Noname manuscript No.
(will be inserted by the editor)

Functional limit theorems for the Pólya urn

Dimitris Cheliotis* · Dimitra Kouloumpou

Received: date / Accepted: date

Abstract For the plain Pólya urn with two colors, black and white, we prove a functional central limit theorem for the number of white balls assuming that the initial number of black balls is large. Depending on the initial number of white balls, the limit is either a pure birth process or a diffusion.

Keywords Pólya urn · Functional limit theorems · Birth processes · Diffusion processes

Mathematics Subject Classification (2010) 60F17 · 60K99 · 60C05

1 Introduction and results

The model. The Pólya urn is the model where in an urn that has initially A_0 white and B_0 black balls we draw, successively, and uniformly at random, a ball from it and then we return the ball back together with k balls of the same color as the one drawn. The number $k \in \mathbb{N}^+$ is fixed. Call A_n and B_n the number of white and black balls respectively after n drawings. The most notable result regarding the asymptotic behavior of the urn is that the proportion of white balls in the urn after n drawings, $A_n/(A_n+B_n)$, converges almost surely as $n \to \infty$ to a random variable with distribution $\text{Beta}(A_0/k, B_0/k)$.

Our aim in this work is to examine whether the entire path $(A_n)_{n\in\mathbb{N}}$, after appropriate natural transformations, converges in distribution to a nontrivial stochastic process.

Standard references for the theory and the applications of the Pólya urn and related models are [9] and [10].

Dimitris Cheliotis*
National and Kapodistrian University of Athens, Department of Mathematics, Panepistimiopolis, Athens 15784, Greece. Email: dcheliotis@math.uoa.gr.

Dimitra Kouloumpou
Hellenic Naval Academy, Section of Mathematics, Mathematical Modeling and Applications Laboratory, Piraeus 18539, Greece. Email: dimkouloumpou@hna.gr.

The setting. We consider an urn whose initial composition depends on $m \in \mathbb{N}^+$. It is $A_0^{(m)}$ and $B_0^{(m)}$ white and black balls respectively. After n drawings, the composition is $A_n^{(m)}, B_n^{(m)}$.

To see a new process arising out of the path $(A_n^{(m)})_{n\in\mathbb{N}}$, we start with an initial number of balls that tends to infinity as $m \to \infty$. More specifically, we assume then that $B_0^{(m)}$ grows linearly with m. Regarding $A_0^{(m)}$, we study three regimes:

a) $A_0^{(m)}$ stays fixed with m.
b) $A_0^{(m)}$ grows to infinity but sublinearly with m.
c) $A_0^{(m)}$ grows linearly with m.

The regime where $A_0^{(m)}$ grows superlinearly with m follows from regime b) by changing the roles of the two colors. We remark on this after Theorem 2.

In the regimes a) and b), the scarcity of white balls has as a result that the time between two consecutive drawings of a white ball is of order $m/A_0^{(m)}$ (the probability of picking a white ball in the first few drawings is approximately $A_0^{(m)}/m$, which is small). We expect then that speeding up time by this factor we will see a birth process. And indeed this is the case as our first two theorems show.

In this work, all processes appearing with index set $[0,\infty)$ and values in some Euclidean space $\mathbb{R}^d$ are elements of $D_{\mathbb{R}^d}[0,\infty)$, the space of functions $f:[0,\infty)\to\mathbb{R}^d$ that are right continuous and have limits from the left at each point of $[0,\infty)$. This space is endowed with the Skorokhod topology (defined in §5 of Chapter 3 of [5]), and convergence in distribution of processes with values on that space is defined through that topology.

We remind the reader that the negative binomial distribution with parameters $\nu \in (0,\infty)$ and $p \in (0,1)$ is the distribution with support in $\mathbb{N}$ and probability mass function

$$f(x) = \binom{x+\nu-1}{x} p^\nu (1-p)^x \tag{1}$$

for all $x \in \mathbb{N}$. When $\nu \in \mathbb{N}^+$, this is the distribution of the number of failures until we obtain the ν-th success in a sequence of independent trials, each having probability of success p. For a random variable X with this distribution, we write $X \sim NB(\nu, p)$.

Since in each drawing we add k balls in the urn, the quantity $k^{-1}\{A_n^{(m)} - A_0^{(m)}\}$, appearing in our first two theorems, counts the number of times in the first n drawings that we selected a white ball.

Theorem 1 *Fix $a_0 \in \mathbb{N}^+$ and $b > 0$. If $A_0^{(m)} = a_0$ for all $m \in \mathbb{N}^+$ and $\lim_{m\to\infty} B_0^{(m)}/m = b$, then the process $(k^{-1}\{A_{[mt]}^{(m)} - A_0^{(m)}\})_{t\geq 0}$ converges in distribution, as $m \to \infty$, to an inhomogeneous in time pure birth process $Z = \{Z(t)\}_{t\geq 0}$ with $Z(0) = 0$ and such that for all $0 \leq t_1 < t_2, j \in \mathbb{N}$,*

$$Z(t_2) - Z(t_1) | Z(t_1) = j \text{ has distribution } NB\Big(\frac{a_0}{k} + j, \frac{t_1 + (b/k)}{t_2 + (b/k)}\Big).$$

In particular, Z has rates $\lambda_{t,j} = (kj + a_0)/(kt + b)$ for all $(t, j) \in [0, \infty) \times \mathbb{N}$.

Theorem 2 *If $A_0^{(m)} =: g_m$ with $g_m \to \infty, g_m = o(m)$ and $\lim_{m\to\infty} B_0^{(m)}/m = b$ with $b > 0$ constant, then the process $(k^{-1}\{A_{[tm/g_m]}^{(m)} - A_0^{(m)}\})_{t\geq 0}$, as $m \to \infty$, converges in distribution to the Poisson process on $[0, \infty)$ with rate $1/b$.*

We return to the discussion at the beginning of the subsection. The regime where $\lim_{m\to\infty} B_0^{(m)}(0)/m = b > 0$ and $A_0^{(m)}/m \to \infty$ is covered by the previous theorem. We need to change the roles of the colors and remark that the role of m as a scaling parameter is played now by $A_0^{(m)}$. The result that we obtain is that the process

$$\frac{1}{k}\left(B^{(m)}_{[tA_0^{(m)}/(bm)]} - B_0^{(m)}\right)_{t\geq 0}$$

converges in distribution, as $m \to \infty$, to the Poisson process on $[0, \infty)$ with rate 1.

Next, we look at regime c), i.e., in the case that at time 0 both black and white balls are of order m. In this case, the normalized process of the number of white balls has a non-random limit, which we determine, and then we study the fluctuations of the process around this limit.

Theorem 3 *Assume that $A_0^{(m)}, B_0^{(m)}$ are such that*

$$\lim_{m\to\infty} \frac{A_0^{(m)}}{m} = a, \lim_{m\to\infty} \frac{B_0^{(m)}}{m} = b,$$

where $a, b \in [0, \infty)$ are not both zero. Then the process $(A_{[mt]}^{(m)}/m)_{t\geq 0}$, as $m \to \infty$, converges in distribution to the deterministic process $X_t = \frac{a}{a+b}(a + b + kt), t \geq 0$.

The limit X is the same as in an urn in which we add at each step k white or black balls with corresponding probabilities $a/(a + b), b/(a + b)$, that is, irrespective of the composition of the urn at that time.

To determine the fluctuations of the process $(A_{[mt]}^{(m)}/m)_{t\geq 0}$ around its $m \to \infty$ limit, X, we let

$$C_t^{(m)} = \sqrt{m}\left(\frac{A_{[mt]}^{(m)}}{m} - X_t\right) \tag{2}$$

for all $m \in \mathbb{N}^+$ and $t \geq 0$.

Theorem 4 *Let $a, b \in [0, \infty)$, not both zero, $\theta_1, \theta_2 \in \mathbb{R}$, and assume that $A_0^{(m)} := [am + \theta_1\sqrt{m}], B_0^{(m)} = [bm + \theta_2\sqrt{m}]$ for all large $m \in \mathbb{N}$. Then the process $(C_t^{(m)})_{t\geq 0}$ converges in distribution, as $m \to \infty$, to the unique strong solution of the stochastic differential equation*

$$Y_0 = \theta_1, \tag{3}$$

$$dY_t = \frac{k}{a + b + kt}\left\{Y_t - \frac{a}{a + b}(\theta_1 + \theta_2)\right\} dt + k\frac{\sqrt{ab}}{a + b}\, dW_t, \tag{4}$$

which is

$$Y_t = \theta_1 + \frac{b\theta_1 - a\theta_2}{(a+b)^2}kt + k\frac{\sqrt{ab}}{a+b}(a+b+kt)\int_0^t \frac{1}{a+b+ks}\,dW_s. \tag{5}$$

W is a standard Brownian motion

In the previous theorem, it is possible to allow other kinds of deviations away from linearity (and not only of order $\sqrt{m}$) for the values of $A_0^{(m)}, B_0^{(m)}$. And then we get a diffusion limit if instead of (2) we look at the process

$$D_t^{(m)} = \sqrt{m}\left(\frac{A_{[mt]}^{(m)}}{m} - \frac{A_0^{(m)}}{m} - kt\frac{A_0^{(m)}}{A_0^{(m)} + B_0^{(m)}}\right) \tag{6}$$

for all $m \in \mathbb{N}^+$ and $t \geq 0$. More specifically, we have the following.

Theorem 5 *Assume that* $\lim_{m\to\infty} \frac{A_0^{(m)}}{m} = a, \frac{B_0^{(m)}}{m} = b$ *where* $a, b \in [0, \infty)$ *are not both zero. Then the process* $(D_t^{(m)})_{t\geq 0}$ *converges in distribution, as* $m \to \infty$*, to the unique strong solution of the stochastic differential equation*

$$V_0 = 0, \tag{7}$$

$$dV_t = \frac{kV_t}{a+b+kt}\,dt + k\frac{\sqrt{ab}}{a+b}\,dW_t, \tag{8}$$

which is

$$V_t = k\frac{\sqrt{ab}}{a+b}(a+b+kt)\int_0^t \frac{1}{a+b+ks}\,dW_s. \tag{9}$$

W is a standard Brownian motion

Remark. Functional central limit theorems for Pólya type urns have been proven with increasing generality in the works [6], [2], [8]. The major difference with our results is that in theirs the initial number of balls, $A_0^{(m)}, B_0^{(m)}$, is fixed (see however the last point in the list, concerning the recent work [3]). More specifically:

1) Gouet ([6]) studies urns with two colors (black and white) in the setting of Bagchi and Pal ([1]). According to that, when a white ball is drawn, we return it in the urn together with a white and b black balls, while if a black ball is drawn, we return it together with c white and d black. The numbers a, b, c, d are fixed integers (possibly negative), the number of balls added to the urn is fixed (that is $a + b = c + d$), and balls are drawn uniformly form the urn. The plain Pólya urn is not studied in that work because, according to the author, it has been studied by Heyde in [7]. However, for the Pólya urn, [7] discusses the central limit theorem and the law of the iterated logarithm. In any case, following the techniques of Heyde and Gouet one can prove the following. Assume for simplicity that $k = 1$ and let $L =: \lim_{n\to\infty} \frac{A_n}{n}$. The limit exists with probability one because of the martingale convergence theorem. Then

$$\left\{\sqrt{n}\left(t\frac{A_{[n/t]}}{n} - L\right)\right\}_{t\geq 0} \xrightarrow{d} \{W_{L'(1-L')t}\}_{t\geq 0}$$

as $n \to \infty$. W is a standard Brownian motion and L' is a random variable independent of W and having the same distribution as L. On the other hand, de-Finetti's theorem gives easily the more or less equivalent statement that, as $n \to \infty$,

$$\left\{\sqrt{n}\left(\frac{A_{[nt]}}{nt} - L\right)\right\}_{t\geq 0} \xrightarrow{d} \{W_{L'(1-L')/t}\}_{t\geq 0}$$

with W, L' as before.

2) Bai, Hu, and Zhang ([2]) work again in the setting of Bagchi and Pal, but now the numbers a, b, c, d depend on the order of the drawing and are random. The requirement that each time we add the same number of balls is relaxed.

3) Janson ([8]) considers urns with many colors, labeled $1, 2, \ldots, l$, where after each drawing, if we pick a ball of color i, we place in the urn balls of every color according to a random vector $(\xi_{i,1}, \ldots, \xi_{i,l})$ whose distribution depends on i ($\xi_{i,j}$ is the number of balls of color j that we add in the urn). Also, each ball is assigned a certain nonrandom activity that depends only on its color, and then the probability to pick a certain color at a drawing equals the ratio of the total of the activities of all balls of that color to the total of the activities of all balls present in the urn at that time. A restriction in that work is that there is a color i_0 so that starting the urn with just one ball and this ball has this color, there is positive probability to see in the future every other color. This excludes the classical Pólya urn that we study.

4) In [3], K. Borovkov studies a Pólya urn with $d+1$ colors, $1, 2, \ldots, d+1$, and proves convergence after appropriate scaling for the path $\{M([nt])\}_{t\in[0,1]}$, as $n \to \infty$, where

$$M(j) := (\xi_1(j), \xi_1(j) + \xi_2(j), \ldots, \sum_{i=1}^{d} \xi_i(j)) \in \mathbb{N}^d$$

and $\xi_i(j)$ is the number of balls of color i present in the urn at time j. The initial total number of balls in the urn is N and the author considers limits as $N, n \to \infty$ with $n/N \to c$ under the regimes $c = 0, c \in (0, \infty), c = \infty$. It assumes that at each drawing we add one ball, i.e., $k = 1$ in our notation.

Its relation to the present work is the following. We study only the case $d = 1$, and then $M(j) = \xi_1(j) = A_j^{(m)}$.

a) Theorems 1 and 2 are not covered by [3] because in Corollary 1 of [3] the changes $A_{[mt]}^{(m)} - A_0^{(m)}, A_{[tm/g_m]}^{(m)} - A_0^{(m)}$ are divided by $\sqrt{m}$ and $\sqrt{m/g_m}$ respectively (and then m is sent to infinity), while in Theorem 1 of [3], these changes are related to certain processes but with an error term of the order of $\log^2 m$. That is, in the scenarios of Theorems 1 and 2, the results of [3] are too rough to capture the birth process that we identify.

b) Theorems 4, 5 follow from Corollary 1(ii) in [3]. For example, under the assumptions of Theorem 4, the Corollary gives that

$$C_t^{(m)} - \theta_1 - \frac{b\theta_1 - a\theta_2}{(a+b)^2} t = H_t + o_{\mathbf{P}}(1)$$

for all $t \in [0,1]$, where the supremum of the error term, $o_{\mathbf{P}}(1)$, over $t \in [0,1]$ goes to zero in probability as $m \to \infty$, while the process H is Gaussian with continous paths, mean function zero, and covariance function

$$\mathrm{Cov}(H_s, H_t) = \frac{ab}{(a+b)^3} s(a+b+t)$$

for all $0 \le s \le t$. The term involving the stochastic integral in (5) also defines a Gaussian process with continuous paths and the same mean and covariance function as H. The justification for Theorem 5 is similar.

A preprint of the present work appeared in the arxiv on May 30, 2019, a few months before the preprint of [3].

2 Jump process limits. Proof of Theorems 1, 2

In the case of Theorem 1 we let $g_m := 1$ for all $m \in \mathbb{N}^+$, and for both theorems we let $v := v_m := m/g_m$ (we suppress the dependence of v on m). Our interest is in the sequence of the processes $(Z^{(m)})_{m \in \mathbb{N}^+}$ with

$$Z^{(m)}(t) = \frac{1}{k}(A^{(m)}_{[vt]} - A^{(m)}_0) \tag{10}$$

for all $t \ge 0$.

To show convergence in distribution, according to Theorem 7.8 of Chapter 3 of [5], it is enough to show that the sequence $(Z^{(m)})_{m\ge 1}$ is tight and its finite dimensional distributions converge. The description of the limiting process is determined on the way.

An easy argument shows that tightness follows from the convergence of the finite dimensional distributions becauce each $Z^{(m)}$ has non decreasing paths. It thus remains to establish the convergence of the finite dimensional distributions.

Notation: (i) For sequences $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}}$ with values in $\mathbb{R}$, we will say that they are asymptotically equivalent, and will write $a_n \sim b_n$ as $n \to \infty$, if $\lim_{n\to\infty} a_n/b_n = 1$. We use the same expressions for functions f, g defined in a neighborhood of ∞ and satisfy $\lim_{x\to\infty} f(x)/g(x) = 1$.
(ii) For $a \in \mathbb{C}$ and $k \in \mathbb{N}^+$, let

$$(a)_k := a(a-1)\cdots(a-k+1), \tag{11}$$

$$a^{(k)} := a(a+1)\cdots(a+k-1), \tag{12}$$

the falling and rising factorial respectively. Also let $(a)_0 := a^{(0)} := 1$.

2.1 Convergence of finite dimensional distributions

By definition, $Z^{(m)}(0) = 0 = Z(0)$ for all $m \in \mathbb{N}^+$.

Since, for each $m \in \mathbb{N}^+$, the process $Z^{(m)}$ is Markov taking values in $\mathbb{N}$ and non decreasing in time, our objective will have been accomplished if we show that the conditional probability

$$\mathbf{P}(Z^{(m)}(t_2) = k_2 | Z^{(m)}(t_1) = k_1) \tag{13}$$

converges as $m \to \infty$ for each $0 \le t_1 < t_2$ and nonnegative integers $k_1 \le k_2$. Define

$$n := [vt_2] - [vt_1], \tag{14}$$

$$x := k_2 - k_1, \tag{15}$$

$$\sigma := \frac{A_0^{(m)} + kk_1}{k}, \tag{16}$$

$$\tau := \frac{k[vt_1] - kk_1 + B_0^{(m)}}{k}. \tag{17}$$

Then, the above probability equals

$$\mathbf{P}(A^{(m)}_{[vt_2]} = kk_2 + a_0 | A^{(m)}_{[vt_1]} = kk_1 + a_0)$$
$$= \binom{n}{x} \frac{k\sigma(k\sigma + k)\cdots(k\sigma + (x-1)k)k\tau(k\tau + k)\cdots(k\tau + (n-x-1)k)}{(k\sigma + k\tau)(k\sigma + k\tau + k)\cdots(k\sigma + k\tau + (n-1)k)} \tag{18}$$

$$= \frac{(n)_x}{x!} \frac{\sigma^{(x)}\tau^{(n-x)}}{(\sigma+\tau)^{(n)}} = \frac{(n)_x}{x!} \sigma^{(x)} \frac{\Gamma(\tau + n - x)}{\Gamma(\tau)} \frac{\Gamma(\sigma+\tau)}{\Gamma(\sigma+\tau+n)}. \tag{19}$$

To compute the limit as $m \to \infty$ of (19), we will use Stirling's approximation for the Gamma function,

$$\Gamma(y) \sim \left(\frac{y}{e}\right)^y \sqrt{\frac{2\pi}{y}} \tag{20}$$

as $y \to \infty$, and its consequence

$$\Gamma(y + a) \sim \Gamma(y) y^a \tag{21}$$

as $y \to \infty$ for all $a \in \mathbb{R}$.

Proof ***(The computation for Theorem 1)*** Recall that $v = m$ in this theorem. Using (21) twice, with the role of a played by $-x$ and σ, we see that the last quantity in (19), for $m \to \infty$, is asymptotically equivalent to

$$\frac{(m(t_2 - t_1))^x}{x!} \sigma^{(x)} \tau^\sigma \frac{(\tau + n)^{-x}}{(\tau + n)^\sigma} \sim \frac{(m(t_2 - t_1))^x}{x!} \sigma^{(x)} \frac{\{m(t_1 + (b/k))\}^\sigma}{\{m(t_2 + (b/k))\}^{\sigma + x}}$$
$$= \frac{(t_2 - t_1)^x}{x!} \sigma^{(x)} \frac{\{t_1 + (b/k)\}^\sigma}{\{t_2 + (b/k)\}^{\sigma + x}}$$
$$= \binom{\sigma + x - 1}{x} \left(\frac{t_2 - t_1}{t_2 + (b/k)}\right)^x \left(1 - \frac{t_2 - t_1}{t_2 + (b/k)}\right)^\sigma. \tag{22}$$

[For reader's convenience, we remark that the asymptotics, as $m \to \infty$, of the relevant quantities are as follows: x, σ are constants while $n \sim (t_2 - t_1)m, \tau \sim (t_1 + (b/k))m$.]
Thus, as $m \to \infty$, the distribution of $\{Z^{(m)}(t_2) - Z^{(m)}(t_1)\}|Z^{(m)}(t_1) = k_1$ converges to the negative binomial distribution with parameters $\sigma, \frac{t_1+(b/k)}{t_2+(b/k)}$ [recall (1)].

*Proof (**The computation for Theorem 2**)* Using (20), we see that the last quantity in (19), for $m \to \infty$, is asymptotically equivalent to

$$
\begin{aligned}
&\frac{(m(t_2-t_1))^x}{x! g_m^x} \frac{g_m^x}{k^x} e^x \frac{(\tau+n-x)^{\tau+n-x}}{\tau^\tau} \frac{(\sigma+\tau)^{\sigma+\tau}}{(\sigma+\tau+n)^{\sigma+\tau+n}} \\
&\sim \frac{m^x(t_2-t_1)^x}{x! k^x} e^x (\tau+n-x)^{-x} \left(\frac{\tau+n-x}{\sigma+\tau+n}\right)^n \\
&\times \left(\frac{\sigma+\tau}{\sigma+\tau+n}\right)^\sigma \left(\frac{(\tau+n-x)(\sigma+\tau)}{\tau(\sigma+\tau+n)}\right)^\tau \\
&\sim \frac{m^x(t_2-t_1)^x}{x! k^x} e^x \tau^{-x} e^{-(t_2-t_1)/b} e^{-(t_2-t_1)/b} e^{-x+(t_2-t_1)/b} \\
&\sim \frac{1}{x!}\left(\frac{t_2-t_1}{b}\right)^x e^{-(t_2-t_1)/b}.
\end{aligned}
$$

[Here, the asymptotics, as $m \to \infty$, of the relevant quantities are as follows: x is constant while $n \sim (t_2 - t_1)m/g_m, \tau \sim (b/k)m, \sigma \sim g_m/k$.]
Thus, as $m \to \infty$, the distribution of

$$\{Z^{(m)}(t_2) - Z^{(m)}(t_1)\}|Z^{(m)}(t_1) = k_1$$

converges to the Poisson distribution with parameter $(t_2 - t_1)/b$.

2.2 Conclusion

It is clear from the form of the finite dimensional distributions that in both Theorems 1, 2 the limiting process Z is a pure birth process that does not explode in finite time. Its rate at the point $(t, j) \in [0, \infty) \times \mathbb{N}$ is

$$\lambda_{t,j} = \lim_{h \to 0^+} \frac{1}{h} \mathbf{P}(Z(t+h) = j+1 | Z(t) = j)$$

and is found as stated in the statement of each theorem.

3 Deterministic and diffusion limits. Proof of Theorems 3, 4, 5.

Theorems 3, 4, 5 are proved with the use of Theorem 7.1 in Chapter 8 of [4], which is concerned with convergence of time-homogeneous Markov chains to diffusions. The chains whose convergence is of interest to us are time inhomogeneous, but we reduce their study to the time-homogenous setting by considering for each such chain $\{Z_n\}_{n\in\mathbb{N}}$ the time homogeneous chain $\{(Z_n, n)\}_{n\in\mathbb{N}}$. The following consequence of the aforementioned theorem suffices for our purposes.

Corollary 1 *Assume that for each $m \in \mathbb{N}^+$, $(Z_n^{(m)})_{n\in\mathbb{N}}$ is a Markov chain in $\mathbb{R}$. For each $m \in \mathbb{N}^+$ and $n \in \mathbb{N}$, let $\Delta Z_n^{(m)} := Z_{n+1}^{(m)} - Z_n^{(m)}$ and*

$$\mu^{(m)}(x,n) := m\boldsymbol{E}(\Delta Z_n^{(m)} \boldsymbol{1}_{|\Delta Z_n^{(m)}|\le 1} | Z_n^{(m)} = x), \tag{23}$$

$$a^{(m)}(x,n) := m\boldsymbol{E}\{(\Delta Z_n^{(m)})^2 \boldsymbol{1}_{|\Delta Z_n^{(m)}|\le 1} | Z_n^{(m)} = x\} \tag{24}$$

for all $x \in \mathbb{R}$ with $\mathbf{P}(Z_n^{(m)} = x) > 0$. Also, for $R > 0$ and for the same m, n as above, let $A(m,n,R) := \{(x,n) : |x| \le R, n/m \le R, \mathbf{P}(Z_n^{(m)} = x) > 0\}$.

Assume that there are continuous functions $\mu : \mathbb{R} \times [0,\infty) \to \mathbb{R}, a : \mathbb{R} \times [0,\infty) \to [0,\infty)$, and $x_0 \in \mathbb{R}$ so that:
For every $R, \varepsilon > 0$, it holds

(i) $\sup_{(x,n)\in A(m,n,R)} |\mu^{(m)}(x,n) - \mu(x, n/m)| \to 0$ *as* $m \to \infty$,
(ii) $\sup_{(x,n)\in A(m,n,R)} |a^{(m)}(x,n) - a(x, n/m)| \to 0$ *as* $m \to \infty$,
(iii) $\sup_{(x,n)\in A(m,n,R)} m\mathbf{P}(|\Delta Z_n^{(m)}| \ge \varepsilon | Z_n^{(m)} = x) \to 0$ *as* $m \to \infty$,

and also

(iv) $Z_0^{(m)} \to x_0$ *as $m \to \infty$ with probability 1,*
(v) for each $x \in \mathbb{R}$, the stochastic differential equation

$$\begin{aligned} dZ_t &= \mu(Z_t, t)\, dt + \sqrt{a(Z_t, t)}\, dB_t, \\ Z_0 &= x, \end{aligned} \tag{25}$$

where B is a one dimensional Brownian motion, has a weak solution which is unique in distribution.

Then, the process $(Z_{[mt]}^{(m)})_{t\ge 0}$ converges in distribution to the weak solution of (25) *with $x = x_0$.*

Proof For each $m \in \mathbb{N}^+$, we consider the process $Y_n^{(m)} := (Z_n^{(m)}, n/m), n \in \mathbb{N}$, which is a time-homogeneous Markov chain with values in $\mathbb{R}^2$, and we apply Theorem 7.1 in Chapter 8 of [4] Conditions $(i), (ii), (iii)$ of that theorem follow from our conditions $(ii), (i), (iii)$ respectively, while condition (A) there translates to the requirement that the martingale problem for the functions μ and $\sqrt{a}$ is well posed, and this follows from condition (v).

The tool we will use in checking that condition (v) of the corollary is satisfied is the well known existence and uniqueness theorem for strong solutions of SDEs which requires that for all $T > 0$, the coefficients $\mu(x,t), \sqrt{a(x,t)}$ are Lipschitz in x uniformly for $t \in [0,T]$ and $\sup_{t\in[0,T]}\{|\mu(0,t)| + a(0,t)\} < \infty$ (e.g., Theorem 2.9 of Chapter 5 or [4]). The same conditions imply uniqueness in distribution.

3.1 Proof of Theorem 3

We will apply Corollary 1. For each $m \in \mathbb{N}^+$, consider the Markov chain $Z_n^{(m)} = \frac{A_n^{(m)}}{m}, n \in \mathbb{N}$. From any given state x of $Z_n^{(m)}$, the chain moves to either of $x+km^{-1}, x$ with corresponding probabilities $p(x,n,m), 1-p(x,n,m)$, where

$$p(x,n,m) := \frac{mx}{A_0^{(m)} + B_0^{(m)} + kn}. \tag{26}$$

In particular, for any $\varepsilon > 0$, is holds $|\Delta Z_n^{(m)}| < 1 \wedge \varepsilon$ for m large enough. Thus, condition (iii) of the corollary is satisfied trivially. Also, for large m, with the notation of the corollary, we have

$$\mu^{(m)}(x,n) = kp(x,n,m), \tag{27}$$

$$a^{(m)}(x,n) = \frac{k}{m} p(x,n,m). \tag{28}$$

And it is easy to see that conditions (i), (ii) are satisfied by the functions a, μ with $a(x,t) = 0$ and $\mu(x,t) = kp(x,t)$ where

$$p(x,t) := \frac{x}{a+b+kt}. \tag{29}$$

Now for each $x \in \mathbb{R}$, the equation

$$\begin{aligned} dZ_t &= kp(Z_t, t)\, dt, \\ Z_0 &= x, \end{aligned} \tag{30}$$

has a unique solution. Thus, Corollary 1 applies. In fact, (30) is a separable ordinary differential equation and its unique solution is the one given in the statement of the theorem.

3.2 Proof of Theorem 4

Call $\lambda := a/(a+b)$. For each $m \in \mathbb{N}^+$, consider the Markov chain

$$Z_n^{(m)} = \sqrt{m}\Big(\frac{A_n^{(m)}}{m} - X_{\frac{n}{m}}\Big), n \in \mathbb{N}.$$

From any given state x of $Z_n^{(m)}$, the chain moves to either of $x - km^{-1/2}\lambda, x + km^{-1/2}(1-\lambda)$ with corresponding probabilities

$$\frac{B_n^{(m)}}{A_n^{(m)} + B_n^{(m)}}, \frac{A_n^{(m)}}{A_n^{(m)} + B_n^{(m)}},$$

where

$$A_n^{(m)} = ma + \lambda kn + x\sqrt{m}, \tag{31}$$

$$B_n^{(m)} = A_0^{(m)} + B_0^{(m)} + kn - A_n^{(m)}. \tag{32}$$

Note that

$$A_0^{(m)} + B_0^{(m)} = (a+b)m + (\theta_1+\theta_2)\sqrt{m} + \delta_m, \tag{33}$$

with $\delta_m \in [0,2)$, and consequently

$$A_n^{(m)} = \lambda(A_n^{(m)} + B_n^{(m)}) + \sqrt{m}(x - \lambda(\theta_1+\theta_2)) - \lambda\delta_m. \tag{34}$$

Again, condition (iii) of Corollary 1 holds trivially, while $\lim_{m\to\infty} Z_0^{(m)} = \theta_1$ (condition (iv)). Then, for large m we have

$$\mu^{(m)}(x,n) = k\sqrt{m}\frac{(1-\lambda)A_n^{(m)} - \lambda B_n^{(m)}}{A_n^{(m)} + B_n^{(m)}} = k\frac{x - \lambda(\theta_1+\theta_2) - \lambda\frac{\delta_m}{\sqrt{m}}}{\frac{A_0^{(m)}+B_0^{(m)}}{m} + k\frac{n}{m}}, \tag{35}$$

$$a^{(m)}(x,n) = k^2\left(\lambda^2\frac{B_n^{(m)}}{A_n^{(m)}+B_n^{(m)}} + (1-\lambda)^2\frac{A_n^{(m)}}{A_n^{(m)}+B_n^{(m)}}\right) \tag{36}$$

$$= k^2\lambda(1-\lambda) + k^2(1-2\lambda)\frac{\sqrt{m}(x-\lambda(\theta_1+\theta_2)) - \lambda\delta_m}{A_n^{(m)}+B_n^{(m)}}. \tag{37}$$

It follows that conditions (i), (ii) are satisfied by the functions μ, a with

$$\mu(x,t) = \frac{k\{x - (\theta_1+\theta_2)\lambda\}}{a+b+kt}, \tag{38}$$

$$a(x,t) = \frac{k^2ab}{(a+b)^2}. \tag{39}$$

For each $x \in \mathbb{R}$, the stochastic differential equation

$$dY_t = \frac{k\{Y_t - (\theta_1+\theta_2)\lambda\}}{a+b+kt}\,dt + k\frac{\sqrt{ab}}{a+b}dW_t, \tag{40}$$

$$Y_0 = x, \tag{41}$$

where W is a standard Brownian motion, has a unique strong solution as the drift and diffusion coefficients are Lipschitz in Y_t and grow at most linearly in Y_t at infinity (both conditions uniformly in t). Thus, Corollary 1 applies and gives that the process $(Z_{[mt]}^{(m)})_{t\geq 0}$ converges in distribution, as $m\to\infty$, to the solution of (40), (41) with $x = \theta_1$. The same is true for $(C_t^{(m)})_{t\geq 0}$ because $\sup_{t\geq 0}|C_t^{(m)} - Z_{[mt]}^{(m)}| = \sup_{t\geq 0}\sqrt{m}\lambda k(t - [mt]/m) = \lambda k/\sqrt{m} \to 0$ as $m\to\infty$.

To solve the stochastic differential equation (40), (41), we set $U_t := \{Y_t - (\theta_1+\theta_2)\lambda\}/(a+b+kt)$. Itô's lemma gives that

$$dU_t = k\frac{\sqrt{ab}}{(a+b)}\frac{1}{a+b+kt}dW_t,$$

and since $U_0 = (b\theta_1 - a\theta_2)/(a+b)^2$, we get

$$U_t = \frac{b\theta_1 - a\theta_2}{(a+b)^2} + k\frac{\sqrt{ab}}{a+b}\int_0^t \frac{1}{a+b+ks}dW_s.$$

This gives (5).

3.3 Proof of Theorem 5

The proof is analogous to that of Theorem 4. Call $\lambda_m := A_0^{(m)}/(A_0^{(m)}+B_0^{(m)})$. For each $m \in \mathbb{N}^+$, consider the Markov chain

$$Z_n^{(m)} = \sqrt{m}\Big(\frac{A_n^{(m)}}{m} - \frac{A_0^{(m)}}{m} - \lambda_m k\frac{n}{m}\Big), n \in \mathbb{N}.$$

From any given state x of $Z_n^{(m)}$, the chain moves to either of $x - km^{-1/2}\lambda_m, x + km^{-1/2}(1-\lambda_m)$ with corresponding probabilities

$$\frac{B_n^{(m)}}{A_n^{(m)}+B_n^{(m)}}, \frac{A_n^{(m)}}{A_n^{(m)}+B_n^{(m)}},$$

where

$$A_n^{(m)} = A_0^{(m)} + \lambda_m kn + x\sqrt{m}, \tag{42}$$
$$B_n^{(m)} = A_0^{(m)} + B_0^{(m)} + kn - A_n^{(m)}. \tag{43}$$

Again, condition (iii) of Corollary 1 holds trivially, while $\lim_{m\to\infty} Z_0^{(m)} = 0$ (condition (iv)). Then, for large m we have

$$\mu^{(m)}(x,n) = k\sqrt{m}\frac{(1-\lambda_m)A_n^{(m)} - \lambda_m B_n^{(m)}}{A_n^{(m)}+B_n^{(m)}} = \frac{kx}{\frac{A_0^{(m)}+B_0^{(m)}}{m} + k\frac{n}{m}}, \tag{44}$$

$$a^{(m)}(x,n) = k^2\left(\lambda_m^2\frac{B_n^{(m)}}{A_n^{(m)}+B_n^{(m)}} + (1-\lambda_m)^2\frac{A_n^{(m)}}{A_n^{(m)}+B_n^{(m)}}\right) \tag{45}$$

$$= k^2\lambda_m(1-\lambda_m) + k^2(1-2\lambda_m)\frac{x\sqrt{m}}{A_n^{(m)}+B_n^{(m)}}. \tag{46}$$

Note now that $\lim_{m\to\infty}\lambda_m = a/(a+b)$ and $\lim_{m\to\infty}(A_n^{(m)}+B_n^{(m)})/m = a+b$. It follows that conditions (i), (ii) are satisfied by the functions μ, a with

$$\mu(x,t) = \frac{kx}{a+b+kt}, \tag{47}$$
$$a(x,t) = \frac{k^2ab}{(a+b)^2}. \tag{48}$$

For each $x \in \mathbb{R}$, the stochastic differential equation

$$dV_t = \frac{kY_t}{a+b+kt}\,dt + k\frac{\sqrt{ab}}{a+b}dW_t, \tag{49}$$
$$V_0 = x, \tag{50}$$

where W is a standard Brownian motion, has a unique strong solution as the drift and diffusion coefficients are Lipschitz in V_t and grow at most linearly in V_t at infinity (both conditions uniformly in t). Thus, Corollary 1 applies and gives that the process $(Z^{(m)}_{[mt]})_{t\geq 0}$ converges in distribution, as $m \to \infty$, to the solution of (49), (50) with $x = 0$. The same is true for $(D^{(m)}_t)_{t\geq 0}$ because $\sup_{t\geq 0} |D^{(m)}_t - Z^{(m)}_{[mt]}| \leq k/\sqrt{m} \to 0$ as $m \to \infty$.

Easily one finds that the solution of the stochastic differential equation (49), (50) with $x = 0$ is (9)

Acknowledgements We thank an anonymous referee for comments that improved significantly the exposition and for suggesting the question of Theorem 5.

The first author was supported by the Hellenic Foundation for Research and Innovation (H. F. R. I.) under the “First Call for H. F. R. I. Research Projects to support Facutly members and Researchers and the procurement of high-cost research equipment grant”, (Project Number: 1034).

References

1. Bagchi, Arunabha, and Asim K. Pal. “Asymptotic normality in the generalized Pólya-Eggenberger urn model, with an application to computer data structures.” SIAM Journal on Algebraic Discrete Methods 6, no. 3 (1985): 394-405.
2. Bai, Zhi-Dong, Feifang Hu, and Li-Xin Zhang. “Gaussian approximation theorems for urn models and their applications.” The Annals of Applied Probability 12, no. 4 (2002): 1149-1173.
3. Borovkov, Konstantin. “Gaussian process approximations for multicolor Pólya urn models.” Journal of Applied Probability 58, no. 1 (2021): 274-286.
4. Durrett, Richard. Stochastic calculus: a practical introduction. CRC press, 1996.
5. Ethier, Stewart N., and Thomas G. Kurtz. Markov processes: characterization and convergence. Vol. 282. John Wiley & Sons, 2009.
6. Gouet, Raul. “Martingale functional central limit theorems for a generalized Pólya urn.” The Annals of Probability (1993): 1624-1639.
7. Heyde, C. C. “On central limit and iterated logarithm supplements to the martingale convergence theorem.” Journal of Applied Probability 14, no. 4 (1977): 758-775.
8. Janson, Svante. “Functional limit theorems for multitype branching processes and generalized Pólya urns.” Stochastic Processes and their Applications 110, no. 2 (2004): 177-245.
9. Johnson, Norman Lloyd, and Samuel Kotz. “Urn models and their application; an approach to modern discrete probability theory.” (1977).
10. Mahmoud, Hosam. Pólya urn models. Chapman and Hall/CRC, 2008.