int_x^\infty\int_0^\infty h(tz)u^{\gamma/\delta}z^{\gamma/\alpha-1}e^{-\alpha^{-1}u^{\alpha/\delta}z}\,dz\,du\\
&=\int_0^\infty h(tz)\,d\nu_x^1(z),
\end{aligned}$$

where $\nu_x^1$ is the measure on $(0,\infty)$ with density

$$\frac{d\nu_x^1(z)}{dz}=\frac{\alpha^{-\gamma/\alpha}}{\Gamma(\gamma/\alpha)}z^{\gamma/\alpha-1}\int_x^\infty u^{\gamma/\delta}e^{-\alpha^{-1}u^{\alpha/\delta}z}\,du;$$

by Fubini's theorem again we have

$$\begin{aligned}
\int_0^\infty z^2\,d\nu_x^1(z)&=\frac{\alpha^{-\gamma/\alpha}}{\Gamma(\gamma/\alpha)}\int_x^\infty\int_0^\infty z^{\gamma/\alpha+1}u^{\gamma/\delta}e^{-\alpha^{-1}u^{\alpha/\delta}z}\,dz\,du\\
&=\frac{\alpha^2\Gamma(\gamma/\alpha+2)}{\Gamma(\gamma/\alpha)}\int_x^\infty u^{\gamma/\delta-(\gamma/\alpha+2)\alpha/\delta}\,du\\
&=\gamma(\gamma+\alpha)\int_x^\infty u^{-2\alpha/\delta}\,du\\
&=\frac{\gamma(\alpha+\gamma)}{2\alpha/\delta-1}x^{1-2\alpha/\delta}<\infty.
\end{aligned}\tag{8.4}$$

Thus, by (7.5) and (7.6), the second factor in (8.1) is the characteristic function of an infinitely divisible distribution with mean $-\gamma(1-\alpha/\delta)^{-1}x^{1-\alpha/\delta}$ (or $-\gamma\ln x$ if $\alpha=\delta$) and Lévy measure $\nu_x^1$. The result follows. □

It will be convenient to work mainly with real variables. We first extend the definitions (7.5)–(7.7) of $g$ to all complex $t$ with $\operatorname{Re} it\le 0$, and then define, for $\operatorname{Re}\tau\ge 0$,

$$\tilde g(\tau,x):=g(i\alpha^{-1}\tau,x)=\int_x^\infty R_2(\tau u^{-\alpha/\delta})\,du+\frac{\gamma/\alpha}{1-\alpha/\delta}\tau x^{1-\alpha/\delta}-x \tag{8.5}$$

(with the usual modification if $\alpha=\delta$). By Section 4, $\int e^{-ty}\,d\mu_x(y)$ is finite for every $x,t>0$, and (8.1) extends to

$$\int e^{-\tau y}\,d\mu_x(y)=(1+\alpha\tau x^{-\alpha/\delta})^{-x_0/\alpha}e^{\tilde g(\alpha\tau,x)+x},\qquad \operatorname{Re}\tau\ge 0. \tag{8.6}$$

By Theorem 8.1,

$$\mathbb{E} f(Z) = \Gamma(y_0/\delta)^{-1} \int_0^\infty x^{y_0/\delta - 1} e^{-x} \int f(y)\, d\mu_x(y)\, dx \tag{8.7}$$

for every measurable function $f$ such that $\mathbb{E} f(Z)$ makes sense; for example if $f \geq 0$ or $\mathbb{E}|f(Z)| < \infty$. In particular, by (8.6), for $\tau \geq 0$,

$$\mathbb{E} e^{-\tau Z} = \Gamma(y_0/\delta)^{-1} \int_0^\infty (1 + \alpha\tau x^{-\alpha/\delta})^{-x_0/\alpha} e^{\tilde{g}(\alpha\tau, x)} x^{y_0/\delta - 1}\, dx. \tag{8.8}$$

Since $\tilde{g}(\tau, x) < -x/2$ for large $x$ by (8.5), it follows that $\mathbb{E} e^{-\tau Z}$ is finite for all $\tau \geq 0$. (Hence it is a continuous function of $\tau$ for $\operatorname{Re}\tau \geq 0$, and analytic for $\operatorname{Re}\tau > 0$.) Hence (8.7) and (8.6) show that (8.8) holds for all complex $\tau$ with $\operatorname{Re}\tau \geq 0$.

It is now straightforward to compute the first moments of $Z$, and in particular to see when they are finite. Similar criteria for finiteness of higher moments, and similar (but more complicated) formulas for them, can be derived by the same method.

**Theorem 8.2.** *Assume $\alpha > \delta/2 > 0$ and $y_0 > 0$. Then $-\infty < \mathbb{E} Z \leq +\infty$, and*

$$\mathbb{E}|Z| < \infty \iff \begin{cases} y_0 > \alpha & \text{if } x_0 > 0, \\ y_0 + \delta > \alpha & \text{if } x_0 = 0 \text{ and } \gamma > 0; \end{cases}$$

$$\mathbb{E} Z^2 < \infty \iff \begin{cases} y_0 > 2\alpha & \text{if } x_0 > 0, \\ y_0 + \delta > 2\alpha & \text{if } x_0 = 0 \text{ and } \gamma > 0. \end{cases}$$

*When the moments are finite, they are given by, provided $\alpha \neq \delta$,*

$$\mathbb{E} Z = x_0 \frac{\Gamma\big((y_0 - \alpha)/\delta\big)}{\Gamma(y_0/\delta)} + \gamma \frac{\Gamma\big((y_0 + \delta - \alpha)/\delta\big)}{(\alpha/\delta - 1)\Gamma(y_0/\delta)},$$

$$\begin{aligned}
\mathbb{E} Z^2 &= x_0(x_0 + \alpha) \frac{\Gamma\big((y_0 - 2\alpha)/\delta\big)}{\Gamma(y_0/\delta)} \\
&\quad + \gamma \Big( \frac{\alpha + \gamma}{2\alpha/\delta - 1} + \frac{2x_0}{\alpha/\delta - 1} \Big) \frac{\Gamma\big((y_0 + \delta - 2\alpha)/\delta\big)}{\Gamma(y_0/\delta)} \\
&\quad + \gamma^2 \frac{\Gamma\big((y_0 + 2\delta - 2\alpha)/\delta\big)}{(\alpha/\delta - 1)^2 \Gamma(y_0/\delta)}.
\end{aligned}$$

When $\alpha = \delta$, the moments are given by appropriate limits of the expressions above; we leave the details to the reader.

*Proof.* Let $Z_+ := \max(Z, 0)$ and $Z_- := \max(-Z, 0)$. As remarked above, $\mathbb{E} e^{-Z} < \infty$, and thus $\mathbb{E} Z_- < \infty$. Consequently, (8.7) applied to $f(Z) = Z_+$ and $Z_-$ separately shows that we can take $f(Z) = Z$ too; thus

$$\mathbb{E} Z = \Gamma(y_0/\delta)^{-1} \int_0^\infty x^{y_0/\delta - 1} e^{-x} \lambda(x)\, dx \tag{8.9}$$

with both sides finite or $+\infty$, where $\lambda(x)$ is given by (8.2). For small $x$, we have by (8.2) $\lambda(x) \asymp x^{-\alpha/\delta}$ if $x_0 > 0$, while $\lambda(x) \asymp x^{1-\alpha/\delta}$ if $x_0 = 0$ and $\gamma > 0$ (and $\alpha \neq \delta$; otherwise $\lambda(x) \asymp |\ln x|$); the criterion for $\mathbb{E}|Z| < \infty$ follows. When $\mathbb{E}\, Z$ is finite, its value follows from (8.9) and (8.2).

Similarly, (8.7) yields, since $\int y^2\, d\mu_x(y) = \lambda(x)^2 + \int y^2\, d\nu_x(y)$,

$$\mathbb{E}\, Z^2 = \Gamma(y_0/\delta)^{-1} \int_0^\infty x^{y_0/\delta-1} e^{-x} \Bigl(\lambda(x)^2 + \int y^2\, d\nu_x(y)\Bigr)\, dx.$$

By (8.3) and (8.4), we have

$$\int_0^\infty y^2\, d\nu_x(y) = x_0 \alpha x^{-2\alpha/\delta} + \frac{\gamma(\alpha+\gamma)}{2\alpha/\delta - 1} x^{1-2\alpha/\delta},$$

and the results for $\mathbb{E}\, Z^2$ follow, using (8.2) and straightforward calculations. □

Observe next that by (8.5), or by (7.5), (7.6), (7.7), also when $\alpha = \delta$,

$$\begin{aligned}\frac{\partial}{\partial x}\tilde g(\tau, x) &= -R_2(\tau x^{-\alpha/\delta}) + (\gamma/\alpha)\tau x^{-\alpha/\delta} - 1 \\ &= -(1 + \tau x^{-\alpha/\delta})^{-\gamma/\alpha}.\end{aligned} \tag{8.10}$$

Hence $|\frac{\partial}{\partial x}\tilde g(\tau, x)| \leq 1$ for $x > 0$, $\operatorname{Re}\tau \geq 0$, and thus the limit $\tilde g(\tau, 0) := \lim_{x\to 0} \tilde g(\tau, x)$ exists uniformly for $\operatorname{Re}\tau \geq 0$, and

$$\tilde g(\tau, x) = \tilde g(\tau, 0) - \int_0^x (1 + \tau u^{-\alpha/\delta})^{-\gamma/\alpha}\, du. \tag{8.11}$$

The integral in (8.5) is, for fixed $x > 0$, a continuous function of $\tau$ for $\operatorname{Re}\tau \geq 0$, and analytic for $\operatorname{Re}\tau > 0$. Hence, so is $\tilde g(\tau, x)$ for every fixed $x > 0$, and, by uniform convergence, also $\tilde g(\tau, 0)$.

**Lemma 8.3.** *Suppose $\alpha > \delta/2 > 0$ and $\operatorname{Re}\tau \geq 0$.*

(i) *If $\alpha \neq \delta$, then*

$$\tilde g(\tau, 0) = \frac{\delta}{\alpha} B\Bigl(\frac{\gamma+\delta}{\alpha}, \frac{-\delta}{\alpha}\Bigr)\tau^{\delta/\alpha} = -\frac{\Gamma\bigl(\frac{\gamma+\delta}{\alpha}\bigr)\Gamma\bigl(1 - \frac{\delta}{\alpha}\bigr)}{\Gamma\bigl(\frac{\gamma}{\alpha}\bigr)}\tau^{\delta/\alpha}.$$

(ii) *If $\alpha = \delta$, then*

$$\tilde g(\tau, 0) = \frac{\gamma}{\alpha}\Bigl(\psi\Bigl(\frac{\gamma}{\alpha} + 1\Bigr) - \psi(2)\Bigr)\tau + \frac{\gamma}{\alpha}\tau\ln\tau.$$

*Proof.* By analytic continuation, it suffices to consider real $\tau > 0$.

If $\alpha < \delta$, then (8.5) and Lemma 4.1(ii) yield, with $v = \tau u^{-\alpha/\delta}$,

$$\begin{aligned}\tilde g(\tau, 0) &= \int_0^\infty R_2(\tau u^{-\alpha/\delta})\, du = \frac{\delta}{\alpha}\tau^{\delta/\alpha}\int_0^\infty R_2(v) v^{-\delta/\alpha-1}\, dv \\ &= \frac{\delta}{\alpha}\tau^{\delta/\alpha} B\Bigl(\frac{\gamma+\delta}{\alpha}, \frac{-\delta}{\alpha}\Bigr).\end{aligned}$$

The case $\alpha > \delta$ is similar, using (7.7) and Lemma 4.1(i).

If $\alpha = \delta$, finally, for $0 < x < \tau$,

$$\tilde{g}(\tau, x) + x = \int_x^\infty R_2(\tau u^{-1})\,du + \frac{\gamma}{\alpha}\tau \ln x = \tau \int_0^{\tau/x} R_2(v) v^{-2}\,dv + \frac{\gamma}{\alpha}\tau \ln x$$

$$= \tau \int_0^1 R_2(v) v^{-2}\,dv + \tau \int_1^{\tau/x} R_1(v) v^{-2}\,dv + \tau \frac{\gamma}{\alpha} \ln \frac{\tau}{x} + \frac{\gamma}{\alpha}\tau \ln x.$$

Hence,

$$\tilde{g}(\tau, 0) = \tau \Big( \int_0^1 R_2(v) v^{-2}\,dv + \int_1^\infty R_1(v) v^{-2}\,dv \Big) + \frac{\gamma}{\alpha}\tau \ln \tau,$$

and the result follows by Lemma 4.1(iii). ∎

**Lemma 8.4.** *Suppose $\alpha > \delta/2 > 0$ and $\alpha \neq \delta$. If $\tau > 0$ and $x > 0$, then $\tilde{g}(\tau, \tau^{\delta/\alpha} x) = \tau^{\delta/\alpha} \tilde{g}(1, x)$.*

*Proof.* Follows by (8.11) with the change of variables $u = \tau^{\delta/\alpha}$ and Lemma 8.3. ∎

The formulas (7.4) and (8.8) can be transformed in various ways using changes of variables and the formulas for $\tilde{g}$ above, for example as follows.

**Theorem 8.5.** *Assume $\alpha > \delta/2 > 0$ and $y_0 > 0$. If $\operatorname{Re} \tau \geq 0$ and (redundant for $\alpha > \delta$) $|\arg \tau| < \alpha\pi/2\delta$, then*

$$\mathbb{E}\, e^{-\tau Z/\alpha} = \frac{\delta/\alpha}{\Gamma(y_0/\delta)} \tau^{y_0/\alpha} \int_0^\infty (1+x)^{-x_0/\alpha} e^{\tilde{g}(\tau, \tau^{\delta/\alpha} x^{-\delta/\alpha})} x^{-y_0/\alpha - 1}\,dx \quad (8.12)$$

*where*

$$\tilde{g}(\tau, \tau^{\delta/\alpha} x^{-\delta/\alpha}) = \tilde{g}(\tau, 0) - \frac{\delta}{\alpha} \tau^{\delta/\alpha} \int_x^\infty (1+u)^{-\gamma/\alpha} u^{-\delta/\alpha - 1}\,du \quad (8.13)$$

*with $\tilde{g}(\tau, 0)$ given by Lemma 8.3.*

*Proof.* Assume first $\tau > 0$. Then (8.12) follows from (8.8) by the change of variables $y = \tau x^{-\alpha/\delta}$ (and then replacing $y$ by $x$), while (8.13) similarly follows from (8.11) and $v = \tau u^{-\alpha/\delta}$. The general case, with (8.13) as a definition of $\tilde{g}(\tau, \tau^{\delta/\alpha} x^{-\delta/\alpha})$, follows by analytic continuation. ∎

In a special case, we find a simple form for the limit.

**Theorem 8.6.** *If $\alpha > \delta/2$, $x_0 = \gamma > 0$ and $y_0 = \delta > 0$, then the limit $Z$ has a spectrally positive $\delta/\alpha$-stable distribution with*

$$\mathbb{E}\, e^{-\tau Z} = e^{\tilde{g}(\alpha\tau, 0)}, \qquad \operatorname{Re} \tau \geq 0, \quad (8.14)$$

*where $\tilde{g}(\alpha\tau, 0)$ is given by Lemma 8.3. The Lévy measure has density*

$$\alpha^{\delta/\alpha} \frac{\delta \Gamma\big(\frac{\gamma+\delta}{\alpha}\big)}{\alpha \Gamma\big(\frac{\gamma}{\alpha}\big)} x^{-\delta/\alpha - 1}, \qquad x > 0.$$

*Proof.* Assume $\tau > 0$. By (8.10),

$$\frac{\partial}{\partial x}\tilde{g}(\tau, \tau^{\delta/\alpha}x^{-\delta/\alpha}) = \frac{\delta}{\alpha}(1+x)^{-\gamma/\alpha}\tau^{\delta/\alpha}x^{-\delta/\alpha-1}$$

and thus (8.12) can be written

$$\mathbb{E}\, e^{-\tau Z/\alpha} = \int_0^\infty \frac{\partial}{\partial x} e^{\tilde{g}(\tau,\tau^{\delta/\alpha}x^{-\delta/\alpha})}\, dx = e^{\tilde{g}(\tau,0)} - e^{\tilde{g}(\tau,\infty)} = e^{\tilde{g}(\tau,0)},$$

because $\tilde{g}(\tau,\infty) = -\infty$ by (8.11).

This proves (8.14) for $\tau > 0$, and the general case follows by analytic continuation. The formulas in Lemma 8.3 now show that $Z$ is $\delta/\alpha$-stable with the given Lévy measure, see e.g. [9, XVII.3(g)]. □

In Theorem 7.1(v), $X^{(m)}$ is normalized without subtraction of a constant; thus the limit $Z \geq 0$ a.s. We can sharpen this and prove that all negative moments of $Z$ are finite.

**Theorem 8.7.** *Suppose $\alpha > \delta > 0$ and $y_0 > 0$. Then $Z > 0$ a.s., and if $r > 0$, then $\mathbb{E}\, Z^{-r} < \infty$. More precisely,*

$$\mathbb{E}\, Z^{-r} = \alpha^{-r}\frac{\Gamma(y_0/\delta + r\alpha/\delta)}{\Gamma(y_0/\delta)\Gamma(r)}\int_0^\infty f(x)^{-y_0/\delta - r\alpha/\delta}(1+x)^{-x_0/\alpha}x^{-y_0/\alpha-1}\, dx, \tag{8.15}$$

*where*

$$\begin{aligned} f(x) &:= x^{-\delta/\alpha} + \frac{\delta}{\alpha}\int_0^x \bigl(1-(1+v)^{-\gamma/\alpha}\bigr)v^{-\delta/\alpha-1}\, dv \\ &= -\tilde{g}(1,0) + \frac{\delta}{\alpha}\int_x^\infty (1+v)^{-\gamma/\alpha}v^{-\delta/\alpha-1}\, dv. \end{aligned} \tag{8.16}$$

*Proof.* We define $f(x) = -\tilde{g}(1, x^{-\delta/\alpha})$; then (8.16) follows from (8.5) (with $v = u^{-\delta/\alpha}$) and (8.13). Note that $f(x) > 0$.

By Fubini's theorem twice, (8.12) and Lemma 8.4,

$$\begin{aligned} \mathbb{E}(Z/\alpha)^{-r}\Gamma(r) &= \mathbb{E}\int_0^\infty e^{-\tau Z/\alpha}\tau^{r-1}\, d\tau \\ &= \frac{\delta/\alpha}{\Gamma(y_0/\delta)}\int_0^\infty\int_0^\infty \tau^{y_0/\alpha+r-1}(1+x)^{-x_0/\alpha}e^{-\tau^{\delta/\alpha}f(x)}x^{-y_0/\alpha-1}\, dx\, d\tau \\ &= \frac{\Gamma(y_0/\delta + r\alpha/\delta)}{\Gamma(y_0/\delta)}\int_0^\infty f(x)^{-y_0/\alpha - r}(1+x)^{-x_0/\alpha}x^{-y_0/\alpha-1}\, dx. \end{aligned}$$

To see that the integral is finite, we observe from (8.16) that $f(x) \sim x^{-\delta/\alpha}$ as $x \to 0$, and $f(x) \sim -\tilde{g}(1,0) > 0$ as $x \to \infty$. □

In the special case of a balanced urn, we can evaluate the integral in (8.15).

**Theorem 8.8.** *Suppose $\alpha = \delta + \gamma$ with $\gamma, \delta > 0$ and $y_0 > 0$. Then, for every $r > 0$,*

$$\mathbb{E}\, Z^{-r} = \alpha^{-r} \frac{\Gamma(y_0/\delta + r\alpha/\delta)\Gamma(x_0/\alpha + y_0/\alpha)}{\Gamma(x_0/\alpha + y_0/\alpha + r)\Gamma(y_0/\delta)}.$$

*Proof.* By Lemma 8.3, $\tilde{g}(1,0) = -1$. Further,

$$\frac{\partial}{\partial v}\bigl(1+v^{-1}\bigr)^{\delta/\alpha} = -\frac{\delta}{\alpha}\bigl(1+v^{-1}\bigr)^{\delta/\alpha-1}v^{-2} = -\frac{\delta}{\alpha}\bigl(1+v\bigr)^{-\gamma/\alpha}v^{-1-\delta/\alpha}$$

and thus (8.16) yields $f(x) = (1+x^{-1})^{\delta/\alpha} = (1+x)^{\delta/\alpha}x^{-\delta/\alpha}$. Hence, the integral in (8.15) becomes the Beta integral

$$\int_0^\infty (1+x)^{-y_0/\alpha - r - x_0/\alpha} x^{y_0/\alpha + r - y_0/\alpha - 1}\, dx,$$

and the result follows by Theorem 8.7 and (4.2). □

## 9. Proof of Theorem 1.3

*Proof.* We may now obtain limit results for $X_n$ and $Y_n$ by using the following lemma to invert the results for $X^{(m)}$ in Theorem 7.1.

**Lemma 9.1.** *Suppose $\delta > 0$ and that $m, n \geq 0$ are integers. If $\alpha > 0$, then the following are equivalent.*

(i) $Y_n < m\delta + y_0$,
(ii) $X^{(m)} > \alpha n + (\gamma - \alpha)m + x_0$.

*If $\alpha < 0$,* (i) *is instead equivalent the opposite inequality*

(ii)′ $X^{(m)} < \alpha n + (\gamma - \alpha)m + x_0$.

*Proof.* Let $N$ be the time of the $m$:th white draw. (i) means that $N > n$, i.e., by Lemma 1.2 used at time $N$, if $\alpha > 0$,

$$X^{(m)} - x_0 + \frac{\alpha - \gamma}{\delta}(m\delta + y_0 - y_0) = N\alpha > n\alpha.$$

This is the same as (ii). If $\alpha < 0$, the inequalities are reversed. □

To prove Theorem 1.3, let in each of the five cases, $\widetilde{Y}_n$ denote the left hand side of the corresponding equation (1.2)–(1.8), and let $x$ be a fixed real number. (We exclude the at most countably many values of $x$ that are discontinuity points of the distribution of $W$ below. We doubt that there may be any such points, but we do not know.) The idea of the proof is to define an integer $m$, depending on $n$, such that $\widetilde{Y}_n < x$ is equivalent to $Y_n < m\delta + y_0$ and then use Lemma 9.1 and Theorem 7.1. The details will vary slightly between the different cases.

In cases (i), (ii), (iii), let $a_n := n^{1/2}$, $(n \ln n)^{1/2}$ and $n^{\alpha/\delta}$, respectively; thus $\widetilde{Y}_n = a_n^{-1}\bigl(Y_n - \delta\frac{\delta-\alpha}{\delta-\alpha+\gamma}n\bigr)$. We define

$$m := \Bigl\lceil \frac{\delta - \alpha}{\delta - \alpha + \gamma}n + \delta^{-1}(a_n x - y_0)\Bigr\rceil = \frac{\delta - \alpha}{\delta - \alpha + \gamma}n + \delta^{-1}a_n x + O(1). \tag{9.1}$$

Thus, because $(Y_n - y_0)/\delta$ always is an integer,

$$\widetilde{Y}_n < x \iff Y_n < \delta\frac{\delta-\alpha}{\delta-\alpha+\gamma}n + a_n x \iff \frac{Y_n - y_0}{\delta} < m,$$

which by Lemma 9.1, temporarily assuming $\alpha > 0$, is equivalent to $X^{(m)} > \alpha n + (\gamma-\alpha)m + x_0$. By (9.1),

$$n = \frac{\delta-\alpha+\gamma}{\delta-\alpha}(m - \delta^{-1}a_n x) + O(1)$$

and thus

$$\begin{aligned}\alpha n + (\gamma-\alpha)m + x_0 &= \Bigl(\alpha\frac{\delta+(\gamma-\alpha)}{\delta-\alpha} + (\gamma-\alpha)\Bigr)m - \frac{\alpha(\delta-\alpha+\gamma)}{\delta(\delta-\alpha)}a_n x + O(1)\\ &= \frac{\delta\gamma}{\delta-\alpha}m - \frac{\alpha(\delta-\alpha+\gamma)}{\delta(\delta-\alpha)}a_n x + O(1).\end{aligned}$$

Further, as $n \to \infty$, (9.1) shows that $m \sim \frac{\delta-\alpha}{\delta-\alpha+\gamma}n$ so $m \to \infty$ and $a_m \to \infty$. Defining $\lambda := \bigl(\frac{\delta-\alpha}{\delta-\alpha+\gamma}\bigr)^{1/2}$ in cases (i) and (ii) and $\lambda := \bigl(\frac{\delta-\alpha}{\delta-\alpha+\gamma}\bigr)^{\alpha/\delta}$ in case (iii), we further have $a_m/a_n \to \lambda$. Consequently, Lemma 9.1 and Theorem 7.1 show that

$$\begin{aligned}\mathbb{P}(\widetilde{Y}_n < x) &= \mathbb{P}(Y_n < m\delta + y_0)\\ &= \mathbb{P}\Bigl(X^{(m)} > \frac{\delta\gamma}{\delta-\alpha}m - \frac{\alpha(\delta-\alpha+\gamma)}{\delta(\delta-\alpha)}a_n x + O(1)\Bigr)\\ &= \mathbb{P}\Bigl(a_m^{-1}\Bigl(X^{(m)} - \frac{\delta\gamma}{\delta-\alpha}m\Bigr) > -\frac{\alpha(\delta-\alpha+\gamma)}{\delta(\delta-\alpha)}\frac{a_n}{a_m}x + o(1)\Bigr)\\ &\to \mathbb{P}\Bigl(Z > -\frac{\alpha(\delta-\alpha+\gamma)}{\delta(\delta-\alpha)}\lambda^{-1}x\Bigr) = \mathbb{P}\Bigl(-\lambda\frac{\delta(\delta-\alpha)}{\alpha(\delta-\alpha+\gamma)}Z < x\Bigr),\end{aligned}$$

where $Z \sim N(0, \sigma_1^2)$ in case (i) and $Z \sim N(0, \gamma(\gamma+\alpha))$ in case (ii); hence

$$\widetilde{Y}_n \xrightarrow{\mathrm{d}} W := -\lambda\frac{\delta(\delta-\alpha)}{\alpha(\delta-\alpha+\gamma)}Z.$$

If $\alpha < 0$, the same argument works with some of the inequalities above reversed (but the final result is the same); we omit the details.

In the exceptional case $\alpha = 0$, $X^{(m)}$ is deterministic and Theorem 7.1 does not help. Instead, we may argue directly and note that the proportion of white balls stays close to $\delta/(\gamma+\delta)$, and thus the number of white draws is has approximatively a $\mathrm{Bi}(n, \delta/(\delta+\gamma))$ distribution, which leads to (1.2). We omit the details, since as remarked in Remark 1.10, the result also follows from [14].

In case (iv), the most complicated case, we define

$$\begin{aligned}m &:= \Bigl\lceil \frac{\alpha}{\gamma}\frac{n}{\ln n}\Bigl(1 + \frac{\ln\ln n}{\ln n} + \frac{\gamma x}{\alpha^2 \ln n}\Bigr) - \frac{y_0}{\alpha}\Bigr\rceil\\ &= \frac{\alpha}{\gamma}\frac{n}{\ln n}\Bigl(1 + \frac{\ln\ln n}{\ln n} + \frac{\gamma x}{\alpha^2\ln n} + O\Bigl(\frac{\ln n}{n}\Bigr)\Bigr).\end{aligned}$$

Thus

$$\ln m = \ln n - \ln\ln n + \ln\frac{\alpha}{\gamma} + o(1) = \ln n\Bigl(1 - \frac{\ln\ln n}{\ln n} + \frac{\ln(\alpha/\gamma)}{\ln n} + o\Bigl(\frac{1}{\ln n}\Bigr)\Bigr)$$

and, multiplying these equations,

$$m\ln m = \frac{\alpha}{\gamma}n\Bigl(1 + \frac{\gamma x/\alpha^2 + \ln(\alpha/\gamma) + o(1)}{\ln n}\Bigr).$$

Hence

$$\begin{aligned}\alpha n &= \gamma m\ln m - \alpha\frac{n}{\ln n}\Bigl(\frac{\gamma x}{\alpha^2} + \ln\frac{\alpha}{\gamma} + o(1)\Bigr)\\ &= \gamma m\ln m - \gamma m\Bigl(\frac{\gamma x}{\alpha^2} + \ln\frac{\alpha}{\gamma} + o(1)\Bigr).\end{aligned}$$

Consequently, by Lemma 9.1 and Theorem 7.1(iv),

$$\begin{aligned}\mathbb{P}(\widetilde{Y}_n < x) &= \mathbb{P}(Y_n < m\delta + y_0) = \mathbb{P}\Bigl(X^{(m)} > \alpha n + (\gamma-\alpha)m + x_0\Bigr)\\ &= \mathbb{P}\Bigl(X^{(m)} > \gamma m\ln m + m\Bigl(-\frac{\gamma^2 x}{\alpha^2} - \gamma\ln\frac{\alpha}{\gamma} + \gamma - \alpha + o(1)\Bigr)\Bigr)\\ &\to \mathbb{P}\Bigl(Z > -\frac{\gamma^2 x}{\alpha^2} - \gamma\ln\frac{\alpha}{\gamma} + \gamma - \alpha\Bigr)\\ &= \mathbb{P}(W < x).\end{aligned}$$

Finally, in case (v), for $x > 0$, define

$$m := \lceil \delta^{-1}(n^{\delta/\alpha}x - y_0)\rceil = \delta^{-1}n^{\delta/\alpha}x + O(1).$$

We now have

$$\alpha n + (\gamma-\alpha)m + x_0 = \alpha n + o(n) = \alpha(x/\delta)^{-\alpha/\delta}m^{\alpha/\delta}(1+o(1))$$

and, recalling that $Z > 0$ a.s. by Theorem 8.7,

$$\begin{aligned}\mathbb{P}(\widetilde{Y}_n < x) &= \mathbb{P}(Y_n < m\delta + y_0) = \mathbb{P}\Bigl(X^{(m)} > \alpha n + (\gamma-\alpha)m + x_0\Bigr)\\ &= \mathbb{P}\Bigl(m^{-\alpha/\delta}X^{(m)} > \alpha(x/\delta)^{-\alpha/\delta} + o(1)\Bigr)\\ &\to \mathbb{P}\Bigl(Z > \alpha(x/\delta)^{-\alpha/\delta}\Bigr) = \mathbb{P}\Bigl((Z/\alpha)^{-\delta/\alpha} < x/\delta\Bigr) = \mathbb{P}(W < x).\end{aligned}$$

This completes the proof of Theorem 1.3. ∎

## 10. Proofs of Theorems 1.6–1.8

*Proof of Theorem 1.6.* An immediate consequence of (1.8) and Theorem 8.7. ∎

*Proof of Theorem 1.7.* An immediate consequence of (1.8) and Theorem 8.8. ∎

*Proof of Theorem 1.8.* Let $U$ have a Mittag-Leffler distribution with density function $f$. Thus $\int_0^\infty x^r f(x)\,dx = \mathbb{E}\,U^r = \Gamma(1+r)/\Gamma(1+r\delta/\alpha)$ for every $r > -1$. Consequently, for every $r \geq 0$,

$$\int_0^\infty x^r c x^{y_0/\delta - x_0/\gamma} f(x/\delta)\,dx = \delta^r \frac{\mathbb{E}\,U^{r+y_0/\delta-x_0/\gamma}}{\mathbb{E}\,U^{y_0/\delta-x_0/\gamma}}$$
$$= \delta^r \frac{\Gamma(1+r+y_0/\delta-x_0/\gamma)\Gamma(1+y_0/\alpha-x_0\delta/\alpha\gamma)}{\Gamma(1+r\delta/\alpha+y_0/\alpha-x_0\delta/\alpha\gamma)\Gamma(1+y_0/\delta-x_0/\gamma)}.$$

It is easily verified that this equals $\mathbb{E}\,W^r$ in Theorem 1.7 in the two cases $x_0 = 0$ and $x_0 = \gamma$. The result follows, since Theorem 1.7 also implies that $W$ has a finite Laplace transform, and thus its distribution is determined by its moments. $\square$

## 11. The diagonal case

In this section we consider the diagonal case $\beta = \gamma = 0$, i.e. only balls of the same colour as the drawn one are added. To avoid trivialities, we assume $\alpha, \delta, x_0, y_0 > 0$.

Theorem 7.1 is valid in this case too, but the limits in (i) and (ii) are degenerate, which means that the normalization is wrong. In cases (iii)–(v), we have $g(t,x) = -x$, since now $R_1(z) = R_2(z) = 0$, and (7.4) implies easily that $Z \overset{\mathrm{d}}{=} \alpha U V^{-\alpha/\delta}$, where $U$ and $V$ are independent Gamma variables as in Theorem 1.4; thus (in all three cases)

$$m^{-\alpha/\delta} X^{(m)} \overset{\mathrm{d}}{\to} \alpha U V^{-\alpha/\delta} \qquad \text{as } m \to \infty. \tag{11.1}$$

This can be seen more easily as follows, which also includes cases (i) and (ii).

In the diagonal case, the branching process $\big(\mathcal{X}(t), \mathcal{Y}(t)\big)$ consists of two independent generalized Yule processes. As is well-known,

$$e^{-\alpha t}\mathcal{X}(t)/\alpha \overset{\text{a.s.}}{\to} U \sim \Gamma(x_0/\alpha, 1) \qquad \text{as } t \to \infty; \tag{11.2}$$

indeed, the a.s. convergence follows because $e^{-\alpha t}\mathcal{X}(t)$ is a positive martingale [3, Theorem III.7.1], and the distribution of the limit is easily found from (5.1). Similarly, interchanging the colours,

$$e^{-\delta t}\mathcal{Y}(t)/\delta \overset{\text{a.s.}}{\to} V \sim \Gamma(y_0/\delta, 1) \qquad \text{as } t \to \infty.$$

Here $U$ and $V$ are independent, because the processes $\mathcal{X}(t)$ and $\mathcal{Y}(t)$ are. Consequently,

$$\frac{\mathcal{X}(t)/\alpha}{(\mathcal{Y}(t)/\delta)^{\alpha/\delta}} \overset{\text{a.s.}}{\to} \frac{U}{V^{\alpha/\delta}} \qquad \text{as } t \to \infty. \tag{11.3}$$

Letting $t$ be the time of the $m$:th white death, we obtain (11.1). We can argue as in Section 9 to obtain Theorem 1.4 from (11.1), but it is easier to use (11.3) directly as follows.

*Proof of Theorem 1.4.* Let $t_n$ be the time of the $n$:th death; thus $(X_n, Y_n) = \bigl(\mathcal{X}(t_n), \mathcal{Y}(t_n)\bigr)$, and since $t_n \to \infty$ we obtain

$$\frac{X_n/\alpha}{(Y_n/\delta)^{\alpha/\delta}} \overset{\text{a.s.}}{\to} \frac{U}{V^{\alpha/\delta}} \qquad \text{as } n \to \infty. \tag{11.4}$$

Note also that (Lemma 1.2)

$$X_n/\alpha + Y_n/\delta = n + x_0/\alpha + y_0/\delta = n + O(1). \tag{11.5}$$

It follows from (11.4) and (11.5) that $X_n \overset{\text{a.s.}}{\to} \infty$ and $Y_n \overset{\text{a.s.}}{\to} \infty$.

If $\alpha < \delta$, then (11.4) implies that $X_n/Y_n \overset{\text{a.s.}}{\to} 0$. Thus, by (11.5), $Y_n/\delta \sim n$, so (11.4) again yields

$$n^{-\alpha/\delta} X_n/\alpha \overset{\text{a.s.}}{\to} UV^{-\alpha/\delta} \qquad \text{as } n \to \infty. \tag{11.6}$$

This yields (1.9) by (11.5).

If $\alpha > \delta$, we interchange the colours; (11.6) then yields (1.11).

If $\alpha = \delta$, (11.4) and (11.5) yield (1.10). □

## 12. The case $\delta = 0$

*Proof of Theorem 1.5.* Since $Y_n = y_0$, $(X_n)_n$ is a Markov chain with transitions $x \to x + \alpha$ and $x \to x + \gamma$. We also consider the branching process $\bigl(\mathcal{X}(t), \mathcal{Y}(t)\bigr)$, where now $\mathcal{Y}(t) = y_0$ is constant while $\mathcal{X}(t)$ is a generalized Yule process as in Section 5 but with immigration: bunches of $\gamma$ black balls are added at the white draws, which occur according to a Poisson process with intensity $y_0$.

(i): We may assume $\alpha = -1$ by Remark 1.11; then $\gamma \geq 1$ and $x_0 \geq 0$ are integers. It is easily seen that the Markov chain (with state space $\{0, 1, 2, \dots\}$) is irreducible, and that it has period $\gamma+1$ because $X_n - x_0 \equiv -n \pmod{\gamma + 1}$.

Consider next the branching process $\mathcal{X}(t)$. In this case, the branching process originating from each new black ball trivially is subcritical and dies out. As a special case of [13, Theorem (7.1.1)] (and not difficult to verify directly), then $\mathcal{X}(t) \overset{\text{d}}{\to} W$ as $t \to \infty$, where $W$ is a compound Poisson distribution with probability generating function

$$\mathbb{E}\, z^W = \exp\biggl(\int_0^\infty \bigl(\Phi(z;t)^\gamma - 1\bigr) y_0 \, dt\biggr), \tag{12.1}$$

with $\Phi$ given by (5.1). Since $\alpha = -1$, we have

$$\Phi(z;t)^\gamma - 1 = \bigl(ze^{-t} + (1 - e^{-t})\bigr)^\gamma - 1 = \sum_{j=0}^{\gamma} \binom{\gamma}{j} (z^j - 1) e^{-jt} (1 - e^{-t})^{\gamma - j}$$

and thus

$$\int_0^\infty \bigl(\Phi(z;t)^\gamma - 1\bigr)\,dt = \sum_{j=1}^{\gamma}\binom{\gamma}{j}(z^j-1)\int_0^\infty e^{-jt}(1-e^{-t})^{\gamma-j}\,dt$$
$$= \sum_{j=1}^{\gamma}\binom{\gamma}{j}(z^j-1)B(j,\gamma-j+1) = \sum_{j=1}^{\gamma}\frac{z^j-1}{j},$$

which by (12.1) gives (1.12).

The branching process $\mathcal{X}(t)$ and the Markov chain $X_n$ have the same transitions, but the intensity of a transition $(x,y)\to(x',y')$ is $(x+y)$ times larger in the branching process. Hence, if $p_k = c(k+y_0)q_k$, with $c>0$ such that $\sum_0^\infty p_k = 1$, then $(p_k)_k$ is a stationary distribution for $X_n$. We have the probability generating function, with $Q(z) := \mathbb{E}\, z^W = \sum_k q_k z^k$,

$$P(z) := \sum_{k=0}^{\infty} p_k z^k = c\bigl(zQ'(z) + y_0 Q(z)\bigr) = c\Bigl(y_0\sum_{j=1}^{\gamma} z^j + y_0\Bigr)Q(z)$$
$$= cy_0\sum_{j=0}^{\gamma} z^j Q(z).$$

Setting $z=1$ we find $c = 1/(y_0(\gamma+1))$, and thus $P(z) = F(z)Q(z)$, where $F(z) := \sum_{j=0}^{\gamma} z^j/(\gamma+1)$ is the probability generating function of a random variable $U$ that is uniformly distributed on $\{0,1,\dots,\gamma\}$. It follows that $p_k$ is the distribution of $U+W$, with $U$ and $W$ independent. Since the Markov chain has a stationary distribution, it is persistent, and if $k\equiv x_0 - j \pmod{\gamma+1}$, then

$$\mathbb{P}(X_{(\gamma+1)n+j} = k) \to (\gamma+1)p_k = \sum_{j=0}^{\gamma} q_{k-j},$$

see [8, Chapter XV.6,7], which shows our claims.

Note that the branching process and the embedded urn process have different limits.

(ii): Let the white draws have numbers $N_1 < N_2 < \dots$, and let $N_0 = 0$. For $N_k \le n < N_{k+1}$, we have $(X_n, Y_n) = (x_0 + k\gamma, y_0)$. Hence, $N_{k+1} - N_k - 1$ has a geometric distribution $\mathrm{Ge}(y_0/(x_0+y_0+k\gamma))$, whence

$$\mathbb{E}(N_{k+1}-N_k) = \frac{x_0+y_0+k\gamma}{y_0} = \frac{\gamma}{y_0}k + O(1),$$
$$\mathrm{Var}(N_{k+1}-N_k) = \frac{x_0+y_0+k\gamma}{y_0}\,\frac{x_0+k\gamma}{y_0} = \frac{\gamma^2}{y_0^2}k^2 + O(k),$$
$$\mathbb{E}(N_{k+1}-N_k)^3 = O(k^3).$$

Further, the variables $N_{k+1}-N_k$, $k\ge 0$, are independent. Consequently,

$$\mathbb{E}\, N_k = \frac{\gamma}{2y_0}k^2 + O(k), \qquad \mathrm{Var}\, N_k = \frac{\gamma^2}{3y_0^2}k^3 + O(k^2),$$

and, by the central limit theorem with Lyapunov's condition,

$$\frac{N_k - (\gamma/2y_0)k^2}{k^{3/2}} \xrightarrow{\text{d}} N(0, \gamma^2/3y_0^2), \qquad \text{as } k \to \infty. \tag{12.2}$$

We invert this result by a standard argument from renewal theory. We may assume $\gamma > 0$ (otherwise no balls are ever added, and the result is trivial).

Let $x$ be a real number and let $k := \lceil \sqrt{2y_0/\gamma} n^{1/2} + (x/\gamma)n^{1/4} - x_0/\gamma \rceil$. Then,

$$\begin{aligned}\mathbb{P}(X_n < \sqrt{2y_0\gamma} n^{1/2} + xn^{1/4}) &= \mathbb{P}(X_n < x_0 + k\gamma) = \mathbb{P}(N_k > n) \\ &= \mathbb{P}\Big(\frac{N_k - (\gamma/2y_0)k^2}{k^{3/2}} > \frac{n - (\gamma/2y_0)k^2}{k^{3/2}}\Big).\end{aligned}$$

As $n \to \infty$,

$$\frac{n - (\gamma/2y_0)k^2}{k^{3/2}} \to -\frac{x}{y_0}\Big(\frac{2y_0}{\gamma}\Big)^{-1/4},$$

and (1.13) follows from (12.2) by simple calculations.

(iii): In this case ($\alpha > 0$), $\mathcal{X}(t)$ is a supercritical branching process with immigration. It follows that

$$e^{-\alpha t}\mathcal{X}(t) \xrightarrow{\text{a.s.}} Z, \tag{12.3}$$

for some random variable $Z \geq 0$, see [13, Theorem(7.1.6)]. (In our case, this can be seen easily by verifying that $e^{-\alpha t}(\mathcal{X}(t) + y_0\gamma/\alpha)$ is an $L^2$-bounded martingale, which implies (12.3).) Further, considering just the black balls descending from the original $x_0$ black balls, or (if $x_0 = 0$) from the first $\gamma$ added, we see by (11.2) that $Z > 0$ a.s.

Let $M(t)$ be the number of white draws up to time $t$, and let $\tau_n$ be the time of the $n$:th draw. Then

$$\mathcal{X}(\tau_n) = X_n = x_0 + (n - M(\tau_n))\alpha + M(\tau_n)\gamma. \tag{12.4}$$

Since $M(t)$ is a Poisson process with intensity $y_0$,

$$M(t)/t \xrightarrow{\text{a.s.}} y_0, \qquad t \to \infty.$$

In particular, $e^{-\alpha t}M(t) \to 0$ a.s., and (12.4) and (12.3) yield, since $\tau_n \to \infty$, a.s.,

$$e^{-\alpha\tau_n} n\alpha = e^{-\alpha\tau_n} X(\tau_n) + o(1) \to Z$$

and thus $-\alpha\tau_n + \ln n \to \ln(Z/\alpha)$. Consequently, if we define $t_n^{\pm} := \ln n/\alpha \pm \ln\ln n$, a.s. $t_n^- < \tau_n < t_n^+$ and thus $M(t_n^-) \leq M(\tau_n) \leq M(t_n^+)$ for all large $n$.

Since $M(t) \sim \text{Po}(y_0 t)$, we have $(M(t) - y_0 t))/t^{1/2} \xrightarrow{\text{d}} N(0, y_0)$ as $t \to \infty$, and thus, with $t_n := \ln n/\alpha$, $(M(t_n^{\pm}) - y_0 t_n))/t_n^{1/2} \xrightarrow{\text{d}} N(0, y_0)$. Consequently, $(M(\tau_n) - y_0 t_n))/t_n^{1/2} \xrightarrow{\text{d}} N(0, y_0)$, and (1.14) follows from (12.4). $\square$

## 13. The diagonal case again

In the diagonal case $\beta = \gamma = 0$, we can also use the embedding in a multi-type branching process in a somewhat different way. Let $p_{k,l}$ be the probability that of the first $k+l$ drawn balls, $k$ are black and $l$ are white. In the notation of Section 6 (if $\alpha, \delta \neq 0$), $p_{k,l} = \tilde{P}_{x_0+k\alpha,\, y_0+l\delta}$. We then have the following formula.

**Theorem 13.1.** *Assume $\alpha, \delta > 0$, $\beta = \gamma = 0$ and $x_0, y_0 > 0$. Then*

$$p_{k,l} = (x_0 + y_0 + k\alpha + l\delta)\frac{\Gamma(k + x_0/\alpha)\Gamma(l + y_0/\delta)}{\Gamma(x_0/\alpha)\Gamma(l + y_0/\delta)\, k!\, l!} \cdot \int_0^\infty e^{-(x_0+y_0)t}(1 - e^{-\alpha t})^k (1 - e^{-\delta t})^l \, dt. \quad (13.1)$$

*Proof.* The corresponding branching process $\big(\mathcal{X}(t), \mathcal{Y}(t)\big)$ consists of two independent generalized Yule processes as in Section 5, since the black and white balls act independently. Hence, by (5.2),

$$\begin{aligned} P_{x_0+k\alpha, y_0+l\delta}(t) &= \mathbb{P}(\mathcal{X}(t) = x_0 + k\alpha)\, \mathbb{P}(\mathcal{Y}(t) = y_0 + l\delta) \\ &= \frac{\Gamma(k + x_0/\alpha)}{\Gamma(x_0/\alpha)\, k!} \frac{\Gamma(l + y_0/\delta)}{\Gamma(l + y_0/\delta)\, l!} e^{-(x_0+y_0)t}(1 - e^{-\alpha t})^k (1 - e^{-\delta t})^l. \end{aligned}$$

The result follows by Lemma 6.3. □

By expanding the powers inside the integral in (13.1), we obtain a closed form expression for $p_{k,l}$ as a complicated alternating double sum; we doubt that this formula is of much use and leave the details to the interested reader.

It ought to be possible to derive local limit theorems from (13.1); again we leave this to the reader.

Instead we will use Theorem 13.1 to give an exact formula for $\mathbb{E}\, Y_n$ in a special example. (At least a few other examples can be treated in the same way, but the general case seems difficult.) Note that if $x_0 = \alpha$ and $y_0 = \delta$, (13.1) simplifies to

$$p_{k,l} = \big((k+1)\alpha + (l+1)\delta\big) \int_0^\infty e^{-(\alpha+\delta)t}(1 - e^{-\alpha t})^k (1 - e^{-\delta t})^l \, dt. \quad (13.2)$$

**Example 13.2.** Let $\alpha = x_0 = 2$, $\delta = y_0 = 1$, $\beta = \gamma = 0$. Thus a black [white] ball is replaced together with 2 [1] of the same colour. (Equivalently, see Remark 1.12, we take $\alpha = \delta = x_0 = y_0 = 1$ and weights 2 and 1; this is an urn where each drawn ball is replaced together with another of the same colour, as in Pólya's original urn, but each black ball is chosen with twice the probability of each white ball.) By (13.2),

$$\begin{aligned}\mathbb{P}(Y_n = k+1) = p_{n-k,k} &= (2n-k+3)\int_0^\infty e^{-3t}(1-e^{-2t})^{n-k}(1-e^{-t})^k\,dt \\ &= (2n-k+3)\int_0^1 x^2(1-x^2)^{n-k}(1-x)^k\,dx \\ &= (2n-k+3)\int_0^1 (1+x)^{n-k}(1-x)^n x^2\,dx. \qquad (13.3)\end{aligned}$$

Define, for fixed $n$ and $r \geq 0$, the polynomials

$$S_r(x) := \sum_{k=0}^n (k+1)^r(1+x)^{n-k}, \qquad (13.4)$$

$$T_r(x) := \sum_{k=0}^n (2n+3-k)(k+1)^r(1+x)^{n-k}. \qquad (13.5)$$

Then, by (13.3),

$$\mathbb{E}\,Y_n^r = \sum_{k=0}^n (k+1)^r\,\mathbb{P}(Y_n = k+1) = \int_0^1 T_r(x)(1-x)^n x^2\,dx. \qquad (13.6)$$

From the definitions,

$$\begin{aligned}S_0(x) &= \frac{(1+x)^{n+1}-1}{x}, \\ S_{r+1}(x) &= (n+1)S_r(x) - (1+x)S_r'(x), \\ T_r(x) &= (n+3)S_r(x) + (1+x)S_r'(x),\end{aligned}$$

which after some calculations give

$$x^2T_1(x) = \big((2n+3)x^2+4(n+1)x+2n-1-2x^{-1}\big)(1+x)^n-(n+2)^2x+1+2x^{-1}. \qquad (13.7)$$

We have, for $a > -1$, by (4.2),

$$\int_0^1 x^a(1-x)^n\,dx = \frac{\Gamma(a+1)\,n!}{\Gamma(n+a+2)}$$

$$\begin{aligned}\int_0^1 x^a(1+x)^n(1-x)^n\,dx &= \int_0^1 x^a(1-x^2)^n\,dx = \tfrac{1}{2}\int_0^1 y^{(a-1)/2}(1-y)^n\,dy \\ &= \frac{\Gamma\big((a+1)/2\big)\,n!}{2\Gamma(n+a/2+3/2)}.\end{aligned}$$

Thus also, with $\psi(z)=\Gamma'(z)/\Gamma(z)$ as in Section 4 and $H_n:=\sum_1^n k^{-1}$, using $\psi(n+1)=\psi(n)+1/n$,

$$\begin{aligned}
\int_0^1 \frac{(1+x)^n-1}{x}(1-x)^n\,dx &= \lim_{\varepsilon\downarrow 0}\int_0^1 x^{\varepsilon-1}\bigl((1-x^2)^n-(1-x)^n\bigr)\,dx\\
&= \lim_{\varepsilon\downarrow 0}\Bigl(\frac{\Gamma(\varepsilon/2)\,n!}{2\Gamma(n+1+\varepsilon/2)}-\frac{\Gamma(\varepsilon)\,n!}{\Gamma(n+1+\varepsilon)}\Bigr)\\
&= \lim_{\varepsilon\downarrow 0}\frac{n!}{\varepsilon}\Bigl(\frac{\Gamma(1+\varepsilon/2)}{\Gamma(n+1+\varepsilon/2)}-\frac{\Gamma(1+\varepsilon)}{\Gamma(n+1+\varepsilon)}\Bigr) = -\tfrac12\frac{d}{dt}\frac{n!\,\Gamma(1+t)}{\Gamma(n+1+t)}\Big|_{t=0}\\
&= \tfrac12\bigl(-\psi(1)+\psi(n+1)\bigr)=\tfrac12 H_n.
\end{aligned}$$

Consequently, we find by (13.6) and (13.7)

$$\begin{aligned}
\mathbb{E}\,Y_n &= (2n+3)\frac{\Gamma(3/2)\,n!}{2\Gamma(n+5/2)}+4(n+1)\frac{\Gamma(1)\,n!}{2\Gamma(n+2)}+(2n-1)\frac{\Gamma(1/2)\,n!}{2\Gamma(n+3/2)}\\
&\qquad -(n+2)^2\frac{\Gamma(2)\,n!}{\Gamma(n+3)}+\frac{1}{n+1}-H_n\\
&= \sqrt{\pi}\frac{n\cdot n!}{\Gamma(n+3/2)}-H_n+1. \qquad (13.8)
\end{aligned}$$

This exact formula leads to the asymptotics (where $\gamma$ is Euler's constant)

$$\mathbb{E}\,Y_n=\sqrt{\pi n}-\ln n+1-\gamma+O(n^{-1/2}); \qquad (13.9)$$

further terms can be found at will, cf. [18, 1.2.11.2]. Note that Theorem 1.4 yields $n^{-1/2}Y_n\xrightarrow{\mathrm{d}}W$, where by Theorem 1.9 $\mathbb{E}\,W=\sqrt{\pi}$; this fits nicely with the leading term in (13.9), although we have not proved moment convergence in general.

Exact and asymptotic formulas for the second and higher moments can be found in the same way.

The logarithmic second order term in (13.9) is a surprise, and shows that even the diagonal case is far from simple.

For some non-diagonal cases with $\gamma>0$ and $\alpha=\gamma+\delta$, similar exact and asymptotic formulas for the mean are given by [19] and [10], [11]. ([10] also treats the second moment.) In the examples worked out in these references, the asymptotic expansions contain only powers of $n$ and no logarithmic term as in our example.

It would be interesting to find similar refined asymptotic results for a non-diagonal case with $\alpha\neq\gamma+\delta$.

## References


[1] M. Abramowitz and I.A Stegun, *Handbook of Mathematical Functions.* Dover, New York, 1972.

[2] K.B. Athreya & S. Karlin, Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist.* **39** (1968), 1801–1817.

[3] K.B. Athreya & P.E. Ney, *Branching Processes.* Springer-Verlag, Berlin, 1972.

[4] A. Bagchi & A.K. Pal, Asymptotic normality in the generalized Pólya–Eggenberger urn model, with an application to computer data structures. *SIAM J. Algebraic Discrete Methods* **6** (1985), no. 3, 394–405.

[5] S. Bernstein, Nouvelles applications des grandeurs aléatoires presqu'indépendantes. (Russian) *Izv. Akad. Nauk SSSR Ser. Mat.* **4** (1940), 137–150.

[6] M. Drmota & V. Vatutin, Limiting distributions in branching processes with two types of particles. In *Classical and modern branching processes (Minneapolis, MN, 1994)*, 89–110, IMA Vol. Math. Appl., 84, Springer, New York, 1997.

[7] F. Eggenberger & G. Pólya, Über die Statistik verketteter Vorgänge. *Zeitschrift Angew. Math. Mech.* **3** (1923), 279–289.

[8] W. Feller, *An Introduction to Probability Theory and Its Applications. Vol. I.* Second edition, Wiley, New York 1957.

[9] W. Feller, *An Introduction to Probability Theory and Its Applications. Vol. II.* Second edition, Wiley, New York 1971.

[10] P. Flajolet, J. Gabarró & H. Pekari, Analytic urns. Preprint, 2003. Available from `http://algo.inria.fr/flajolet/Publications/publist.html` (The original version; not the revised!)

[11] P. Flajolet & V. Puyhaubert, in preparation.

[12] B. Friedman, A simple urn model. *Comm. Pure Appl. Math.* **2** (1949), 59–70.

[13] P. Jagers, *Branching Processes with Biological Applications.* Wiley, Chichester, London, 1975.

[14] S. Janson, Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Process. Appl.* **110** (2004), no. 2, 177–245.

[15] M. Jiřina, Stochastic branching processes with continuous state space. *Czechoslovak Math. J.* **8 (83)** (1958), 292–313.

[16] N.L. Johnson & S. Kotz, *Urn models and their application.* Wiley, New York, 1977.

[17] O. Kallenberg, *Foundations of modern probability.* 2nd ed., Springer-Verlag, New York, 2002.

[18] D.E. Knuth, *The Art of Computer Programming. Vol. 1: Fundamental Algorithms.* 3nd ed., Addison-Wesley, Reading, Mass., 1997.

[19] S. Kotz, H. Mahmoud & P. Robert, On generalized Pólya urn models. *Statist. Probab. Lett.* 49 (2000), no. 2, 163–173.

[20] R. Pemantle & S. Volkov, Vertex-reinforced random walk on $Z$ has finite range. *Ann. Probab.* **27** (1999), no. 3, 1368–1388.

[21] H. Pollard, The completely monotonic character of the Mittag-Leffler function $E_a(-x)$. *Bull. Amer. Math. Soc.* **54** (1948), 1115–1116.

[22] G. Pólya, Sur quelques points de la théorie des probabilités. *Ann. Inst. Poincaré* **1** (1931), 117–161.

[23] V. Puyhaubert, *Modèles d'urnes et phénomènes de seuils en combinatoire analytique.* Ph. D. thesis, l'École Polytechnique, 2005.

[24] R.T. Smythe, Central limit theorems for urn models. *Stochastic Process. Appl.* **65** (1996), no. 1, 115–137.

Department of Mathematics, Uppsala University, PO Box 480, S-751 06 Uppsala, Sweden

*E-mail address*: `svante.janson@math.uu.se`

*URL*: `http://www.math.uu.se/~svante/`