Functional limit theorems for the Pólya and q-Pólya urns

Dimitris Cheliotis* Dimitra Kouloumpou †

January 31, 2020

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. We also prove analogous results for the q-Pólya urn, which is an urn where, when picking a ball, the balls of one color have priority over those of the other.

1 Introduction and results

1.1 The models

The Pólya urn. This is the model where in an urn that has initially r white and s black balls we draw, successively, uniformly, and 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 its asymptotic behavior 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 $\text{Beta}(r/k, s/k)$. Our aim in this work is to examine whether the entire path $(A_n)_{n\geq 0}$ after appropriate natural transformations converges to a stochastic process.

Standard references for the theory and the applications of Pólya urn and related models are [12] and [15].

The q-Pólya urn. This is a q-analog of the Pólya urn (see [8], [13] for more on q-analogs) introduced in [14] and studied further in [4] (see also [5]). A q-analog of a mathematical object A is another object $A(q)$ so that when $q \to 1$, $A(q)$ "tends" to A. Take $q \in (0, \infty)\backslash\{1\}$. The q-analog of any $x \in \mathbb{C}$ is defined as

$$[x]_q := \frac{q^x - 1}{q - 1}. \tag{1.1}$$

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

†Techological Educational Institute of Pelloponnese, Antikalamos 24100, Kalamata, Greece. Email: dimkouloubou@teipel.gr.

2010 Mathematics Subject Classification: 60F17; 60K99; 60C05.

Keywords: Pólya urn, q-Pólya urn, q-calculus, functional limit theorems.

Note that $\lim_{q\to 1}[x]_q = x$. Now consider an urn that contains a finite number of white and black balls. We perform a sequence of additions of balls in the urn according to the following rule. If at a given time the urn contains r white and s black balls ($r, s \in \mathbb{N}, r+s>0$), then we add k white balls with probability

$$\mathbf{P}_q(\text{white}) = \frac{[r]_q}{[r+s]_q}. \tag{1.2}$$

Otherwise, we add k black balls, and this has probability

$$\mathbf{P}_q(\text{black}) = 1 - \mathbf{P}_q(\text{white}) = q^r \frac{[s]_q}{[r+s]_q}. \tag{1.3}$$

To understand how the q-Pólya urn works, it helps to realize the probabilities $\mathbf{P}_q(\text{white})$, $\mathbf{P}_q(\text{black})$ through a natural experiment.

If $q \in (0,1)$, then we put the balls in a line with the r white coming first and the s black following. To pick a ball, we go through the line, starting from the beginning and picking each ball with probability $1-q$ independently of what happened with the previous balls. If we finish the line without picking a ball, we start from the beginning. Once we pick a ball, we return it to its position together with k balls of the same color. Given these rules, the probability of picking a white ball is

$$(1-q^r)\sum_{j=0}^{\infty}(q^{r+s})^j = \frac{1-q^r}{1-q^{r+s}} = \frac{[r]_q}{[r+s]_q}, \tag{1.4}$$

which is (1.2), because before picking a white ball, we will go through the entire list a random number of times, say j, without picking any ball and then, going through the white balls, we pick one (probability $1-q^r$).

If $q>1$, we place in the line first the black balls and we go through the list picking each ball with probability $1-q^{-1}$. According to the above computation, the probability of picking a black ball is

$$\frac{[s]_{q^{-1}}}{[r+s]_{q^{-1}}} = q^r \frac{[s]_q}{[r+s]_q},$$

which is (1.3).

We extend the notion of drawing a ball from a q-Pólya urn to the case where exactly one of r, s is infinity. Then the probability to pick a white (resp. black) ball is determined again by (1.2) (resp. (1.3)), where this is understood as the limit of the right hand side as r or s goes to ∞. For example, assuming that $r=\infty$ and $s \in \mathbb{N}$, we have $\mathbf{P}_q(\text{white}) = 1$ if $q<1$ and $\mathbf{P}_q(\text{white}) = q^{-s}$ if $q>1$. Again these probabilities are realized through the experiment described above. Thus, we can run the process even if we start with an infinite number of balls from one color and finite from the other.

1.2 Pólya urn. Scaling limits

For the results of this section, 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 of $(A_n^{(m)})_{n\geq 0}$, we start with an initial number of balls that tends to infinity as $m \to \infty$. 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 by regime b) by changing the roles of the two colors.

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.

All processes appearing in this work 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 of each point of $[0,\infty)$. This space is endowed with the Skorokhod topology, 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

$$f(x) = \binom{x+\nu-1}{x} p^\nu (1-p)^x \tag{1.5}$$

for all $x \in \mathbb{N}$. When $\nu \in \mathbb{N}^+$, this is the distribution of the number of failures until we see 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 $X \sim NB(\nu, p)$.

Theorem 1.1. *Fix $w_0 \in \mathbb{N}^+$ and $b \geq 0$. If $A_0^{(m)} = w_0$ 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 $Z = (Z_t)_{t\geq 0}$ such that for all $0 \leq t_1 < t_2, j \in \mathbb{N}$, the random variable $Z(t_2) - Z(t_1)|Z(t_1) = j$ has distribution $NB\big(\frac{w_0}{k} + j, \frac{t_1+(b/k)}{t_2+(b/k)}\big)$. In particular, Z has rates $\lambda_{t,j} = (kj + w_0)/(kt + b)$ for all $(t, j) \in [0,\infty) \times \mathbb{N}$.*

Theorem 1.2. *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 $1/b$.*

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 1.3. *Assume that $A_0^{(m)}, B_0^{(m)}$ are such 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 $(A_{[mt]}^{(m)}/m)_{t\geq 0}$, 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 $(A^{(m)}_{[mt]}/m)_{t\geq 0}$ around its $m\to\infty$ limit, X, we let

$$C^{(m)}_t = \sqrt{m}\left(\frac{A^{(m)}_{[mt]}}{m} - X_t\right)$$

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

Theorem 1.4. *Let $a,b\in[0,\infty)$, not both zero, $\theta_1,\theta_2\in\mathbb{R}$, and assume that $A^{(m)}_0 := [am+\theta_1\sqrt{m}], B^{(m)}_0 = [bm+\theta_2\sqrt{m}]$ for all large $m\in\mathbb{N}$. Then the process $(C^{(m)}_t)_{t\geq 0}$ converges in distribution, as $m\to\infty$, to the unique strong solution of the stochastic differential equation*

$$Y_0 = \theta_1, \tag{1.6}$$

$$dY_t = \frac{k}{a+b+kt}\left\{Y_t - \frac{a}{a+b}(\theta_1+\theta_2)\right\}dt + k\frac{\sqrt{ab}}{a+b}\,dW_t, \tag{1.7}$$

which is

$$Y_t = \theta_1 + \frac{b\theta_1 - a\theta_2}{(a+b)^2}kt + k\frac{\sqrt{ab}}{a+b}(a+b+kt)\int_0^t \frac{1}{a+b+ks}\,dW_s. \tag{1.8}$$

W is a standard Brownian motion

Remark. Functional central limit theorems for Pólya type urns have been proven with increasing generality in the works [9], [2], [11]. The major difference with our results is that in theirs, the initial number of balls, $A^{(m)}_0, B^{(m)}_0$, is fixed. More specifically:

1) Gouet ([9]) 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 [10]. However, for the Pólya urn, [10] 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 ([11]) 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 the recent preprint [3], K. Borovkov studies a Pólya urn with $d+1$ colors, $1, 2, \ldots, d+1$, and identifies a scaling limit for the path $(M([nt]))_{t\geq 0}$, 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 $c = 0, c \in (0, \infty), c = \infty$. The case $c = 0$ is related to Theorem 1.2 while the case $c \in (0, \infty)$ is related to Theorems 1.1, 1.4 (see Corollary 1 in [3]).

1.3 q-Pólya urn. Basic results

We recall some notation from q-calculus (see [5], [13]). For $q \in (0, \infty)\backslash\{1\}, x \in \mathbb{C}, k \in \mathbb{N}^+$, we define

$$[x]_q := \frac{q^x - 1}{q - 1} \qquad \text{the } q\text{-number of } x, \tag{1.9}$$

$$[k]_q! := [k]_q[k-1]_q \cdots [1]_q \qquad \text{the } q\text{-factorial}, \tag{1.10}$$

$$[x]_{k,q} := [x]_q[x-1]_q \cdots [x-k+1]_q \qquad \text{the } q\text{-factorial of order } k, \tag{1.11}$$

$$\begin{bmatrix} x \\ k \end{bmatrix}_q := \frac{[x]_{k,q}}{[k]_q!} \qquad \text{the } q\text{-binomial coefficient} \tag{1.12}$$

$$(x; q)_\infty := \prod_{i=0}^{\infty}(1 - xq^i) \text{ when } q \in [0, 1) \qquad \text{the } q\text{-Pochhammer symbol}, \tag{1.13}$$

We extend these definitions in the case $k = 0$ by letting $[0]_q! = 1, [x]_{0,q} = 1$.

Now consider a q-Pólya urn that has initially r white and s black balls, where $r \in \mathbb{N} \cup \{\infty\}$ and $s \in \mathbb{N}$. Call X_n the number of drawings that give white ball in the first n drawings. Its distribution is specified by the following.

Fact 1: Let $\hat{r} := r/k$ and $\hat{s} := s/k$.

(i) If $r \in \mathbb{N}$, then the probability mass function of X_n is

$$\mathbf{P}\left(X_n = x\right) = q^{k(n-x)(\hat{r}+x)} \frac{\begin{bmatrix} -\hat{r} \\ x \end{bmatrix}_{q^{-k}} \begin{bmatrix} -\hat{s} \\ n-x \end{bmatrix}_{q^{-k}}}{\begin{bmatrix} -\hat{r}-\hat{s} \\ n \end{bmatrix}_{q^{-k}}} = q^{-sx} \frac{\begin{bmatrix} \hat{r}+x-1 \\ x \end{bmatrix}_{q^{-k}} \begin{bmatrix} \hat{s}+n-x-1 \\ n-x \end{bmatrix}_{q^{-k}}}{\begin{bmatrix} \hat{r}+\hat{s}+n-1 \\ n \end{bmatrix}_{q^{-k}}} \tag{1.14}$$

$$= q^{-kx(\hat{s}+n-x)} \frac{\begin{bmatrix} -\hat{r} \\ x \end{bmatrix}_{q^{k}} \begin{bmatrix} -\hat{s} \\ n-x \end{bmatrix}_{q^{k}}}{\begin{bmatrix} -\hat{r}-\hat{s} \\ n \end{bmatrix}_{q^{k}}} \tag{1.15}$$

for all $x \in \mathbb{N}$.

(ii) If $r = \infty$ and $q > 1$, then the probability mass function of X_n is

$$\mathbf{P}\left(X_n = x\right) = q^{-sx}(1-q^{-k})^{n-x} \begin{bmatrix} \hat{s}+n-x-1 \\ n-x \end{bmatrix}_{q^{-k}} \frac{[n]_{q^{-k}}!}{[x]_{q^{-k}}!} \tag{1.16}$$

for all $x \in \mathbb{N}$.

Relation (1.14) is (3.1) in [4] where it is proved through recursion. In Section 2 we give an alternative proof.

According to the experiment described in Section 1.1, the balls that are placed first in the line have an advantage to be picked (the white if $q \in (0,1)$, the black if $q > 1$). In fact, this leads to the extinction of drawings from the balls of the other color; there is a point after which the number of balls in the urn of that color stays fixed to a random number. In the next theorem, we identify the distribution of this number. We treat the case $q > 1$.

Theorem 1.5 (Extinction of the second color)**.** *Assume that $q > 1, r \in \mathbb{N} \cup \{\infty\}, s \in \mathbb{N}$. As $n \to \infty$, with probability one, $(X_n)_{n\geq 1}$ converges to a random variable X with values in $\mathbb{N}$ and probability mass function*

(i)

$$f(x) = q^{-sx} \begin{bmatrix} \frac{r}{k}+x-1 \\ x \end{bmatrix}_{q^{-k}} \frac{(q^{-s};q^{-k})_\infty}{(q^{-r-s};q^{-k})_\infty} \tag{1.17}$$

for all $x \in \mathbb{N}$ in the case $r \in \mathbb{N}$ and

(ii)

$$f(x) = \left(\frac{q^{-s}}{1-q^{-k}}\right)^x \frac{1}{[x]_{q^{-k}}!}(q^{-s};q^{-k})_\infty \tag{1.18}$$

for all $x \in \mathbb{N}$ in the case $r = \infty$.

When $r \in \mathbb{N}$ and $k|r$, X has the negative q-binomial distribution of the second kind with parameters $r/k, q^{-s}, q^{-k}$ (see §3.1 in [5] for its definition). When $r = \infty$, X has the Euler distribution with parameters $q^{-s}/(1-q^{-k}), q^{-k}$ (see §3.3 in [5] again).

1.4 q-Pólya urn. Scaling limits

As in Section 1.2, we consider an urn whose composition after n drawings is $A_n^{(m)}$ white and $B_n^{(m)}$ black balls. $m \in \mathbb{N}^+$ is a parameter. Our objective is to find limits of the entire path of the process $(A_n^{(m)})_{n\in\mathbb{N}}$ analogous to the ones of Section 1.2 for the Pólya urn. Assume that $q > 1$.

If we keep q fixed, nothing new appears because: (a) If $A_0^{(m)}, B_0^{(m)}$ are fixed for all m, then after some point we pick only black balls (Theorem 1.5(i)). (b) If $\lim_{m\to\infty} B_0^{(m)} = \infty$ then

the process converges to the one where we pick only black balls. (c) If $B_0^{(m)}$ is fixed for all m and $\lim_{m\to\infty} A_0^{(m)} = \infty$ then the process converges to the one where $r = \infty$ and again, after some point, we pick only black balls (Theorem 1.5(ii)).

Interesting limits appear once we take $q = q_m$ to depend on m and approach 1 as $m \to \infty$. We study two regimes for q_m. In the first, the distance of q_m from 1 is $\Theta(1/m)$ while in the second, the distance is $o(1/m)$.

1.4.1 The regime $q = 1 + \Theta(m^{-1})$

Assume that $q_m = c^{1/m}$ with $c > 1$.

Theorem 1.6. *Fix $w_0 \in \mathbb{N}^+$ and $b \geq 0$. If $A_0^{(m)} = w_0$ 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 with starting value 0 and such that for all $0 \leq t_1 < t_2, j \in \mathbb{N}$, the random variable $Z(t_2) - Z(t_1)|Z(t_1) = j$ has distribution $NB\big(\frac{w_0}{k} + j, \frac{1-c^{-b-kt_1}}{1-c^{-b-kt_1}}\big)$. In particular, Z has rates*

$$\lambda_{t,j} = \frac{w_0 + jk}{c^{b+kt} - 1} \log c \tag{1.19}$$

for all $(t, j) \in [0, \infty) \times \mathbb{N}$.

Theorem 1.7. *Assume that $A_0^{(m)} = g_m$ and $\lim_{m\to\infty} B_0^{(m)}/m = b$, where $b \in (0, \infty)$ and $g_m \in \mathbb{N}^+, g_m \to \infty, g_m = o(m)$ as $m \to \infty$. Then the process $(k^{-1}(A_{[tm/g_m]}^{(m)} - A_0^{(m)}))_{t\geq 0}$ converges in distribution, as $m \to \infty$, to the Poisson process on $[0, \infty)$ with rate*

$$\frac{\log c}{c^b - 1}. \tag{1.20}$$

Theorem 1.8. *Assume that $A_0^{(m)}, B_0^{(m)}$ are such that $\lim_{m\to\infty} A_0^{(m)}/m = a$, $\lim_{m\to\infty} B_0^{(m)}/m = b$, where $a, b \in [0, \infty)$ are not both zero. Then the process $\big(A_{[mt]}/m\big)_t \geq 0$ converges in distribution, as $m \to +\infty$, to the unique solution of the differential equation*

$$\hat{X}_0 = a, \tag{1.21}$$

$$d\hat{X}_t = k\frac{1 - c^{\hat{X}_t}}{1 - c^{a+b+kt}} dt, \tag{1.22}$$

which is

$$\hat{X}_t := a - \frac{1}{\log c} \log\left(\frac{c^b - 1 + c^{-kt}(1 - c^{-a})}{c^b - c^{-a}}\right). \tag{1.23}$$

As for the Pólya urn, we determine the fluctuations of the process $(A_{[mt]}^{(m)}/m)_{t\geq 0}$ around its $m \to \infty$ limit, $\hat{X}$. Let

$$\hat{C}_t^{(m)} = \sqrt{m}\left(\frac{A_{[mt]}^{(m)}}{m} - \hat{X}_t\right)$$

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

Theorem 1.9. *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 $(\hat{C}_t^{(m)})_{t\geq 0}$ converges in*

distribution, as $m \to \infty$, to the unique solution of the stochastic differential equation

$$\begin{aligned} \hat{Y}_0 &= \theta_1, \\ d\hat{Y}_t &= \frac{k \log c}{c^{a+b+kt}-1}\left\{\frac{(c^{a+b}-1)\hat{Y}_t - c^b(c^a-1)(\theta_1+\theta_2)}{c^b-1+c^{-kt}(1-c^{-a})}\right\} dt \\ &\quad + k\sqrt{(c^a-1)(c^b-1)}\frac{c^{(a+kt)/2}}{c^{a+b+kt}-c^{a+kt}+c^a-1}\, dW_t, \end{aligned} \tag{1.24}$$

which is

$$\begin{aligned} \hat{Y}_t =& \frac{c^{a+b+kt}-1}{c^{a+b+kt}-c^{a+kt}+c^a-1}\Bigg(\theta_1 - (\theta_1+\theta_2)\frac{c^{a+b}(c^a-1)}{c^{a+b}-1}\frac{c^{kt}-1}{c^{a+b+kt}-1} \\ &+ k\sqrt{(c^a-1)(c^b-1)}\int_0^t \frac{c^{(a+kt)/2}}{c^{a+b+kt}-1}\, dW_s\Bigg). \end{aligned} \tag{1.25}$$

W is a standard Brownian motion

1.4.2 The regime $q = 1 + o(m^{-1})$

In this regime, we let $q = q(m) := c^{\varepsilon_m/m}$ where $c > 1$ and $\varepsilon_m \to 0^+$ as $m \to \infty$. With computations analogous to those of the results of the previous subsection, it is easy to see that Theorems 1.1, 1.2 , 1.3, 1.4 hold exactly the same for the q-Pólya urn in this regime.

1.5 q-Pólya urn with many colors

In this paragraph, we give a q-analog for the Pólya urn with more than two colors. The way to do the generalization is inspired by the experiment we used in order to explain relation (1.2).

Let $l \in \mathbb{N}, l \geq 2$, and $q \in (0,1)$. Assume that we have an urn containing w_i balls of color i for each $i \in \{1, 2, \ldots, l\}$. To draw a ball from the urn, we do the following. We order the balls in a line, first those of color 1, then those of color 2, and so on. Then we visit the balls, one after the other, in the order that they have been placed, and we select each with probability $1-q$ independently of what happened with the previous balls. If we go through all balls without picking any, we repeat the same procedure starting from the beginning of the line. Once a ball is selected, the drawing is completed. We return the ball to its position together with another k of the same color. For each $i = 0, 1, \ldots, l$, let $s_i = \sum_{1 \leq j \leq i} w_j$. Notice that s_l is the total number of balls in the urn. Then, working as for (1.4), we see that

$$\mathbf{P}(\text{color } i \text{ is drawn}) = q^{s_{i-1}}\frac{1-q^{w_i}}{1-q^{s_l}} = \frac{q^{s_{i-1}}-q^{s_i}}{1-q^{s_l}} = q^{s_{i-1}}\frac{[w_i]_q}{[s_l]_q}. \tag{1.26}$$

Call p_i the number in the last display for all $i = 1, 2, \ldots, l$. Note that when $q \to 1$, p_i converges to w_i/s_l, which is the probability for the usual Pólya urn with l colors. It is clear that for any given $q \in (0, \infty)\backslash\{1\}$, the numbers $p_1, p_2, \ldots, p_l$ are non-negative and add to 1 (the second fraction in (1.26) shows this). We define then for this q the q-Pólya urn with colors $1, 2, \ldots, l$ to be the sequential procedure in which, at each step, we add k balls of a color picked randomly among $\{1, 2, \ldots, l\}$ so that the probability that this color is i is p_i .

When $q > 1$, these probabilities come out of the experiment described above but in which we place the balls in reverse order (that is, first those of color l, then those of color $l-1$, and so on) and we go through the list selecting each ball with probability $1 - q^{-1}$. It is then easy to see that the probability to pick a ball of color i is p_i.

Theorem 1.10. *Assume that $q \in (0,1)$ and that we start with $w_1, w_2, \ldots, w_l$ balls from colors 1, 2, $\ldots, l$ respectively, where $w_1, w_2, \ldots, w_l \in \mathbb{N}$ are not all zero. Call $X_{n,i}$ the number of times in the first n drawings that we picked color i. The probability mass function for the vector $(X_{n,2}, X_{n,3}, \ldots, X_{n,l})$ is*

$$\mathbf{P}\left(X_{n,2} = x_2, \ldots, X_{n,l} = x_l\right) = q^{\sum_{i=2}^{l} x_i \sum_{j=1}^{i-1}(w_j + k x_j)} \frac{\prod_{i=1}^{l} \left[{-\frac{w_i}{k} \atop x_i}\right]_{q^{-k}}}{\left[{-\frac{w_1+w_2\ldots+w_l}{k} \atop n}\right]_{q^{-k}}} \tag{1.27}$$

$$= \begin{bmatrix} n \\ x_1, x_2, \ldots, x_l \end{bmatrix}_{q^{-k}} \frac{q^{\sum_{i=2}^{l} x_i \sum_{j=1}^{i-1}(w_j + k x_j)} \prod_{i=1}^{l} \left[-\frac{w_i}{k}\right]_{x_i, q^{-k}}}{\left[-\frac{w_1+w_2+\ldots+w_l}{k}\right]_{n, q^{-k}}} \tag{1.28}$$

for all $x_2, \ldots, x_l \in \{0, 1, 2, \ldots, n\}$ with $x_2 + \cdot + x_l \leq n$, where $x_1 := n - \sum_{i=2}^{l} x_2$ and $\left[{n \atop x_1, x_2, \ldots, x_l}\right]_{q^{-k}} := \frac{[n]_{q^{-k}}!}{[x_1]_{q^{-k}}! \cdot \ldots \cdot [x_l]_{q^{-k}}!}$ is the q-multinomial coefficient.

It follows from Theorem 1.5 that when $q \in (0,1)$, after some random time, we will be picking only balls of color 1. So that the number of times that we pick each of the other colors $2, 3, \ldots, l$, say $X_2, X_3, \ldots, X_n$ are finite. We determine the joint distribution of these numbers.

Theorem 1.11. *Under the assumptions of Theorem 1.10, as $n \to +\infty$, with probability one, the vector $(X_{n,2}, X_{n,3}, \ldots, X_{n,l})$ converges to a random vector $(X_2, X_3, \ldots, X_l)$ with values in $\mathbb{N}^{l-1}$ and probability mass function*

$$f\left(x_2, x_3, \ldots, x_l\right) = q^{\sum_{i=2}^{l} x_i \sum_{j=1}^{i-1} w_j} \prod_{i=2}^{l} \begin{bmatrix} x_i + \frac{w_i}{k} - 1 \\ x_i \end{bmatrix}_{q^k} \frac{(q^{w_1}; q^k)_\infty}{(q^{w_1+\cdots+w_l}; q^k)_\infty} \tag{1.29}$$

for all $x_2, \ldots, x_l \in \mathbb{N}$.

Note that the random variables $X_2, \ldots, X_l$ are independent although $(X_{n,2}, X_{n,3}, \ldots, X_{n,l})$ are dependent.

Next, we look for a scaling limit for the path of the process. Assume that $c \in (0,1)$ and $q_m = c^{1/m}$. Let $A_{j,i}^{(m)}$ be the number of balls of color i in this urn after j drawings.

Theorem 1.12. *Let m be a positive integer and assume that in the q-Pólya urn with l different colors of balls it holds $\frac{1}{m}\left(A_{0,1}^{(m)}, A_{0,2}^{(m)}, \ldots, A_{0,l}^{(m)}\right) \overset{m\to\infty}{\to} (a_1, a_2, \ldots, a_l)$, where $a_1, \ldots, a_l \in [0, \infty)$ are not all zero. Set $\sigma_0 = 0$ and $\sigma_i := \sum_{j \leq i} a_j$ for all $i = 1, 2, \ldots, l$. Then the process $\frac{1}{m}\left(A_{[mt],1}^{(m)}, A_{[mt],2}^{(m)}, \ldots, A_{[mt],l}^{(m)}\right)_{t \geq 0}$ converges in distribution, as $m \to +\infty$, to $(X_{t,1}, X_{t,2}, \ldots, X_{t,l})_{t \geq 0}$ with*

$$X_{t,i} = a_i + \frac{1}{\log c} \log \frac{(1 - c^{\sigma_l + kt}) - c^{\sigma_{i-1}}(1 - c^{kt})}{(1 - c^{\sigma_l + kt}) - c^{\sigma_i}(1 - c^{kt})} \tag{1.30}$$

for all $i = 1, 2, \ldots, l$.

As in the case of two colors, we study the regime where $q_m = c^{\epsilon_m/m}$, with $c \in (0,1)$ and $\epsilon_m \to 0^+$.

Theorem 1.13. *Let m be a positive integer and assume that in the q-Pólya urn with l different colors of balls that $\frac{1}{m}\left(A_{0,1}^{(m)}, A_{0,2}^{(m)}, \ldots, A_{0,l}^{(m)}\right) \overset{m\to\infty}{\to} (a_1, a_2, \ldots, a_l)$, where $a_1, \ldots, a_l \in [0,\infty)$ are not all zero. Then the process $\frac{1}{m}\big(A_{[mt],1}^{(m)}, A_{[mt],2}^{(m)}, \ldots, A_{[mt],l}^{(m)}\big)_{t\geq 0}$ converges in distribution, as $m \to +\infty$, to $(X_t)_{t\geq 0}$ with*

$$X_t = \left(1 + \frac{kt}{a_1 + \cdots + a_l}\right)(a_1, a_2, \ldots, a_l) \tag{1.31}$$

for all $t \geq 0$.

Remark. Discussing this preprint with Prof. Ch. Charalambides, we were informed that he has also considered this q-Pólya urn with many colors in a work that is near completion.

Orientation. In Section 2, we prove Fact 1 and Theorem 1.5, which are basic results for the q-Pólya urn. Section 3 (Section 4) contains the proofs of the theorems for the Pólya and q-Pólya urns that give convergence to a jump process (to a continuous process). Finally, Section 5 contains the proofs for the results that refer to the q-Pólya urn with arbitrary, finite number of colors.

2 q-Pólya urn. Prevalence of a single color

In this section, we prove the claims of Section 1.3. Before doing so, we mention three properties of the q-binomial coefficient. For all $q \in (0,\infty)\backslash\{1\}, x \in \mathbb{C}, n, k \in \mathbb{N}$ with $k \leq n$ it holds

$$[-x]_q = -q^{-x}[x]_q, \tag{2.1}$$

$$\begin{bmatrix} -x \\ k \end{bmatrix}_q = (-1)^k q^{-k(k+2x-1)/2} \begin{bmatrix} x+k-1 \\ k \end{bmatrix}_q, \tag{2.2}$$

$$\begin{bmatrix} x \\ k \end{bmatrix}_{q^{-1}} = q^{-k(x-k)} \begin{bmatrix} x \\ k \end{bmatrix}_q, \tag{2.3}$$

$$\sum_{1\leq i_1<i_2<\cdots<i_k\leq n} q^{i_1+i_2+\cdots+i_k} = q^{\binom{k+1}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q. \tag{2.4}$$

The first is trivial, the second follows from the first, the third is easily shown, while the last is Theorem 6.1 in [13].

Proof of Fact 1. (i) The probability to get black balls exactly at the drawings $i_1 < i_2 < \cdots < i_{n-x}$ is

$$g(i_1, i_2, \ldots, i_{n-x}) = \frac{\prod_{j=0}^{x-1}[r+jk]_q \prod_{j=0}^{n-x-1}[s+jk]_q}{\prod_{j=0}^{n-1}[r+s+jk]_q} q^{\sum_{\nu=1}^{n-x} r+(i_\nu-\nu)k}. \tag{2.5}$$

To see this, note that, due to (1.2) and (1.3), the required probability would be equal to the above fraction if in (1.3) the term q^w were absent. This term appears whenever we draw a black ball. Now, when we draw the ν-th black ball, there are $r + (i_\nu - \nu)k$ white balls in the urn, and this explains the exponent of q in (2.5).

Since $[x+jk]_q = \frac{1-q^{x+jk}}{1-q} = [-\frac{x}{k}-j]_{q^{-k}}[-k]_q$ for all $x, j \in \mathbb{R}$, the fraction in (2.5) equals

$$\frac{[-\hat{r}]_{x,q^{-k}}[-\hat{s}]_{n-x,q^{-k}}}{[-\hat{r}-\hat{s}]_{n,q^{-k}}}. \tag{2.6}$$

Then

$$\sum_{1\le i_1<i_2<\cdots<i_{n-x}\le n} q^{\sum_{\nu=1}^{n-x} r+(i_\nu-\nu)k} \tag{2.7}$$

$$= q^{(n-x)r-k(n-x)(n-x+1)/2} \sum_{1\le i_1<i_2<\cdots<i_{n-x}\le n} (q^k)^{i_1+i_2+\cdots+i_{n-x}} \tag{2.8}$$

$$= q^{(n-x)r-k(n-x)(n-x+1)/2} q^{k\binom{n-x+1}{2}} \begin{bmatrix} n \\ x \end{bmatrix}_{q^k} \tag{2.9}$$

$$= q^{(n-x)r} q^{kx(n-x)} \begin{bmatrix} n \\ x \end{bmatrix}_{q^{-k}} = q^{k(n-x)(\hat{r}+x)} \begin{bmatrix} n \\ x \end{bmatrix}_{q^{-k}}. \tag{2.10}$$

The second equality follows from (2.4) and the equality $\left[{n \atop x}\right]_{q^k} = \left[{n \atop n-x}\right]_{q^k}$. The third, from (2.3). Thus, the sum $\sum_{1\le i_1<i_2<\cdots<i_{n-x}\le n} g(i_1, i_2, \ldots, i_{n-x})$ equals the first expression in (1.14). Then (1.15) and the second expression in (1.14) follow by using (2.3) and (2.2) respectively.

(ii) In this scenario, we take $r \to \infty$ in the last expression in (1.14). We will explain shortly why this gives the probability we want. Since $q^{-k} \in (0,1)$, we have $\lim_{t\to\infty}[t]_{q^{-k}} = (1-q^{-k})^{-1}$ and thus, for each $\nu \in \mathbb{N}$, it holds

$$\lim_{t\to\infty} \begin{bmatrix} t+\nu-1 \\ \nu \end{bmatrix}_{q^{-k}} = \frac{1}{[\nu]_{q^{-k}}!} \frac{1}{(1-q^{-k})^\nu}. \tag{2.11}$$

Applying this twice in the last expression in (1.14) (there $\hat{r} = r/k \to \infty$), we get as limit the right hand side of (1.16).

Now, to justify that passage to the limit $r \to \infty$ in (1.14) gives the required result, we argue as follows. For clarity, denote the probability $\mathbf{P}_q(\text{white})$ when there are w white and b black balls in the urn by $\mathbf{P}_q^{w,b}(\text{white})$. And when there are r white and s black balls in the urn in the beginning of the procedure, denote the probability of the event $X_n = x$ by $\mathbf{P}^{r,s}(X_n = x)$. It is clear that the probability $\mathbf{P}^{r,s}(X_n = x)$ is a continuous function (in fact, a polynomial) of the quantities

$$\mathbf{P}_q^{r+ki,s+kj}(\text{white}) : i = 0, 1, \ldots, x-1, j = 0, 1, \ldots, n-x-1,$$

for all values of $r \in \mathbb{N} \cup \{\infty\}, s \in \mathbb{N}$. In $\mathbf{P}^{\infty,s}(X_n = x)$, each such quantity, $\mathbf{P}_q^{\infty,m}(\text{white})$, equals $\lim_{r\to\infty} \mathbf{P}^{r,m}(\text{white})$. Thus, $\mathbf{P}^{\infty,s}(X_n = x) = \lim_{r\to\infty} \mathbf{P}^{r,s}(X_n = x)$. ■

Before proving Theorem 1.5, we give a simple argument that shows that eventually we will be picking only black balls. That is, the number $X := \lim_{n\to\infty} X_n$ of white balls drawn in an infinite sequence of drawings is finite. It is enough to show it in the case that $r = \infty$ and $s = 1$ since, by the experiment that realizes the q-Pólya urn, we have (using the notation from the proof of Fact 1 (ii))

$$\mathbf{P}^{r,s}(X = \infty) \le \mathbf{P}^{\infty,1}(X = \infty).$$

For each $n \in \mathbb{N}^+$, call E_n the event that at the n-th drawing we pick a white ball, B_n the number of black balls present in the urn after that drawing (also, $B_0 := 1$), and write $\hat{q} := 1/q$. Then $\mathbf{P}(E_n) = \mathbf{E}(\mathbf{P}(E_n|B_{n-1})) = \mathbf{E}(\hat{q}^{B_{n-1}})$. We will show that this decays exponentially with n. Indeed, since at every drawing there is probability at least $1 - \hat{q}$ to pick a black ball, we can construct in the same probability space the random variables $(B_n)_{n\geq 1}$ and $(Y_i)_{i\geq 1}$ so that the Y_i are i.i.d. with $Y_1 \sim \text{Bernoulli}(1 - \hat{q})$ and $B_n \geq 1 + k(Y_1 + \cdots + Y_n)$ for all $n \in \mathbb{N}^+$. Consequently,

$$\mathbf{P}(E_n) \leq \mathbf{E}(\hat{q}^{1+k(Y_1+\cdots+X_{n-1})}) = \hat{q}\{\mathbf{E}(\hat{q}^{kY_1})\}^{n-1}.$$

This implies that $\sum_{n=1}^{\infty} \mathbf{P}(E_n) < \infty$, and the first Borel-Cantelli lemma gives that $\mathbf{P}^{\infty,1}(X_\infty = \infty) = 0$.

Proof of Theorem 1.5. Since $(X_n)_{n\geq 1}$ is increasing, it converges to a random variable X with values in $\mathbb{N} \cup \{\infty\}$. In particular, it converges to this variable in distribution. Our aim is to take the limit as $n \to \infty$ in the last expression in (1.14) and in (1.16) in order to determine the distribution of X. Note that for $a \in \mathbb{R}$ and $\theta \in [0, 1)$ it is immediate that (recall (1.13) for the notation)

$$\lim_{n\to\infty} \begin{bmatrix} a+n \\ n \end{bmatrix}_\theta = \frac{(\theta^{a+1};\theta)_\infty}{(\theta;\theta)_\infty}. \tag{2.12}$$

(i) Taking $n \to \infty$ in the last expression in (1.14) and using (2.12), we get the required expression, (1.17), for f. Then relation (2.2) in [4] (or (8.1) in [13]) shows that $\sum_{x\in\mathbb{N}} f(x) = 1$, so that it is a probability mass function of a random variable X with values in $\mathbb{N}$.
(ii) This follows after taking limit in (1.16) and using (2.12) and $\lim_{n\to\infty}(1 - q^{-k})^n [n]_{q^{-k}}! = (q^{-k}; q^{-k})_\infty$. ■

3 Jump process limits. Proof of Theorems 1.1, 1.2, 1.6, 1.7

In the case of Theorems 1.1, 1.6, we let $g_m := 1$ for all $m \in \mathbb{N}^+$, and in all four theorems we let $v := v_m := m/g_m$. Our interest is in the sequence of the processes $(Z^{(m)})_{m\geq 1}$ with

$$Z^{(m)}(t) = \frac{1}{k}(A^{(m)}_{[vt]} - A^{(m)}_0) \tag{3.1}$$

for all $t \geq 0$.

We apply Theorem 7.8 in [7], that is, we show that the sequence $(Z^{(m)})_{m\geq 1}$ is tight and its finite dimensional distributions converge. Tightness gives that there is a subsequence of this sequence that converges in distribution to a process $Z = (Z_t)_{t\geq 0}$ with paths in the space $D_{\mathbb{R}}[0, \infty)$ of real valued functions on $[0, \infty)$ right continuous with left limits. Then tightness together with convergence of finite dimensional distributions shows that the whole sequence $(Z^{(m)})_{m\geq 1}$ converges in distribution to Z.

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

$$(a)_k := a(a-1)\cdots(a-k+1), \tag{3.2}$$

$$a^{(k)} := a(a+1)\cdots(a+k-1), \tag{3.3}$$

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

3.1 Convergence of finite dimensional distributions

Since for each $m \geq 1$ the process $Z^{(m)}$ is Markov taking values in $\mathbb{N}$ and increasing in time, it is enough to show that the conditional probability

$$\mathbf{P}(Z^{(m)}(t_2) = k_2 | Z^{(m)}(t_1) = k_1) \tag{3.4}$$

converges as $m \to \infty$ for each $0 \leq t_1 < t_2$ and nonnegative integers $k_1 \leq k_2$.
Consider first the case of Pólya urn and define

$$n := [vt_2] - [vt_1], \tag{3.5}$$

$$x := k_2 - k_1, \tag{3.6}$$

$$\sigma := \frac{A_0^{(m)} + kk_1}{k}, \tag{3.7}$$

$$\tau := \frac{k[vt_1] - kk_1 + B_0^{(m)}}{k}. \tag{3.8}$$

Then, the above probability equals

$$\begin{aligned}
&\mathbf{P}(A_{[vt_2]}^{(m)} = kk_2 + w_0 | A_{[vt_1]}^{(m)} = kk_1 + w_0) \\
&= \binom{n}{x} \frac{k\sigma(k\sigma+k)\cdots(k\sigma+(x-1)k)k\tau(k\tau+k)\cdots(k\tau+(n-x-1)k)}{(k\sigma+k\tau)(k\sigma+k\tau+k)\cdots(k\sigma+k\tau+(n-1)k)} \qquad (3.9) \\
&= \frac{(n)_x}{x!}\frac{\sigma^{(x)}\tau^{(n-x)}}{(\sigma+\tau)^{(n)}} = \frac{(n)_x}{x!}\sigma^{(x)}\frac{\Gamma(\tau+n-x)}{\Gamma(\tau)}\frac{\Gamma(\sigma+\tau)}{\Gamma(\sigma+\tau+n)}. \qquad (3.10)
\end{aligned}$$

To compute the limit as $m \to \infty$ of (3.10), we will use Stirling's approximation for the Gamma function,

$$\Gamma(y) \sim \left(\frac{y}{e}\right)^y \sqrt{\frac{2\pi}{y}} \tag{3.11}$$

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

$$\Gamma(y+a) \sim \Gamma(y)y^a \tag{3.12}$$

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

Theorem 1.1. Recall that $v = m$ in this theorem. Using (3.12) twice, with the role of a played by $-x$ and σ, we see that the last quantity in (3.10), for $m \to \infty$, is asymptotically equivalent to

$$\begin{aligned}
\frac{(m(t_2-t_1))^x}{x!}\sigma^{(x)}\tau^\sigma\frac{(\tau+n)^{-x}}{(\tau+n)^\sigma} &\sim \frac{(m(t_2-t_1))^x}{x!}\sigma^{(x)}\frac{\{m(t_1+(b/k))\}^\sigma}{\{m(t_2+(b/k))\}^{\sigma+x}} \\
&= \frac{(t_2-t_1)^x}{x!}\sigma^{(x)}\frac{\{t_1+(b/k)\}^\sigma}{\{t_2+(b/k)\}^{\sigma+x}} \\
&= \binom{\sigma+x-1}{x}\left(\frac{t_2-t_1}{t_2+(b/k)}\right)^x\left(1-\frac{t_2-t_1}{t_2+(b/k)}\right)^\sigma. \qquad (3.13)
\end{aligned}$$

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 (1.5)]. ■

Theorem 1.2. Using (3.11), we see that the last quantity in (3.10), 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 it was crucial that $b > 0$. 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 $(t_2 - t_1)/b$. ■

Now we treat the cases of Theorems 1.6, 1.7, which concern the q-Pólya urn. Define again n, x, σ, τ as in (3.5)-(3.8), and $r := q_m^{-k} = c^{-k/m}$. Then, the probability in (3.4), with the help of the last expression in (1.14), is computed as

$$\begin{aligned}
&r^{\tau x} \begin{bmatrix} \sigma+x-1 \\ x \end{bmatrix}_r \frac{\begin{bmatrix} \tau+n-x-1 \\ n-x \end{bmatrix}_r}{\begin{bmatrix} \sigma+\tau+n-1 \\ n \end{bmatrix}_r} \\
&= r^{\tau x} \begin{bmatrix} \sigma+x-1 \\ x \end{bmatrix}_r \Big(\prod_{i=n-x+1}^{n} (1-r^i) \Big) \frac{1}{\prod_{i=n-x}^{n-1}(1-r^{\tau+i})} \frac{[\tau+n-1]_{n,r}}{[\sigma+\tau+n-1]_{n,r}}.
\end{aligned} \tag{3.14}$$

The last ratio is

$$\prod_{i=0}^{n-1} \frac{1-r^{\tau+i}}{1-r^{\sigma+\tau+i}} = \prod_{i=0}^{n-1} \left(1 - (1-r^\sigma) r^\tau \frac{r^i}{1-r^{\sigma+\tau+i}}\right). \tag{3.15}$$

Denote by $1 - a_{m,i}$ the i-th term of the product. The logarithm of the product equals

$$-(1-r^\sigma) r^\tau \sum_{i=0}^{n-1} \frac{r^i}{1-r^{\sigma+\tau+i}} + o(1) \tag{3.16}$$

as $m \to \infty$. To justify this, note that $1 - r^\sigma \sim \frac{1}{m}(A_0^{(m)} + kk_1) \log c$ and $r^{\tau+i}/(1-r^{\sigma+\tau+i}) \le 1/(1-c^{-b})$ for all $i \in \mathbb{N}$. Thus, for all large m, $|a_{m,i}| < 1/2$ for all $i = 0, 1, \ldots, n-1$, and the error in approximating the logarithm of $1 - a_{m,i}$ by $-a_{m,i}$ is at most $|a_{m,i}|^2$ (by Taylor's expansion, we have $|\log(1-y) + y| \le |y|^2$ for all y with $|y| \le 1/2$). The sum of all errors is at most $n \max_{0 \le i < n} |a_{m,i}|^2$, which goes to zero as $m \to \infty$ because $1 - r^\sigma \sim C/n$ for some appropriate constant $C > 0$.

We will compute the limit of (3.14) as $m \to \infty$ under the assumptions of Theorems 1.6, 1.7.

Theorem 1.6. As $m \to \infty$, the first term of the product in (3.14) converges to $c^{-x(b+kt_1)}$. The q-binomial coefficient converges to $\binom{k^{-1}w_0+k_2-1}{k_2-k_1}$. The third term converges to $(1-c^{-k(t_2-t_1)})^x$, while the denominator of the fourth term converges to $(1-\rho_2)^x$, where we set $\rho_i := c^{-b-kt_i}$ for $i = 1, 2$. The expression preceding $o(1)$ in (3.16) is asymptotically equivalent to

$$-\frac{k}{m}\sigma(\log c)\rho_1 \sum_{i=0}^{n-1} \frac{c^{-ki/m}}{1-r^{\sigma+\tau}c^{-ki/m}} \tag{3.17}$$

$$= -\rho_1 k\sigma(\log c)\frac{1}{m}\sum_{i=0}^{n-1} \frac{c^{-ki/m}}{1-\rho_1 c^{-ki/m}} + o(1) \tag{3.18}$$

$$= -\rho_1 k\sigma \log c \int_0^{t_2-t_1} \frac{1}{c^{ky}-\rho_1}\, dy + o(1) = \sigma \log \frac{1-\rho_1}{1-\rho_2} + o(1). \tag{3.19}$$

The first equality is true because $\lim_{m\to\infty} r^{\sigma+\tau} = \rho_1$ and the function $x \mapsto c^{-ki/m}/(1 - xc^{-ki/m})$ has derivative bounded uniformly in i, m when x is confined to a compact subset of $[0, 1)$. Thus, the limit of (3.14), as $m \to \infty$, is

$$\binom{\sigma+x-1}{x}\left(\frac{\rho_1-\rho_2}{1-\rho_2}\right)^x \left(\frac{1-\rho_1}{1-\rho_2}\right)^\sigma, \tag{3.20}$$

which means that, 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 $\sigma, (1-\rho_1)/(1-\rho_2)$. ■

Theorem 1.7. Now the term $r^{\tau x}$ converges to c^{-xb}, while

$$\begin{bmatrix}\sigma+x-1\\ x\end{bmatrix}_r \Big(\prod_{i=n-x+1}^{n}(1-r^i)\Big) = \frac{\prod_{i=0}^{x-1}(1-r^{\sigma+i})}{\prod_{i=1}^{x}(1-r^i)}\Big(\prod_{i=n-x+1}^{n}(1-r^i)\Big) \tag{3.21}$$

$$\sim \frac{\prod_{i=0}^{x-1}(\sigma+i)}{\prod_{i=1}^{x} i}\frac{((t_2-t_1)k\log c)^x}{g_m^x} \sim \frac{1}{x!}((t_2-t_1)\log c)^x. \tag{3.22}$$

The denominator of the fourth term in (3.14) converges to $(1-c^{-b})^x$. The expression in (3.16) is asymptotically equivalent to

$$-r^\tau(1-r^\sigma)\sum_{i=0}^{n-1}\frac{r^i}{1-r^{\sigma+\tau+i}} \sim -c^{-b}\frac{g_m}{m}\log c\frac{n}{1-c^{-b}} \sim -\frac{\log c}{c^b-1}(t_2-t_1). \tag{3.23}$$

In the first $\sim$, we used the fact that the terms of the sum, as $m \to \infty$, converge uniformly in i to $(1-c^{-b})^{-1}$. Thus, the limit of (3.14), as $m \to \infty$, is

$$\frac{1}{x!}\left(\frac{\log c}{c^b-1}(t_2-t_1)\right)^x e^{-\frac{\log c}{c^b-1}(t_2-t_1)}, \tag{3.24}$$

which means that, 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 $\frac{t_2-t_1}{c^b-1}\log c$. ■

For use in the following section, we define

$$U(t_1, t_2, k_1, x) := \lim_{m\to\infty} \mathbf{P}(Z^{(m)}(t_2) = k_1 + x|Z^{(m)}(t_1) = k_1) \tag{3.25}$$

for all $0 \le t_1 \le t_2, k_1 \in \mathbb{N}, x \in \mathbb{N}$. The results of this section show that U as a function of $x \in \mathbb{N}$ is a probability mass function of an appropriate random variable with values in $\mathbb{N}$.

3.2 Tightness

We apply Corollary 7.4 of Chapter 3 in [7]. According to it, it is enough to show that

(i) For each $t \geq 0$, it holds $\lim_{R\to\infty} \overline{\lim}_{m\to\infty} \mathbf{P}(|Z^{(m)}(t)| \geq R) = 0$.

(ii) For each $T, \varepsilon > 0$, it holds $\lim_{\delta\to 0} \overline{\lim}_{m\to\infty} \mathbf{P}(w'(Z^{(m)}, \delta, T) \geq \varepsilon) = 0$.

Here, for any function $f : [0, \infty) \to \mathbb{R}$, we define

$$w'(f, \delta, T) := \inf_{\{t_i\}} \max_i \sup_{s,t\in[t_{i-1},t_i)} |f(s) - f(t)|,$$

where the infimum is over all partitions of the form $0 = t_0 < t_1 < \cdots t_{n-1} < T \leq t_n$ with $t_i - t_{i-1} > \delta$ for all $i = 1, 2, \ldots, n$.

The first requirement holds because $Z^{(m)}(t)$ converges in distribution as we showed in the previous subsection. The second requirement, since $Z^{(m)}$ is a jump process with jump sizes only 1, is equivalent to

$$\lim_{\delta\to 0^+} \overline{\lim_{m\to\infty}} \mathbf{P}\left(\begin{array}{c}\text{There are at least two jump times}\\ \text{of } Z^{(m)} \text{ in } [0,T] \text{ with distance} \leq \delta\end{array}\right) = 0. \tag{3.26}$$

Call $A_{m,\delta}$ the event inside the probability and for $j = 1, 2, \ldots, [T/\delta]$ define $I_j := ((j-1)\delta, (j+1)\delta]$. Then, for each $\ell \in \mathbb{N}$, the probability $\mathbf{P}(A_{m,\delta} \cap \{Z^{(m)}(T) \leq \ell\})$ is bounded above by

$$\sum_{j=1}^{[T/\delta]} \mathbf{P}(\{Z^{(m)}(T) \leq \ell\} \cap \{\text{There are at least two jump times of } Z^{(m)} \text{ in } I_j\}) \tag{3.27}$$

$$\leq \sum_{j=1}^{[T/\delta]} \mathbf{P}(\{Z^{(m)}(T) \leq \ell\} \cap \{Z^{(m)}((j+1)\delta) - Z^{(m)}((j-1)\delta) \geq 2\}) \tag{3.28}$$

$$\leq \sum_{j=1}^{[T/\delta]} \max_{0\leq\mu\leq\ell} \mathbf{P}(Z^{(m)}((j+1)\delta) - Z^{(m)}((j-1)\delta) \geq 2 | Z^{(m)}((j-1)\delta) = \mu). \tag{3.29}$$

The limit of the last quantity as $m \to \infty$, with the use of the function U of (3.25), is written as

$$\sum_{j=1}^{[T/\delta]} \max_{0\leq\mu\leq\ell} \sum_{x=2}^{\infty} U((j-1)\delta, (j+1)\delta, \mu, x) \leq \frac{T}{\delta} \max_{\substack{0\leq\mu\leq\ell\\ 1\leq j\leq[T/\delta]}} \sum_{x=2}^{\infty} U((j-1)\delta, (j+1)\delta, \mu, x). \tag{3.30}$$

Claim: The max in (3.30) is bounded above by $\delta^2 C(\ell, T)$ for an appropriate constant $C(\ell, T) \in (0, \infty)$ that does not depend on m or δ.

Assuming the claim and taking $m \to \infty$ in $\mathbf{P}(A_{m,\delta}) = \mathbf{P}(A_{m,\delta} \cap \{Z^{(m)}(T) \leq \ell\}) + \mathbf{P}(A_{m,\delta} \cap \{Z^{(m)}(T) > \ell\})$, we get

$$\overline{\lim_{m\to\infty}} \mathbf{P}(A_{m,\delta}) \leq \delta C(\ell, T) + \overline{\lim_{m\to\infty}} \mathbf{P}(\{Z^{(m)}(T) > \ell\}).$$

Now let $\varepsilon > 0$. Because of the validity of (i) in the tightness requirements, there is ℓ large enough so that the second term is $< \varepsilon$. Fixing this ℓ and taking $\delta \to 0$ in the inequality, we get (3.26).

Proof of the claim: We establish the above claim for each of the Theorems 1.1, 1.2, 1.6, 1.7. We use the following bounds. If X, Y are random variables with $X \sim \text{Poisson}(\lambda)$ and $Y \sim NB(\nu, p)$ then

$$\mathbf{P}(X \geq 2) \leq \lambda^2, \tag{3.31}$$

$$\mathbf{P}(Y \geq 2) \leq \frac{\nu(\nu+1)}{2}(1-p)^2. \tag{3.32}$$

The first inequality is elementary, while the second is true because the difference of the two sides

$$\mathbf{P}(Y \geq 2) - \frac{\nu(\nu+1)}{2}(1-p)^2 = 1 - p^\nu - rp^\nu(1-p) - \frac{\nu(\nu+1)}{2}(1-p)^2$$

is an increasing function of p in $[0,1]$ with value 0 at $p = 1$.

According to the results of Section 3.1, the sum after the max in (3.30) equals $\mathbf{P}(X \geq 2)$ where

$$X \sim \begin{cases} NB\big(\frac{w_0}{k} + \mu, \frac{t_1+(b/k)}{t_2+(b/k)}\big) & \text{for Theorem 1.1,} \\ \text{Poisson}\big(\frac{2\delta}{b}\big) & \text{for Theorem 1.2,} \\ NB\big(\frac{w_0}{k} + \mu, \frac{1-c^{-b-kt_1}}{1-c^{-b-kt_2}}\big) & \text{for Theorem 1.6,} \\ \text{Poisson}\big(2\delta\frac{\log c}{c^b-1}\big) & \text{for Theorem 1.7,} \end{cases} \tag{3.33}$$

and $t_1 := (j-1)\delta, t_2 := (j+1)\delta$. The claim then follows from (3.31) and (3.32).

3.3 Conclusion

It is clear from the form of the finite dimensional distributions that in all Theorems 1.1, 1.2, 1.6, 1.7 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.

4 Deterministic and diffusion limits. Proof of Theorems 1.3, 1.4, 1.8, 1.9

These theorems are proved with the use of Theorem 7.1 in Chapter 8 of [6], 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 4.1. *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*

$$\mu^{(m)}(x, n) := m\boldsymbol{E}(\Delta Z_n^{(m)} \boldsymbol{1}_{|\Delta Z_n^{(m)}| \leq 1} | Z_n^{(m)} = x), \tag{4.1}$$

$$a^{(m)}(x, n) := m\boldsymbol{E}((\Delta Z_n^{(m)})^2 \boldsymbol{1}_{|\Delta Z_n^{(m)}| \leq 1} | Z_n^{(m)} = x) \tag{4.2}$$

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 $\mu, a : \mathbb{R} \times [0,\infty) \to \mathbb{R}$ 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

$$\begin{aligned} dZ_t &= \mu(Z_t,t)\,dt + \sqrt{a(Z_t,t)}\,dB_t, \\ Z_0 &= x, \end{aligned} \tag{4.3}$$

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 (4.3) *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 [6] 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 [6]). The same conditions imply uniqueness in distribution.

4.1 Proof of Theorems 1.3, 1.8

We will apply Corollary 4.1. 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

$$p(x,n,m) := \begin{cases} \frac{mx}{A_0^{(m)} + B_0^{(m)} + kn} & \text{in the case of Theorem 1.3,} \\ \frac{1 - q_m^{mx}}{1 - q_m^{A_0^{(m)} + B_0^{(m)} + kn}} & \text{in the case of Theorem 1.8.} \end{cases}$$

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

$$\mu^{(m)}(x,n) = kp(x,n,m), \tag{4.4}$$

$$a^{(m)}(x,n) = \frac{k}{m}p(x,n,m). \tag{4.5}$$

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

$$p(x,t) := \begin{cases} \frac{x}{a+b+kt} & \text{in the case of Theorem 1.3,} \\ \frac{1-c^x}{1-c^{a+b+kt}} & \text{in the case of Theorem 1.8.} \end{cases} \tag{4.6}$$

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

$$\begin{aligned} dZ_t &= kp(Z_t,t)\,dt \\ Z_0 &= x, \end{aligned} \tag{4.7}$$

has a unique solution. Thus, Corollary 4.1 applies. For both theorems, 1.3 and 1.8, (4.7) is a separable ordinary differential equation and its unique solution is the one given in their statements.

4.2 Proof of Theorems 1.4, 1.9

Proof of Theorem 1.4. Call $\lambda := a/(a+b)$. For each $m \in \mathbb{N}^+$, consider the Markov chain

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

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

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

where

$$A_n^{(m)} = ma + \lambda kn + x\sqrt{m}, \tag{4.8}$$

$$B_n^{(m)} = A_0^{(m)} + B_0^{(m)} + kn - A_n^{(m)}. \tag{4.9}$$

Note that

$$A_0^{(m)} + B_0^{(m)} = (a+b)m + (\theta_1 + \theta_2)\sqrt{m} + \delta_m, \tag{4.10}$$

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

$$A_n^{(m)} = \lambda(A_n^{(m)} + B_n^{(m)}) + \sqrt{m}(x - \lambda(\theta_1 + \theta_2)) - \lambda\delta_m. \tag{4.11}$$

Again, condition (iii) of Corollary 4.1 holds trivially, while $\lim_{m\to\infty} Z_0^{(m)} = \theta_1$ (condition (iv)). Then, for large m we have

$$\mu^{(m)}(x,n) = k\sqrt{m}\frac{(1-\lambda)A_n^{(m)} - \lambda B_n^{(m)}}{A_n^{(m)} + B_n^{(m)}} = k\frac{x - \lambda(\theta_1+\theta_2) - \lambda\frac{\delta_m}{\sqrt{m}}}{\frac{A_0^{(m)}+B_0^{(m)}}{m} + k\frac{n}{m}}, \tag{4.12}$$

$$a^{(m)}(x,n) = k^2\left(\lambda^2\frac{B_n^{(m)}}{A_n^{(m)} + B_n^{(m)}} + (1-\lambda)^2\frac{A_n^{(m)}}{A_n^{(m)} + B_n^{(m)}}\right) \tag{4.13}$$

$$= k^2\lambda(1-\lambda) + k^2(1-2\lambda)\frac{\sqrt{m}(x-\lambda(\theta_1+\theta_2)) - \lambda\delta_m}{A_n^{(m)} + B_n^{(m)}}. \tag{4.14}$$

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

$$\mu(x,t) = \frac{k\{x - (\theta_1+\theta_2)\lambda\}}{a+b+kt}, \tag{4.15}$$

$$a(x,t) = \frac{k^2ab}{(a+b)^2}. \tag{4.16}$$

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

$$dY_t = \frac{k\{Y_t - (\theta_1+\theta_2)\lambda\}}{a+b+kt}\,dt + k\frac{\sqrt{ab}}{a+b}dW_t, \tag{4.17}$$

$$Y_0 = x, \tag{4.18}$$

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 4.1 applies and gives that the process $(Z_{[mt]}^{(m)})_{t\geq 0}$ converges in distribution, as $m \to \infty$, to the solution of (4.17), (4.18) with $x = \theta_1$.

The same is true for $(C_t^{(m)})_{t\geq 0}$ because $\sup_{t\geq 0}|C_t^{(m)} - Z_{[mt]}^{(m)}| \leq k/\sqrt{m}$. To solve the last stochastic differential equation, 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 (1.8). ■

Proof of Theorem 1.9. The proof is analogous to that of Theorem 1.4, only the algebra is a little more involved. For each $m \in \mathbb{N}^+$, consider the Markov chain

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

From any given state x of $Z_n^{(m)}$, the chain moves to either of

$$x + km^{-1/2} + \sqrt{m}(\hat{X}_{n/m} - \hat{X}_{(n+1)/m}), \tag{4.19}$$

$$x + \sqrt{m}(\hat{X}_{n/m} - \hat{X}_{(n+1)/m}) \tag{4.20}$$

with corresponding probabilities $p(x,n,m), 1-p(x,n,m)$, where

$$p(x,n,m) = \frac{[A_n^{(m)}]_{q_m}}{[A_0^{(m)} + B_0^{(m)} + kn]_{q_m}} \tag{4.21}$$

and

$$A_n^{(m)} = m\hat{X}_{n/m} + x\sqrt{m}, \tag{4.22}$$

$$B_n^{(m)} = A_0^{(m)} + B_0^{(m)} + kn - A_n^{(m)}. \tag{4.23}$$

For convenience, let $\Delta\hat{X}_{n/m} = \hat{X}_{(n+1)/m} - \hat{X}_{n/m}$. We compute

$$\mathbf{E}\left[\Delta Z_n^{(m)} | Z_n^{(m)} = x\right] = \frac{k}{\sqrt{m}} p(x,n,m) - \sqrt{m}\Delta\hat{X}_{n/m}, \tag{4.24}$$

$$\mathbf{E}\left[(\Delta Z_n^{(m)})^2 | Z_n^{(m)} = x\right] = \left(\frac{k^2}{m} - 2k\Delta\hat{X}_{n/m}\right) p(x,n,m) + m(\Delta\hat{X}_{n/m})^2. \tag{4.25}$$

The asymptotics of these expectations are as follows.

Claim: Fix $R > 0$. For n such that $\tau := n/m \le R$ and as $m \to \infty$, we have

$$(a) \quad \mathbf{E}\left[\Delta Z_n^{(m)} | Z_n^{(m)} = x\right]$$
$$= \frac{1}{m}\frac{k\log c}{c^{a+b+k\tau}-1}\left(c^{\hat{X}_\tau}x - \frac{(c^{\hat{X}_\tau}-1)c^{a+b+k\tau}}{c^{a+b+k\tau}-1}(\theta_1+\theta_2)\right) + O(\frac{1}{m^{3/2}}) \tag{4.26}$$

$$(b) \quad \mathbf{E}\left[(\Delta Z_n^{(m)})^2 | Z_n^{(m)} = x\right] = \frac{1}{m}k^2 g(\tau)\{1-g(\tau)\} + O(\frac{1}{m^{3/2}}) \tag{4.27}$$

where $g(t) := \frac{c^{\hat{X}_t}-1}{c^{a+b+kt}-1}$ for all $t \ge 0$.

Proof of the claim. We examine the asymptotics of $p(x,n,m)$ and $\Delta\hat{X}_{n/m}$. As $\tau \le R$ and $m \to \infty$, we have

$$p(x,n,m) = \frac{c^{\hat{X}_\tau + \frac{1}{\sqrt{m}x}} - 1}{c^{\frac{A_0^{(m)}+B_0^{(m)}}{m} + k\tau} - 1} = \frac{c^{\hat{X}_\tau + \frac{1}{\sqrt{m}x}} - 1}{c^{a+b+k\tau + \frac{\theta_1+\theta_2}{\sqrt{m}} + O(\frac{1}{m})} - 1} \tag{4.28}$$

$$= g(\tau) + \frac{\log c}{c^{a+b+k\tau}-1}\left(c^{\hat{X}_\tau}x - \frac{(c^{\hat{X}_\tau}-1)c^{a+b+k\tau}}{c^{a+b+k\tau}-1}(\theta_1+\theta_2)\right)\frac{1}{\sqrt{m}} \tag{4.29}$$
$$+ O(\frac{1}{m}).$$

The third equality follows from a Taylor's development. Also

$$\Delta\hat{X}_{n/m} = \hat{X}'_{n/m}\frac{1}{m} + O(m^{-2}) = kg(\tau)\frac{1}{m} + O(m^{-2}). \tag{4.30}$$

For $\hat{X}'$ we used the differential equation, (1.22), that $\hat{X}$ satisfies instead of the explicit expression for it. Substituting these expressions in (4.24), (4.25), we get the claim.

Relation (1.23) implies that $c^{\hat{X}_\tau} = (c^{a+b}-1)/\{c^b - 1 + c^{-k\tau}(1-c^{-a})\}$, and this gives that the parenthesis following $\frac{1}{m}$ in equation (a) of the claim above equals

$$\frac{(c^{a+b}-1)x - c^b(c^a-1)(\theta_1+\theta_2)}{c^b - 1 + c^{-k\tau}(1-c^{-a})} \tag{4.31}$$

and also that

$$g(\tau)\{1-g(\tau)\} = \frac{(c^a-1)(c^b-1)c^{a+k\tau}}{(c^{a+b+k\tau}-c^{a+k\tau}+c^a-1)^2}. \tag{4.32}$$

Thus, the claim implies that conditions (i), (ii) of Corollary 4.1 are satisfied by the functions

$$\mu(x,t) = \frac{k\log c}{c^{a+b+kt}-1}\left\{\frac{(c^{a+b}-1)x - c^b(c^a-1)(\theta_1+\theta_2)}{c^b-1+c^{-kt}(1-c^{-a})}\right\}, \tag{4.33}$$

$$a(x,t) = k^2(c^a-1)(c^b-1)\frac{c^{a+kt}}{(c^{a+b+kt}-c^{a+kt}+c^a-1)^2}. \tag{4.34}$$

Again, condition (iii) of the corollary holds trivially, while $\lim_{m\to\infty} Z_0^{(m)} = \theta_1$ (condition (iv)). Finally, for each $x \in \mathbb{R}$ and for the choice of μ, a above, equation (4.3) has a strong solution and uniqueness in distribution holds. Thus, the process $(Z_{[mt]}^{(m)})_{t\geq 0}$ converges, as $m \to \infty$, to the unique solution of the stochastic differential equation (1.24). And with analogous arguments as in Theorem 1.4, we get that the process $(\hat{C}_t^{(m)})_{t\geq 0}$ converges to the same process. To solve that equation, we remark that a solution of an equation of the form

$$dY_t = (\alpha(t)Y_t + \beta(t))dt + \gamma(t)dW_t \tag{4.35}$$

with $\alpha, \beta, \gamma : [0,\infty) \to \mathbb{R}$ continuous functions is given by

$$Y_t = e^{\int_0^t \alpha(s)\,ds}\left(Y_0 + \int_0^t \beta(s)e^{-\int_0^s \alpha(r)\,dr}\,ds + \int_0^t \gamma(s)e^{-\int_0^s \alpha(r)\,dr}\,dW_s\right). \tag{4.36}$$

[To discover the formula, we apply Itó's rule to $Y_t \exp\{-\int_0^t \alpha(s)\,ds\}$ and use (4.35).] Applying this formula for the values of α, β, γ dictated by (1.24) we arrive at (1.25). ■

5 Proofs for the q-Pólya urn with many colors

Proof of Theorem 1.10. First, the equality of the expressions in (1.27), (1.28) follows from the definition of the q-multinomial coefficient.

We will prove (1.27) by induction on l. When $l = 2$, (1.27) holds because of (1.14). In that relation, we have $x_1 = x, x_2 = n - x$. Assuming that (1.27) holds for $l \geq 2$ we will prove the case $l+1$. The probability $\mathbf{P}(X_{n,2} = x_2,..., X_{n,l+1} = x_{l+1})$ equals

$$\begin{aligned}
&\mathbf{P}\left(X_{n,3} = x_3, \ldots, X_{n,l+1} = x_{l+1}\right)\mathbf{P}(X_{n,2} = x_2 \mid X_{n,3} = x_3, \ldots, X_{n,l+1} = x_{l+1}) \qquad (5.1)\\
&= q^{\sum_{i=3}^{l+1} x_i \sum_{j=1}^{i-1}(w_j+kx_j)} \frac{\left[{-\frac{w_1+w_2}{k} \atop x_1+x_2}\right]_{q^{-k}} \prod_{i=3}^{l+1}\left[{-\frac{w_i}{k} \atop x_i}\right]_{q^{-k}}}{\left[{-\frac{w_1+\ldots w_{l+1}}{k} \atop n}\right]_{q^{-k}}} q^{x_2(w_1+kx_1)} \frac{\left[{-\frac{w_1}{k} \atop x_1}\right]_{q^{-k}}\left[{-\frac{w_2}{k} \atop x_2}\right]_{q^{-k}}}{\left[{-\frac{w_1+w_2}{k} \atop x_1+x_2}\right]_{q^{-k}}}\\
&= q^{\sum_{i=2}^{l+1} x_i \sum_{j=1}^{i-1}(w_j+kx_j)} \frac{\prod_{i=1}^{l+1}\left[{-\frac{w_i}{k} \atop x_i}\right]_{q^{-k}}}{\left[{-\frac{w_1+\ldots w_{l+1}}{k} \atop n}\right]_{q^{-k}}}.
\end{aligned}$$

This finishes the induction provided that we can justify these two equalities. The second is obvious, so we turn to the first. The first probability in (5.1) is specified by the inductive hypothesis. That is, given the description of the experiment, in computing this probability it is as if we merge colors 1 and 2 into one color which is placed in the line before the remaining

$l-1$ colors. This color has initially a_1+a_2 balls and we require that in the first n drawings we choose it x_1+x_2 times. The second probability in (5.1) is specified by the $l=2$ case of (1.27), which we know. More specifically, since the number of drawings from colors $3,4,\dots,l+1$ is given, it is as if we have an urn with just two colors $1,2$ that have initially w_1 and w_2 balls respectively. We do x_1+x_2 drawings with the usual rules for a q-Pólya urn, placing in a line all balls of color 1 before all balls of color 2, and we want to pick x_1 times color 1 and x_2 times color 2. ■

Proof of Theorem 1.11. The components of $(X_{n,2},X_{n,3},\dots,X_{n,l})$ are increasing in n, and from Theorem 1.5 we have that each of them has finite limit (we treat all colors $2,\dots,l$ as one color). Thus the convergence of the vector with probability one to a random vector with values is $\mathbb{N}^{l-1}$ follows. In particular, we also have convergence in distribution, and it remains to compute the distribution of the limit. Let $x_1:=n-(x_2+\cdots+x_l)$. Then the probability in (1.27) equals

$$\mathbf{P}\left(X_{n,2}=x_2,\dots,X_{n,l}=x_l\right)=q^{-\sum_{1\le i<j\le l}w_jx_i}\frac{\prod_{i=1}^{l}\left[{\frac{w_i}{k}+x_i-1 \atop x_i}\right]_{q^{-k}}}{\left[{\frac{\sum_{i=1}^{l}w_i}{k}+n-1 \atop n}\right]_{q^{-k}}} \tag{5.2}$$

$$=q^{\sum_{1\le j<i\le l}x_iw_j}\frac{\prod_{i=1}^{l}\left[{\frac{w_i}{k}+x_i-1 \atop x_i}\right]_{q^{k}}}{\left[{n+\frac{\sum_{i=1}^{l}w_i}{k}-1 \atop n}\right]_{q^{k}}} \tag{5.3}$$

$$=q^{\sum_{i=2}^{l}\left(x_i\sum_{j=1}^{i-1}w_j\right)}\left\{\prod_{i=2}^{l}\left[{\frac{w_i}{k}+x_i-1 \atop x_i}\right]_{q^{k}}\right\}\frac{\left[{x_1+\frac{w_1}{k}-1 \atop x_1}\right]_{q^{k}}}{\left[{n+\frac{\sum_{i=1}^{l}w_i}{k}-1 \atop n}\right]_{q^{k}}}. \tag{5.4}$$

In the first equality, we used (2.2) while in the second we used (2.3). When we take $n\to\infty$ in (5.4), the only terms involving n are those of the last fraction, and (2.12) determines their limit. Thus, the limit of (5.4) is found to be the function $f(x_2,\dots,x_l)$ in the statement of the theorem. ■

Proof of Theorem 1.12. For each $m\in\mathbb{N}^+$, we consider the discrete time-homogeneous Markov chain

$$Z_n^{(m)}:=\left(\frac{n}{m},\frac{A_{n,2}^{(m)}}{m},\frac{A_{n,3}^{(m)}}{m},\dots,\frac{A_{n,l}^{(m)}}{m}\right),n\in\mathbb{N}.$$

From any given state $(t,x):=(t,x_2,x_3,\dots,x_l)$ that $Z^{(m)}$ finds itself it moves to one of

$$\left(t+\frac{1}{m},x_2,\dots,x_i+\frac{1}{m},\dots,x_l\right),i=2,\dots,l,$$
$$\left(t+\frac{1}{m},x_2,\dots,x_i,\dots,x_l\right)$$

with corresponding probabilities

$$p_i(x_2,\dots,x_l,t,m)=q^{ms_{i-1}(t)}\frac{[mx_i]_q}{[ms_l(t)]_q},i=2,\dots,l, \tag{5.5}$$

$$p_1(x_2,\dots,x_l,t,m)=\frac{[mx_1(t)]_q}{[ms_l(t)]_q}, \tag{5.6}$$

where $s_i(t) = x_1(t) + \sum_{1<j\le i} x_j$ for $i \in \{1,2,\dots,l\}$ and $x_1(t) := m^{-1}\sum_{j=1}^{l} A_{0,j}^{(m)} + kt - \sum_{2\le j\le l} x_i$. These follow from (1.26) once we count the number of balls of each color present at the state (t,x). To do this, we note that $Z_n^{(m)} = (t,x)$ implies that $n = mt$ drawings have taken place so far, the total number of balls is $A_{0,1}^{(m)} + \cdots + A_{0,l}^{(m)} + kmt$, and the number of balls of color i, for $2 \le i \le l$, is mx_i. Thus, the number of balls of color 1 is $A_{0,1}^{(m)} + \cdots + A_{0,l}^{(m)} + kmt - m\sum_{2\le j\le l} x_i = mx_1(t)$. The required relations follow.

Let $x_1 := \lim_{m\to\infty} x_1(t) = \sigma_l + kt - \sum_{2\le j\le l} x_i$ and $s_i := \lim_{m\to\infty} s_i(t) = \sum_{1\le j\le i} x_i$ for all $i \in \{1,2,\dots,l\}$. Then, since $q = c^{1/m}$, for fixed $(t,x_2,\dots,x_l) \in [0,\infty)^l$ with $(x_2,\dots,x_l) \neq 0$, we have

$$\lim_{m\to\infty} p_i(x_2,\dots,x_l,t,m) = c^{s_{i-1}}\frac{[x_i]_c}{[s_l]_c} \tag{5.7}$$

for all $i = 2,\dots,l$. We also note the following.

$$Z_{n+1,1}^{(m)} - Z_{n,1}^{(m)} = \frac{1}{m}, \tag{5.8}$$

$$\mathbf{E}\left[Z_{n+1,i}^{(m)} - Z_{n,i}^{(m)} | Z_n^{(m)} = (t,x_2,\dots,x_l)\right] = \frac{k}{m} p_i(x_2,\dots,x_l,t,m), \tag{5.9}$$

$$\mathbf{E}\left[(Z_{n+1,i}^{(m)} - Z_{n,i}^{(m)})^2 | Z_n^{(m)} = (t,x_2,\dots,x_l)\right] = \frac{k^2}{m^2} p_i(x_2,\dots,x_l,t,m), \tag{5.10}$$

$$\mathbf{E}\left[(Z_{n+1,i}^{(m)} - Z_{n,i}^{(m)})(Z_{n+1,j}^{(m)} - Z_{n,j}^{(m)}) | Z_n^{(m)} = (t,x_2,\dots,x_l)\right] = 0 \tag{5.11}$$

for $i,j = 2,3,\dots,l$ with $i \neq j$.

Therefore, with similar arguments as in the proof of Theorem 1.3, as $m \to +\infty$, $(Z_{[mt]}^{(m)})_{t\ge 0}$ converges in distribution to Y, the solution of the ordinary differential equation

$$\begin{aligned} dY_t &= b(Y_t)dt, \\ Y_0 &= (0,a_2,\dots,a_l), \end{aligned} \tag{5.12}$$

where $b(t,x_2,\dots,x_l) = \left(1, b^{(2)}(t,x), b^{(3)}(t,x),\dots,b^{(l)}(t,x)\right)$ with

$$b^{(i)}(t,x) = kc^{s_{i-1}}\frac{[x_i]_c}{[s_l]_c}$$

for $i = 2,3,\dots,l$. Note that $s_l = \sigma_l + kt$ does not depend on x.

Since $A_{[mt],1}^{(m)} + A_{[mt],2}^{(m)} + \cdots + A_{[mt],l}^{(m)} = kmt + A_{0,1}^{(m)} + A_{0,2}^{(m)} + \cdots + A_{0,l}^{(m)}$, we get that the process $(A_{[mt],1}^{(m)}/m, A_{[mt],2}^{(m)}/m + \cdots + A_{[mt],l}^{(m)}/m)_{t\ge 0}$ converges in distribution to a process $(X_{t,1}, X_{t,2},\dots,X_{t,l})_{t\ge 0}$ so that $X_{t,1}+\cdots+X_{t,l} = a_1+a_2+\cdots+a_l+kt$, while the $X_{t,i}$, $i = 2,\dots,l$, satisfy the system

$$X_{t,i}' = kc^{\sigma_l + kt - \sum_{j=i}^{l} X_{t,i}}\frac{1-c^{X_{t,i}}}{1-c^{\sigma_l+kt}} \qquad \text{for all } t > 0, \tag{5.13}$$

$$X_{0,i} = a_i, \tag{5.14}$$

with $i = 2,3,\dots,l$. Letting $Z_{r,i} = c^{X_{\frac{1}{k\log c}\log r,i}}$ for all $r \in (0,1]$ and $i \in \{1,2,\dots,l\}$, we have for the $Z_{r,i}$, $i \in \{2,3,\dots,l\}$ the system

$$\frac{Z_{r,i}'}{1-Z_{r,i}} = \frac{\sigma_l}{1-\sigma_l r}\frac{1}{\prod_{i<j\le l} Z_{r,j}}, \tag{5.15}$$

$$Z_{1,i} = c^{a_i}. \tag{5.16}$$

In the case $i = l$, the empty product equals 1. It is now easy to prove by induction (starting from $i = l$ and going down to $i = 2$) that

$$Z_{r,i} = \frac{c^{\sigma_l - \sigma_{i-1}}(1 - c^{\sigma_l} r) - c^{\sigma_l}(1 - r)}{c^{\sigma_l - \sigma_i}(1 - c^{\sigma_l} r) - c^{\sigma_l}(1 - r)} \tag{5.17}$$

for all $r \in (0, 1]$. Since $Z_{r,1} Z_{r,2} \cdots Z_{r,l} = c^{\sigma_l} r$, we can check that (5.17) holds for $i = 1$ too. The fraction in (5.17) equals

$$c^{a_i} \frac{(1 - c^{\sigma_l} r) - c^{\sigma_{i-1}}(1 - r)}{(1 - c^{\sigma_l} r) - c^{\sigma_i}(1 - r)}. \tag{5.18}$$

Recalling that $X_{t,i} = (\log c)^{-1} \log Z_{c^{kt}}$, we get (1.30) for all $i \in \{1, 2, \ldots, l\}$. ■

Proof of Theorem 1.13. This is proved in the same way as Theorem 1.12. We keep the same notation as there. The only difference now is that $\lim_{m\to\infty} p_i(t, x_2, \ldots, x_l, m) = x_i/s_l$. As a consequence, the system of ordinary differential equations for the limit process $Y_t := (t, X_{t,2}, \ldots, X_{t,l})$ is (5.12) but with

$$b^{(i)}(t, x) = \frac{k x_i}{s_l}.$$

Recall that $s_l = \sigma_l + kt$. Thus, for $i = 2, 3, \ldots, l$, the process $X_{t,i}$ satisfies $X'_{t,i} = kX_{t,i}/(\sigma_l + kt), X_{0,i} = a_i$, which give immediately the last $l - 1$ coordinates of (1.31). The formula for the first coordinate follows from $X_{t,1} + X_{t,2} + \cdots + X_{t,l} = kt + \sigma_l$. ■

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." arXiv preprint arXiv:1912.09665 (2019).

[4] Charalambides, Ch A. "A q-Pólya urn model and the q-Pólya and inverse q-Pólya distributions." Journal of Statistical Planning and Inference 142, no. 1 (2012): 276-288.

[5] Charalambides, Charalambos A. Discrete q-distributions. John Wiley & Sons, 2016.

[6] Durrett, Richard. Stochastic calculus: a practical introduction. CRC press, 1996.

[7] Ethier, Stewart N., and Thomas G. Kurtz. Markov processes: characterization and convergence. Vol. 282. John Wiley & Sons, 2009.

[8] G. Gasper and M. Rahman. Basic Hypergeometric Series. Cambridge University Press, Cambridge, 1990

[9] Gouet, Raul. "Martingale functional central limit theorems for a generalized Pólya urn." The Annals of Probability (1993): 1624-1639.

[10] 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.

[11] Janson, Svante. "Functional limit theorems for multitype branching processes and generalized Pólya urns." Stochastic Processes and their Applications 110, no. 2 (2004): 177-245.

[12] Johnson, Norman Lloyd, and Samuel Kotz. "Urn models and their application; an approach to modern discrete probability theory." (1977).

[13] Kac, Victor, and Pokman Cheung. Quantum calculus. Springer Science & Business Media, 2001.

[14] Kupershmidt, Boris A. "q-probability: I. Basic discrete distributions." Journal of Nonlinear Mathematical Physics 7, no. 1 (2000): 73-93.

[15] Mahmoud, Hosam. Pólya urn models. Chapman and Hall/CRC, 2008.