Limit distributions for
Urn models with multiple drawings

Markus Kuba
Institut für Angewandte Mathematik und Naturwissenschaften

Joint work with
Hosam Mahmoud; Basile Morcrette

AofA'15
Strobl, Austria – 11.06.2015
Urn models

- Introduction
Urn models

- Introduction
Urn contains $W = n$ white and $B = m$ black balls.

- Every discrete time steps a ball is drawn at random:
  \[ p_{\text{white}} = \frac{n}{n+m}, \quad p_{\text{black}} = \frac{m}{n+m}. \]

- Color inspection:
  White - $a$ white and $b$ black balls are added/removed; Black - $c$ white and $d$ black balls are added/removed; $a, b, c, d \in \mathbb{Z}$.

- $2 \times 2$ ball replacement matrix
  \[ M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \]
Pólya-Eggenberger urns: Setup

Urn contains $W = n$ white and $B = m$ black balls.

- Every discrete time steps a ball is drawn at random:
  \[ p_{\text{white}} = \frac{n}{n+m}, \quad p_{\text{black}} = \frac{m}{n+m}. \]

- Color inspection:
  - White - $a$ white and $b$ black balls are added/removed; Black -
  - $c$ white and $d$ black balls are added/removed; $a, b, c, d \in \mathbb{Z}$.

- $2 \times 2$ ball replacement matrix
  \[ M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \]
Pólya-Eggenberger urns: Setup

Urn contains \( W = n \) white and \( B = m \) black balls.

- Every discrete time steps a ball is drawn at random:
  \[
  p_{\text{white}} = \frac{n}{n+m}, \quad p_{\text{black}} = \frac{m}{n+m}.
  \]

- Color inspection:
  White - \( a \) white and \( b \) black balls are added/removed; Black - \( c \) white and \( d \) black balls are added/removed; \( a, b, c, d \in \mathbb{Z} \).

- 2 × 2 ball replacement matrix
  \[
  M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.
  \]
Urn contains $W = n$ white and $B = m$ black balls.

- Every discrete time steps a ball is drawn at random:
  
  \[ p_{\text{white}} = \frac{n}{n+m}, \quad p_{\text{black}} = \frac{m}{n+m}. \]

- Color inspection:
  - White - $a$ white and $b$ black balls are added/removed; Black - $c$ white and $d$ black balls are added/removed; $a, b, c, d \in \mathbb{Z}$.

- $2 \times 2$ ball replacement matrix

\[
M = \begin{pmatrix}
  a & b \\
  c & d
\end{pmatrix}.
\]
Pólya-Eggenberger urns: Setup

Urn contains $W = n$ white and $B = m$ black balls.

- Every discrete time steps a ball is drawn at random:
  $$p_{\text{white}} = \frac{n}{n+m}, \quad p_{\text{black}} = \frac{m}{n+m}.$$

- Color inspection:
  White - $\alpha$ white and $\beta$ black balls are added/removed; Black - $\gamma$ white and $\delta$ black balls are added/removed; $\alpha, \beta, \gamma, \delta \in \mathbb{Z}$.

- $2 \times 2$ ball replacement matrix
  $$M = \begin{pmatrix} W & B \\ \{W\} & \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \end{pmatrix}. $$
Pólya-Eggenberger urns: examples

Pólya urn: \( M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \).

Beta limit law; Normal limit for martingale tail sum
Hall and Heyde; Gräbel.

Triangular urns: \( M = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \)
with \( a, d, \in \mathbb{N} \) and \( b \in \mathbb{N}_0 \).

Generalized Mittag-Leffler limit law
Janson; Goldschmidt; Mahmoud.
Pólya-Eggenberger urns: examples

**Pólya urn:** \( M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \).

Beta limit law; Normal limit for martingale tail sum
Hall and Heyde; **Grübel.**

**Triangular urns:** \( M = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \)

with \( a, d, \in \mathbb{N} \) and \( b \in \mathbb{N}_0 \).

Generalized Mittag-Leffler limit law
Janson; **Goldschmidt; Mahmoud.**
Pólya urn: \( M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \).

Beta limit law; Normal limit for martingale tail sum
Hall and Heyde; Grübel.

Triangular urns: \( M = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \)

with \( a, d, \in \mathbb{N} \) and \( b \in \mathbb{N}_0 \).

Generalized Mittag-Leffler limit law
Janson; Goldschmidt; Mahmoud.
Pólya-Eggenberger urns: examples

Ball replacement matrix $M = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$

Initial configuration: $n = 7$ yellow (white), $m = 6$ black balls
Pólya-Eggenberger urns: examples

Ball replacement matrix $M = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$

Initial configuration: $n = 7$ yellow (white), $m = 6$ black balls

$p_{\text{yellow}} = \frac{7}{13}$

$p_{\text{black}} = \frac{6}{13}$
Pólya-Eggenberger urns: examples

Ball replacement matrix $M = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$

Initial configuration: $n = 7$ yellow (white), $m = 6$ black balls
Ball replacement matrix $M = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$

Initial configuration: $n = 7$ yellow (white), $m = 6$ black balls
Pólya-Eggenberger urns: examples

Ball replacement matrix $M = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$

Initial configuration: $n = 7$ yellow (white), $m = 6$ black balls
Ball replacement matrix $M = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$

Initial configuration: $n = 7$ yellow (white), $m = 6$ black balls

$p_{\text{yellow}} = \frac{9}{16}$

$p_{\text{black}} = \frac{7}{16}$
Pólya-Eggenberger urns: examples

Ball replacement matrix \( M = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \)

Initial configuration: \( n = 7 \) yellow (white), \( m = 6 \) black balls
Pólya-Eggenberger urns: examples

Ball replacement matrix $M = \begin{pmatrix} \frac{2}{1} \frac{1}{-1} \end{pmatrix}$

Initial configuration: $n = 7$ yellow (white), $m = 6$ black balls
Pólya-Eggenberger urns: examples

Ball replacement matrix $M = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}$

Initial configuration: $n = 7$ yellow (white), $m = 6$ black balls
We make the following assumptions:

1. **Tenable urns**: process of drawing and adding/removing balls can be continued *ad infinitum*.

2. **Balance condition**:

\[ M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \]

with

\[ \sigma = a + b = c + d \geq 1. \]

Consequently: total number of balls \( T_n \) is non-random:

\[ T_n = W_n + B_n = W_0 + T_0 + n \cdot \sigma. \]

3. **Two colors**: white and black balls.
Main question: Given $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, and start with $W_0 \in \mathbb{N}_0$ white and $B_0 \in \mathbb{N}_0$ black balls: configuration $W_n (B_n)$ after $n$ draws?

Main task: Analyse

$$W_n^{(d)} = W_{n-1} + a \cdot I_n(\{W\}) + c \cdot I_n(\{B\}), \quad n \geq 1,$$

where

$$\mathbb{P}(I_n(\{W\}) = 1 | \mathcal{F}_{n-1}) = \frac{W_{n-1}}{W_{n-1} + B_{n-1}} = \frac{W_{n-1}}{T_{n-1}}$$

and

$$\mathbb{P}(I_n(\{B\}) = 1 | \mathcal{F}_{n-1}) = \frac{B_{n-1}}{W_{n-1} + B_{n-1}} = \frac{B_{n-1}}{T_{n-1}}.$$
Pólya-Eggenberger urns: existing works

**Analytic combinatorics/Symbolic methods:** generating functions, method of moments, etc.

Flajolet, Gabarró and Pekari; Brennan and Prodinger; Stadje; Dumas, Flajolet and Puyhaubert; Hwang, K. and Panholzer; Morcrette; Mahmoud and Morcrette; . . .

**Probabilistic methods:** stochastic processes, martingales, contraction method, etc.

Bagci and Pal, Neininger and Knape; Kingman; Kingman and Volkov; Mahmoud; Johnson, Kotz and Mahmoud; Pittel; Janson; Pouyanne; Chauvin, Mailer and Pouyanne; Mailer; Chauvin, Gardy, Pouyanne and Ton-That; Grübel; Higueras, Moler, Plo and San Miguel; Tsukiji and Mahmoud; Chen and Wei; Renlund; . . .
Analytic combinatorics/Symbolic methods: generating functions, method of moments, etc.

Flajolet, Gabarró and Pekari; Brennan and Prodinger; Stadje; Dumas, Flajolet and Puyhaubert; Hwang, K. and Panholzer; Morcrette; Mahmoud and Morcrette; ... 

Probabilistic methods: stochastic processes, martingales, contraction method, etc.

Bagci and Pal, Neiningers and Knape; Kingman; Kingman and Volkov; Mahmoud; Johnson, Kotz and Mahmoud; Pittel; Janson; Pouyanne; Chauvin, Mailler and Pouyanne; Mailler; Chauvin, Gardy, Pouyanne and Ton-That; Grübel; Higueras, Moler, Plo and San Miguel; Tsukiji and Mahmoud; Chen and Wei; Renlund; ...

Balanced urn models: partial differential equations; History counting approach
Pólya-Eggenberger urns: existing works

Analytic combinatorics/Symbolic methods: generating functions, method of moments, etc.

Flajolet, Gabarró and Pekari; Brennan and Prodinger; Stadje; Dumas, Flajolet and Puyhaubert; Hwang, K. and Panholzer; Morcrette; Mahmoud and Morcrette; . . .

Probabilistic methods: stochastic processes, martingales, contraction method, etc.

Bagci and Pal, Neininger and Knape; Kingman; Kingman and Volkov; Mahmoud; Johnson, Kotz and Mahmoud; Pittel; Janson; Pouyanne; Chauvin, Mailler and Pouyanne; Mailler; Chauvin, Gardy, Pouyanne and Ton-That; Grübel; Higuera, Moler, Plo and San Miguel; Tsukiji and Mahmoud; Chen and Wei; Renlund; . . .
Analytic combinatorics/Symbolic methods: generating functions, method of moments, etc.

Flajolet, Gabarró and Pekari; Brennan and Prodinger; Stadje; Dumas, Flajolet and Puyhaubert; Hwang, K. and Panholzer; Morcrette; Mahmoud and Morcrette; . . .

Probabilistic methods: stochastic processes, martingales, contraction method, etc.

Bagci and Pal, Neininger and Knape; Kingman; Kingman and Volkov; Mahmoud; Johnson, Kotz and Mahmoud; Pittel; Janson; Pouyanne; Chauvin, Mailer and Pouyanne; Mailer; Chauvin, Gardy, Pouyanne and Ton-That; Grübel; Hígueras, Moler, Plo and San Miguel; Tsukiji and Mahmoud; Chen and Wei; Renlund; . . .

unbalanced \(r\)-color urns: general results; also covers 2-color triangular urn models.
Pólya-Eggenberger urns: existing works

**Analytic combinatorics/Symbolic methods:** generating functions, method of moments, etc.

Flajolet, Gabarró and Pekari; Brennan and Prodinger; Stadje; Dumas, Flajolet and Puyhaubert; Hwang, K. and Panholzer; Morcrette; Mahmoud and Morcrette; . . .

**Probabilistic methods:** stochastic processes, martingales, contraction method, etc.

Bagci and Pal, Neininger and Knape; Kingman; Kingman and Volkov; Mahmoud; Johnson, Kotz and Mahmoud; Pittel; Janson; Pouyanne; Chauvin, Mailer and Pouyanne; Mailer; Chauvin, Gardy, Pouyanne and Ton-That; Grübel; Higueras, Moler, Plo and San Miguel; Tsukiji and Mahmoud; Chen and Wei; Renlund; . . .
Pólya-Eggenberger urns: existing works

**Analytic combinatorics/Symbolic methods:** generating functions, method of moments, etc.

Flajolet, Gabarró and Pekari; Brennan and Prodinger; Stadje; Dumas, Flajolet and Puyhaubert; Hwang, K. and Panholzer; Morcrette; Mahmoud and Morcrette; ... 

**Probabilistic methods:** stochastic processes, martingales, contraction method, etc.

Bagci and Pal, Neininger and Knape; Kingman; Kingman and Volkov; Mahmoud; Johnson, Kotz and Mahmoud; Pittel; Janson; Pouyanne; Chauvin, Mailer and Pouyanne; Mailer; Chauvin, Gardy, Pouyanne and Ton-That; Grübel; Higueras, Moler, Plo and San Miguel; Tsukiji and Mahmoud; Chen and Wei; Renlund; ... 

contraction method for balanced urns; smoothing equations equations for large urns.
Analytic combinatorics/Symbolic methods: generating functions, method of moments, etc.

Flajolet, Gabarró and Pekari; Brennan and Prodinger; Stadje; Dumas, Flajolet and Puyhaubert; Hwang, K. and Panholzer; Morcrette; Mahmoud and Morcrette; . . .

Probabilistic methods: stochastic processes, martingales, contraction method, etc.

Bagci and Pal, Neininger and Knape; Kingman; Kingman and Volkov; Mahmoud; Johnson, Kotz and Mahmoud; Pittel; Janson; Pouyanne; Chauvin, Mailer and Pouyanne; Mailer; Chauvin, Gardy, Pouyanne and Ton-That; Grübel; Higuera, Moler, Plo and San Miguel; Tsukiji and Mahmoud; Chen and Wei; Renlund; . . .
Fundamental result

Theorem (Trichotomy of limit laws - Janson)

For a balanced two-color urn model $M = \begin{pmatrix} a_0 & b_0 \\ a_1 & b_1 \end{pmatrix}$ let $\Lambda = \frac{\lambda_1}{\lambda_2}$ denote the ratio of the two eigenvalues of $M$.

1. for $\Lambda \leq \frac{1}{2}$: $\frac{W_n - E(W_n)}{\sqrt{V(W_n)}} \to \mathcal{N}(0, 1)$.

2. for $\Lambda > \frac{1}{2}$: $\frac{W_n - E(W_n)}{n^{\Lambda}} \to \mathcal{L}$.

3. for $b_0 \cdot a_1 = 0$ the urn is triangular and $\frac{W_n}{n^{\Lambda}} \to T$. 
...New problems.


Remark 4.5: "Mahmoud has initiated the study of urn models where several, say 2, balls are drawn at the same time, and balls are added depending on the drawn combination of types. It may be possible to study such models too by the methods of this paper, first considering the corresponding continuous time model, but we have not pursued this. (This case is substantially more complicated than the standard case treated here; for example, the continuous time model will explode in finite time.)"
Pólya-Eggenberger urns

**Remark**

**PHILIPPE FLAJOLET, Talk by Basile Morcrette and Nicolas Pouyanne, 2011**

*Conference - Philippe Flajolet and Analytic Combinatorics, slides;*

---

**Urn with multiple draws:**

*The Bernoulli-Laplace process*

- **Chamber A**
  - $a$ white
  - $b$ black

- **Chamber B**
  - $b$ white
  - $a$ black

- **One urn**
  - $a$ white, $b$ black

---

\[
x^ay^b \xrightarrow{\mathcal{G}} a^2x^{-1}y^1x^ay^b + b^2x^1y^{-1}x^ay^b + 2abx^0y^0x^ay^b
\]

\[
\Theta_u := u\partial_u \text{ (pick & replace a ball)}
\]

\[
\mathcal{G} := x^{-1}y^1\Theta_x^2 + x^1y^{-1}\Theta_y^2 + 2x^0y^0\Theta_x\Theta_y
\]

\[
H_n(x, y) := \mathcal{G}^n \circ x^aoy^{b_0}
\]

\[
\mathbb{P} \{ (a_0, b_0) \rightsquigarrow (a, b) \} = \frac{1}{N^{2n}} [x^ay^b] H_n(x, y)
\]

[Flajolet, July 2010]
Pólya-Eggenberger urns

Remark

PHILIPPE FLAJOLET, Talk by Basile Morcrette and Nicolas Pouyanne, 2011
Conference - Philippe Flajolet and Analytic Combinatorics, slides;

Urns with multiple draws:
the Bernoulli-Laplace process

\[
\begin{align*}
&\text{Chamber A} & \text{Chamber B} & \text{One urn} \\
& a \text{ white} & b \text{ white} & a \text{ white, } b \text{ black} \\
& b \text{ black} & & \\
\end{align*}
\]

\[
\begin{pmatrix}
W \\
B \\
WW \\
BW \\
BB
\end{pmatrix} =
\begin{pmatrix}
-1 & 1 \\
0 & 0 \\
1 & -1
\end{pmatrix}
\]

\[
a + b = N
\]

\[
x^a y^b \xrightarrow{G} a^2 x^{-1} y^1 x^a y^b + b^2 x^1 y^{-1} x^a y^b + 2ab x^0 y^0 x^a y^b
\]

\[
\Theta_u := u \partial_u \text{ (pick & replace a ball)}
\]

\[
G := x^{-1} y^1 \Theta_x^2 + x^1 y^{-1} \Theta_y^2 + 2 x^0 y^0 \Theta_x \Theta_y
\]

\[
H_n(x, y) := G^n \circ x^{a_0} y^{b_0}
\]

\[
P \{ (a_0, b_0) \sim (a, b) \} = \frac{1}{N^{2n}} \left[ x^a y^b \right] H_n(x, y)
\]

[Flajolet, July 2010]
Urn models

Multiple drawings
- linear affine class
Urn models

Multiple drawings
- linear affine class
Previously: Urn contains \( w \) white and \( b \) black balls. \( M = \begin{pmatrix} a_0 & b_0 \\ a_1 & b_1 \end{pmatrix} \)

Draw at random a single ball: \( p_{\text{white}} = \frac{w}{w+b}, \ p_{\text{black}} = \frac{b}{w+b} \).

Multiple drawings: We draw \( m \geq 1 \) balls in order to obtain our sample. Depending on the drawn multiset of white/black balls we add/remove balls according to the \((m+1) \times 2\) ball replacement matrix

\[
M = \begin{pmatrix}
  a_0 & b_0 \\
  a_1 & b_1 \\
  \vdots & \vdots \\
  a_{m-1} & b_{m-1} \\
  a_m & b_m \\
\end{pmatrix}
\]

Balance: \( \sigma = a_0 + b_0 = \cdots = a_m + b_m \).
We sample $m = 2$ balls.

Let

$$M = \begin{pmatrix}
\{WW\} & W \\
\{WB\} & B \\
\{BB\} & B
\end{pmatrix} = \begin{pmatrix}
a_0 & b_0 \\
a_1 & b_1 \\
a_2 & b_2
\end{pmatrix}.$$
We sample $m = 2$ balls.

$$M = \begin{pmatrix}
\{WW\} & \{WB\} & \{BB\} \\
W & a_0 & b_0 \\
B & a_1 & b_1 \\
& a_2 & b_2
\end{pmatrix}.$$
We sample $m = 2$ balls.

\[
M = \begin{pmatrix}
\{WW\} & \{WB\} & \{BB\} \\
a_0 & a_1 & a_2 \\
b_0 & b_1 & b_2
\end{pmatrix}.
\]
We sample $m = 2$ balls.

$$M = \begin{pmatrix} \{WW\} & \{WB\} & \{BB\} \\ a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \end{pmatrix}.$$
We sample $m = 2$ balls.

$$
M = \begin{pmatrix}
\{WW\} & W & B \\
\{WB\} & a_0 & b_0 \\
\{BB\} & a_1 & b_1 \\
& a_2 & b_2
\end{pmatrix}.
$$
We sample $m = 2$ balls.

$$M = \begin{pmatrix} a_0 & b_0 \\ a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}.$$
We sample $m = 2$ balls.

\[ M = \begin{pmatrix} \{WW\} & \{WB\} & \{BB\} \\ a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \end{pmatrix}. \]
Sampling schemes - unordered samples:
Urn contains \( w \) white and \( b \) black balls.

**Sampling without replacement**

\[
p\{k \text{ times white, } (m-k) \text{ times black}\} = \frac{1}{(b + w)^m} \binom{m}{k} w^k b^{m-k},
\]

with \( x^s = x(x-1) \ldots (x-s+1) \).

**Sampling with replacement,**

\[
p\{k \text{ times white, } (m-k) \text{ times black}\} = \frac{1}{(b + w)^m} \binom{m}{k} w^k b^{m-k}.
\]
Distributions: We consider balanced urn models and study the number of white balls $W_n$:

$$W_n \overset{(d)}{=} W_{n-1} + \sum_{k=0}^{m} a_{m-k} \cdot \mathbb{I}_n(W^kB^{m-k}), \quad n \geq 1.$$ 

$\mathbb{I}_n(W^kB^{m-k})$, indicator variables $k$ white and $m-k$ black balls. 

$\mathcal{F}_{n-1}$ sigma-field generated by first $n-1$ draws 

For sampling without replacement $\mathbb{P}\{\mathbb{I}_n(W^kB^{m-k}) = 1|\mathcal{F}_{n-1}\}$ is given by

$$\frac{\binom{W_{n-1}}{k} \binom{B_{n-1}}{m-k}}{\binom{T_{n-1}}{m}} = \frac{\binom{W_{n-1}}{k} \binom{T_{n-1} - W_{n-1}}{m-k}}{\binom{T_{n-1}}{m}}$$

and for sampling with replacement

$$\binom{m}{k} \frac{W^k_{n-1}B^{m-k}_{n-1}}{T^m_{n-1}} = \binom{m}{k} \frac{W^k_{n-1}(T_{n-1} - W_{n-1})^{m-k}}{T^m_{n-1}}.$$
Distributions: We consider balanced urn models and study the number of white balls $W_n$:

$$W_n \overset{(d)}{=} W_{n-1} + \sum_{k=0}^{m} a_{m-k} \cdot \mathbb{I}_n(W^kB^{m-k}), \quad n \geq 1.$$  

$\mathbb{I}_n(W^kB^{m-k})$, indicator variables $k$ white and $m-k$ black balls.

$\mathcal{F}_{n-1}$ sigma-field generated by first $n-1$ draws.

For sampling without replacement $\mathbb{P}\{\mathbb{I}_n(W^kB^{m-k}) = 1 | \mathcal{F}_{n-1}\}$ is given by

$$\binom{W_{n-1}}{k} \binom{B_{n-1}}{m-k} \binom{T_{n-1}}{m-k} = \binom{W_{n-1}}{k} \binom{T_{n-1} - W_{n-1}}{m-k} \binom{T_{n-1}}{m}$$

and for sampling with replacement

$$\binom{m}{k} \frac{W^k_{n-1}B^{m-k}_{n-1}}{T^m_{n-1}} = \binom{m}{k} \frac{W^k_{n-1}(T_{n-1} - W_{n-1})^{m-k}}{T^m_{n-1}}.$$
Distributions: We consider balanced urn models and study the number of white balls $W_n$:

$$W_n \overset{(d)}{=} W_{n-1} + \sum_{k=0}^{m} a_{m-k} \cdot \mathbb{I}_n(W^k B^{m-k}), \quad n \geq 1.$$ 

$\mathbb{I}_n(W^k B^{m-k})$, indicator variables $k$ white and $m-k$ black balls.

$\mathcal{F}_{n-1}$ sigma-field generated by first $n-1$ draws

For sampling without replacement $P\{\mathbb{I}_n(W^k B^{m-k}) = 1 | \mathcal{F}_{n-1}\}$ is given by

$$\binom{W_{n-1}}{k} \binom{B_{n-1}}{m-k} \binom{T_{n-1} - W_{n-1}}{m-k} \binom{T_{n-1}}{T_{n-1} - W_{n-1}}.$$ 

and for sampling with replacement

$$\binom{m}{k} \frac{W_{n-1}^k B_{n-1}^{m-k}}{T_{n-1}^m} = \binom{m}{k} \frac{W_{n-1}^k (T_{n-1} - W_{n-1})^{m-k}}{T_{n-1}^m}.$$ 

Pólya urn: \[ M = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix}, \ c \in \mathbb{N}. \]

Generalization: If we draw \( \{W^k S^{m-k}\} \) we add \( k \cdot c \) white and \( (m - k) \cdot c \) black balls, with \( c \in \mathbb{N} \). \((m + 1) \times 2\)-matrix

\[
M = \begin{pmatrix}
mc & 0 \\
(m-1)c & c \\
... & ... \\
c & (m-1)c \\
0 & mc
\end{pmatrix}
\]

Urn is balanced: \( \sigma = mc \), such that \( T_n = W_n + B_n = nmc + T_0 \).
Pólya urn: \( M = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix} \), \( c \in \mathbb{N} \).

Generalization: If we draw \( \{W^k S^{m-k}\} \) we add \( k \cdot c \) white and \( (m-k) \cdot c \) black balls, with \( c \in \mathbb{N} \). \((m+1) \times 2\)-matrix

\[
M = \begin{pmatrix} mc & 0 \\ (m-1)c & c \\ \cdots & \cdots \\ c & (m-1)c \\ 0 & mc \end{pmatrix}
\]

Urn is balanced: \( \sigma = mc \), such that \( T_n = W_n + B_n = nmc + T_0 \).
Friedman urn: $M = \begin{pmatrix} 0 & c \\ c & 0 \end{pmatrix}$, $c \in \mathbb{N}$.

Generalization: If we draw $\{W^k S^{m-k}\}$ we add $(m-k) \cdot c$ white and $k \cdot c$ black balls, with $c \in \mathbb{N}$. $(m+1) \times 2$-matrix

$$M = \begin{pmatrix} 0 & mc \\ c & (m-1)c \\ \vdots & \vdots \\ (m-1)c & c \\ mc & 0 \end{pmatrix}$$

Urn is balanced: $\sigma = mc$, such that $T_n = W_n + B_n = nmc + T_0$. 
Friedman urn: \[ M = \begin{pmatrix} 0 & c \\ c & 0 \end{pmatrix}, \ c \in \mathbb{N}. \]

Generalization: If we draw \( \{W^k S^{m-k}\} \) we add \( (m-k) \cdot c \) white and \( k \cdot c \) black balls, with \( c \in \mathbb{N} \). (\( m + 1 \) \( \times \) 2-matrix

\[
M = \begin{pmatrix}
0 & mc \\
c & (m-1)c \\
\vdots & \vdots \\
(m-1)c & c \\
mc & 0
\end{pmatrix}
\]

Urn is balanced: \( \sigma = mc \), such that \( T_n = W_n + B_n = nmc + T_0 \).
Urn models with multiple drawings: linear affine models.

Given

\[ M = \begin{pmatrix} a_0 & b_0 \\ a_1 & b_1 \\ \vdots & \vdots \\ a_{m-1} & b_{m-1} \\ a_m & b_m \end{pmatrix} \]

one cannot easily obtain the expected value.

**Difficulty:**

\[ \mathbb{E}(W_n) = \sum_{k=0}^{m} f_{n,k} \mathbb{E}(W_{n-1}^k), \]

for certain sequences \( f_{n,k} \) depending only on \( n \) and \( m \) (and \( k \)). When is it possible to **calculate** the expected value of \( W_n \) (in a **simple** way)?
Urns models with multiple drawings: linear affine models.

Given

\[
M = \begin{pmatrix}
    a_0 & b_0 \\
    a_1 & b_1 \\
    \vdots & \vdots \\
    a_{m-1} & b_{m-1} \\
    a_m & b_m
\end{pmatrix}
\]

one cannot easily obtain the expected value.

Difficulty:

\[
\mathbb{E}(W_n) = \sum_{k=0}^{m} f_{n,k} \mathbb{E}(W_{n-1}^k),
\]

for certain sequences \( f_{n,k} \) depending only on \( n \) and \( m \) (and \( k \)).

When is it possible to calculate the expected value of \( W_n \) (in a simple way)?
Urn models with multiple drawings: linear affine models.

Given

\[
M = \begin{pmatrix}
    a_0 & b_0 \\
    a_1 & b_1 \\
    \vdots & \vdots \\
    a_{m-1} & b_{m-1} \\
    a_m & b_m \\
\end{pmatrix}
\]

one cannot easily obtain the expected value.

**Difficulty:**

\[
\mathbb{E}(W_n) = \sum_{k=0}^{m} f_{n,k} \mathbb{E}(W_{n-1}^k),
\]

for certain sequences \( f_{n,k} \) depending only on \( n \) and \( m \) (and \( k \)).

When is it possible to **calculate** the expected value of \( W_n \) (in a **simple** way)?
Proposition (Mahmoud and K.)

Given matrix $M$, the numbers $a_{m-1}$, $a_m$, and the balance factor $\sigma = a_k + b_k \geq 0$. $W_n$ satisfies a linear affine relation of the form

$$
E[W_n|\mathcal{F}_{n-1}] = \alpha_n W_{n-1} + \beta_n, \quad n \geq 1,
$$

if and only if, for $0 \leq k \leq m$, the numbers $a_k$ satisfy the condition

$$
a_k = (m-k)a_{m-1} - (m-k-1)a_m.
$$

(equivalently, $a_k = a_0 + hk$, with $h$ an integer guaranteeing tenability.) Then

$$
\alpha_n = \frac{T_{n-1} + m(a_{m-1} - a_m)}{T_{n-1}}, \quad \text{and} \quad \beta_n = a_m, \quad n \geq 1.
$$

Remark

For $m = 1$ we reobtain the ordinary balanced urns.
Proof (sketch). For sampling with replacement:

\[ E[W_n|\mathcal{F}_{n-1}] = W_{n-1} + \sum_{k=0}^{m} a_{m-k} E[I_n(W^k B^{m-k}) = 1|\mathcal{F}_{n-1}] \]

\[ = W_{n-1} + \sum_{k=0}^{m} a_{m-k} \binom{m}{k} \frac{W^k_{n-1} B^{m-k}_{n-1}}{T^m_{n-1}} \]

\[ = W_{n-1} + \sum_{k=0}^{m} (k(a_{m-1} - a_m) + a_m) \binom{m}{k} \frac{W^k_{n-1} B^{m-k}_{n-1}}{T^m_{n-1}} \]

\[ = W_{n-1} + \frac{1}{T^m_{n-1}} \left( m(a_{m-1} - a_m)W_{n-1}(W_{n-1} + B_{n-1})^{m-1} + a_m(W_{n-1} \right) \]

\[ = W_{n-1} + \frac{1}{T^m_{n-1}} \left( m(a_{m-1} - a_m)W_{n-1} T_{n-1}^{m-1} + a_m T_{n-1}^m \right). \]

Taking the expectation gives the "if-part" of the stated result.
Proof (sketch). For sampling with replacement:

\[ E[W_n | \mathcal{F}_{n-1}] = W_{n-1} + \sum_{k=0}^{m} a_{m-k} E[I_n(W^k B^{m-k}) = 1 | \mathcal{F}_{n-1}] \]

\[ = W_{n-1} + \sum_{k=0}^{m} a_{m-k} \binom{m}{k} \frac{W_{n-1}^k B_{n-1}^{m-k}}{T_{n-1}^m} \]

\[ = W_{n-1} + \sum_{k=0}^{m} (k(a_{m-1} - a_m) + a_m) \binom{m}{k} \frac{W_{n-1}^k B_{n-1}^{m-k}}{T_{n-1}^m} \]

\[ = W_{n-1} + \frac{1}{T_{n-1}^m} \left( m(a_{m-1} - a_m)W_{n-1}(W_{n-1} + B_{n-1})^{m-1} + a_m(W_{n-1} \right. \]

\[ = W_{n-1} + \frac{1}{T_{n-1}^m} \left( m(a_{m-1} - a_m)W_{n-1}T_{n-1}^{m-1} + a_mT_{n-1}^m \right) \]

Taking the expectation gives the "if-part" of the stated result.
Proof (sketch). For sampling with replacement:

\[
\mathbb{E} [W_n | \mathcal{F}_{n-1}] = W_{n-1} + \sum_{k=0}^{m} a_{m-k} \mathbb{E} [\mathbb{I}_n (W^k B^{m-k}) = 1 | \mathcal{F}_{n-1}]
\]

\[
= W_{n-1} + \sum_{k=0}^{m} a_{m-k} \binom{m}{k} \frac{W_{n-1}^k B_{n-1}^{m-k}}{T_{n-1}^m}
\]

\[
= W_{n-1} + \sum_{k=0}^{m} (k(a_{m-1} - a_m) + a_m) \binom{m}{k} \frac{W_{n-1}^k B_{n-1}^{m-k}}{T_{n-1}^m}
\]

\[
= W_{n-1} + \frac{1}{T_{n-1}^m} \left( m(a_{m-1} - a_m) W_{n-1} \right) (W_{n-1} + B_{n-1})^{m-1} + a_m (W_{n-1} - 1)
\]

\[
= W_{n-1} + \frac{1}{T_{n-1}^m} \left( m(a_{m-1} - a_m) W_{n-1} T_{n-1}^{m-1} + a_m T_{n-1}^m \right).
\]

Taking the expectation gives the ”if-part” of the stated result.
Urn models with multiple drawings.

Proof (sketch). For sampling with replacement:

\[ \mathbb{E}[W_n | \mathcal{F}_{n-1}] = W_{n-1} + \sum_{k=0}^{m} a_{m-k} \mathbb{E}[\mathbb{I}_n(W^k B^{m-k}) = 1 | \mathcal{F}_{n-1}] \]

\[ = W_{n-1} + \sum_{k=0}^{m} a_{m-k} \binom{m}{k} \frac{W^k_{n-1} B^{m-k}_{n-1}}{T^m_{n-1}} \]

\[ = W_{n-1} + \sum_{k=0}^{m} (k(a_{m-1} - a_m) + a_m) \binom{m}{k} \frac{W^k_{n-1} B^{m-k}_{n-1}}{T^m_{n-1}} \]

\[ = W_{n-1} + \frac{1}{T^m_{n-1}} \left( m(a_{m-1} - a_m) W_{n-1} (W_{n-1} + B_{n-1})^{m-1} + a_m (W_{n-1} + B_{n-1})^m \right) \]

\[ = W_{n-1} + \frac{1}{T^m_{n-1}} \left( m(a_{m-1} - a_m) W_{n-1} T_{n-1}^{m-1} + a_m T_{n-1}^m \right). \]

Taking the expectation gives the "if-part" of the stated result.
Proof (sketch). For sampling with replacement:

\[
\mathbb{E}[W_n | \mathcal{F}_{n-1}] = W_{n-1} + \sum_{k=0}^{m} a_{m-k} \mathbb{E}[\mathbb{I}_n(W^k B^{m-k}) = 1 | \mathcal{F}_{n-1}]
\]

\[
= W_{n-1} + \sum_{k=0}^{m} a_{m-k} \binom{m}{k} \frac{W_{n-1}^k B_{n-1}^{m-k}}{T_{n-1}^m}
\]

\[
= W_{n-1} + \sum_{k=0}^{m} (k(a_{m-1} - a_m) + a_m) \binom{m}{k} \frac{W_{n-1}^k B_{n-1}^{m-k}}{T_{n-1}^m}
\]

\[
= W_{n-1} + \frac{1}{T_{n-1}^m} \left( m(a_{m-1} - a_m) W_{n-1} (W_{n-1} + B_{n-1})^{m-1} + a_m (W_{n-1} \right)
\]

\[
= W_{n-1} + \frac{1}{T_{n-1}^m} \left( m(a_{m-1} - a_m) W_{n-1} T_{n-1}^{m-1} + a_m T_{n-1}^m \right).
\]

Taking the expectation gives the "if-part" of the stated result.
Urn models with multiple drawings.

Although only three parameters $a_{m-1}$, $a_m$ and $\sigma$, we **unify earlier treated** models:

**Example**

- **Case** $m = 2$: MAHMOUD's condition $a_0 = 2a_1 - a_2$
- **Case** $a_m = mc$, $a_{m-1} = (m - 1)c$ and $\sigma = mc$: generalized Friedman urn model
- **Case** $a_m = 0$, $a_{m-1} = c$ and $\sigma = mc$: generalized Pólya urn model
- **Case** $a_m = 1$ and $a_{m-1} = 0$ and $\sigma = 1$ we obtain $a_k = -(m - k) + 1$, urn model for logical circuits.
Urn models

- A few results
Urn models

- A few results
Proposition (Mahmoud and K.)

For balanced affine urn schemes the expected value is given by

$$E[W_n] = \frac{a_m(n + \frac{T_0}{\sigma})}{1 - \Lambda} + \left(W_0 - \frac{a_m T_0}{\sigma} \right) \frac{(n - 1 + \frac{T_0}{\sigma} + \Lambda)}{(n - 1 + \frac{T_0}{\sigma})},$$

as well as the asymptotic expansion

$$E[W_n] = \frac{a_m}{1 - \Lambda} n + \left(W_0 - \frac{a_m T_0}{\sigma} \right) \frac{\Gamma(\frac{T_0}{\sigma})}{\Gamma(\frac{T_0}{\sigma} + \Lambda)} n^{\Lambda} + O(1),$$

Moreover, for $\Lambda = 1$ we obtain $E[W_n] = W_0 \frac{n \sigma + T_0}{T_0}$. 
Pólya-Eggenberger urns

Theorem (Mahmoud and K.)

For balanced affine urn schemes, the variance satisfies the following expansions.

Small-index urns, the case $\Lambda < \frac{1}{2}$:

$$\mathbb{V}[W_n] = \frac{a_m b_0 \Lambda^2}{m(1 - 2\Lambda)(1 - \Lambda)^2} n + o(n).$$

Critical-index urns, the case $\Lambda = \frac{1}{2}$:

$$\mathbb{V}[W_n] = \frac{a_m b_0}{m} n \log n + \mathcal{O}(n),$$

Large-index urns, the case $\Lambda > \frac{1}{2}$:

$$\mathbb{V}[W_n] = Cn^{2\Lambda} + \mathcal{O}(n),$$

with the constant $C$ being model-dependent given by an infinite sum.
Theorem (continued)

**Constant C:**

\[
C = \frac{W_0^2}{\psi_0} + \sum_{j=1}^{\infty} \frac{\beta_j E[W_{j-1}] + a_m^2 - \psi_j 2a_m j^{-\Lambda} \left( \frac{amT_0}{1-\Lambda} \right)^j + \left( W_0 - \frac{amT_0}{1-\Lambda} \right) \frac{\Gamma(T_0/\sigma)}{\Gamma(T_0/\sigma + \Lambda)} \psi_j}{\psi_j} + 2a_m^2 \zeta(2\Lambda - 1) + 2a_m \left( W_0 - \frac{amT_0}{1-\Lambda} \right) \frac{\Gamma(T_0/\sigma)}{\Gamma(T_0/\sigma + \Lambda)} \zeta(\Lambda) - \left( W_0 - \frac{amT_0}{1-\Lambda} \right)^2 \frac{\Gamma^2(T_0/\sigma)}{\Gamma^2(T_0/\sigma + \Lambda)},
\]

with \( \zeta(z) \) denoting the Riemann zeta function and \( \beta_j, \psi_j \) certain sequences:

\[
\psi_n = \begin{cases} 
\frac{\Gamma(n + \lambda_1)\Gamma(n + \lambda_2)}{\Gamma(n + \frac{T_0}{\sigma}) \Gamma(n + \frac{T_0 - 1}{\sigma})} & \text{without replacement,} \\
\frac{\Gamma(n + \mu_1)\Gamma(n + \mu_2)}{\Gamma(n + \frac{T_0}{\sigma})^2} & \text{with replacement,}
\end{cases}
\]

and \( \lambda_{1,2} = \frac{m(a_{m-1} - a_m) + T_0 - 1/2 \pm 1/2 \sqrt{1 + 4m(a_{m-1} - a_m)(a_{m-1} - a_m + 1)}}{\sigma} \),

\[
\beta_n = (a_{m-1} - a_m)^2 \left( \frac{m}{T_{n-1}} - \frac{m^2}{T_{n-1}^2} \right) + \frac{2m am(a_{m-1} - a_m)}{T_{n-1}} + 2a_m \quad \text{for sampling without replacement,}
\]

or \( \mu_{1,2} = \frac{m(a_{m-1} - a_m) + T_0 \pm (a_{m-1} - a_m) \sqrt{m}}{\sigma} \), and

\[
\beta_n = \frac{(a_{m-1} - a_m)^2 m}{T_{n-1}} + \frac{2m am(a_{m-1} - a_m)}{T_{n-1}} + 2a_m \quad \text{for sampling with replacement;}
\]
Theorem (Trichotomy of limit laws for affine linear models - Mahmoud and K.)

For a balanced linear affine two-color urn model $M$ with sample size $m \geq 1$ let $\Lambda = \frac{\lambda_1}{\lambda_2}$ denote the ratio of the two eigenvalues of $M$.

1. for $\Lambda \leq \frac{1}{2}$: \[ \frac{W_n - \mathbb{E}(W_n)}{\sqrt{\text{Var}(W_n)}} \rightarrow \mathcal{N}(0, 1). \]

2. for $\Lambda > \frac{1}{2}$: \[ \frac{W_n - \mathbb{E}(W_n)}{n^\Lambda} \rightarrow \mathcal{L}. \]

3. for $b_0 \cdot a_m = 0$ the urn is triangular and \[ \frac{W_n}{n^\Lambda} \rightarrow \mathcal{T}. \]

There exists explicit expressions for all moments of $\mathbb{E}(L^s)$ and $\mathbb{E}(T^s)$ as nested infinite sums; almost sure convergence for large-index urns and triangular urns.
Urn models

- Multiple drawings
  Analytic combinatorics
Ur unjust models
- Multiple drawings
  Analytic combinatorics
Analytic combinatorial approach à la Flajolet:

- Symbolic description of the problem, replacement of balls modelled by differential operator
- Generating functions
- Translate description to higher order linear partial differential equation
- Use analytic machinery to analyze parameters of interest
Using the idea DUMAS, FLAJOLET AND PUYHAUBERT we study urn histories using differential operators:

\( \partial_z \): differential operator with respect to \( z \): \( \partial_z(z^n) = n \cdot z^{n-1} \)

\( \Theta_z = z \cdot \partial_z \), such that \( \Theta_z(z^n) = n \cdot z^n \)

Assume we have \( w \) white, \( b \) black balls \( \implies \) encoded by \( x^w y^b \):

\[
\frac{\binom{m}{k}}{(b+w)^m} y^{m-k} \partial_y^{m-k} x^k \partial_x^k (x^w y^b) = \frac{\binom{m}{k} w^k b^{m-k}}{(b+w)^m} x^w y^b,
\]

\[
\frac{\binom{m}{k}}{(b+w)^m} \Theta_x^k \Theta_y^{m-k} (x^w y^b) = \frac{\binom{m}{k} w^k b^{m-k}}{(b+w)^m} x^w y^b.
\]
Analysis using Analytic combinatorics

Let $H_0(x, y) = \mathbb{E}(x^{W_0} y^{B_0}) H_0$ and $H(x, y, z) = \sum_{n \geq 0} H_n(x, y) \frac{z^n}{(n!)^m}$ denote the complete history generating function.

Proposition (Morcrette and K.)

Starting with $W_0$ white and $B_0$ black balls the generating function of all urn histories $H(x, y; z)$ satisfies

$$D \ast H(x, y, z) = \frac{1}{z} \cdot \Theta_z^m \ast H(x, y, z),$$

with

$$D_M = \sum_{k=0}^{m} \binom{m}{k} x^{a_{m-k}} y^{b_{m-k}} m^{-k} \partial_x^k \partial_y^{m-k},$$

and

$$D_R = \sum_{k=0}^{m} \binom{m}{k} x^{a_{m-k}} y^{b_{m-k}} \Theta_x^k \Theta_y^{m-k}.$$
Proposition (Linear affine model - AC version)

The history generating function \( H_x(1, 1, z) \) of the expected value \( \mathbb{E}(W_n) \) satisfies an ordinary differential equation if the coefficients satisfy an affinity condition: \( a_k = a_0 + h k \), with \( h \) an integer guaranteeing tenability. Then,

\[
\mathbb{E}_x\mathbb{E}_y \partial_x \mathcal{D} = \begin{cases} 
(m(1 - h) \cdot \mathcal{V}^{m-1} + \mathcal{V}^m) \mathbb{E}_x \mathbb{E}_y \partial_x + (a_0 + hm) \mathcal{V}^m \mathbb{E}_x \mathbb{E}_y, \\
(m(1 - h) \cdot \mathcal{V}^{m-1} + \mathcal{V}^m) \mathbb{E}_x \mathbb{E}_y \partial_x + (a_0 + hm) \mathcal{V}^m \mathbb{E}_x \mathbb{E}_y,
\end{cases}
\]

with \( \mathcal{V} = \sigma \Theta_z + T_0 \). Thus,

\[
(m(1 - h) \cdot \mathcal{V}^{m-1} + \mathcal{V}^m - \frac{1}{z} \cdot \Theta_z^m) H_x(1, 1, z) = -(a_0 + hm) \mathcal{V}^m H(1, 1, z)
\]

for sampling without replacement and

\[
(m(1 - h) \cdot \mathcal{V}^{m-1} + \mathcal{V}^m - \frac{1}{z} \cdot \Theta_z^m) H_x(1, 1, z) = -(a_0 + hm) \mathcal{V}^m H(1, 1, z)
\]

for sampling with replacement (with \( H(1, 1, z) \) explicitly given and \( \mathbb{E}_x f(x) = f(1) \)).
Urn models

- Outlook
Urn models

- Outlook
Outlook and conclusion

Conclusion

- Characterization of linear affine urns
- Extension of the trichotomy - small, large and triangular urns - from $m = 1$ to $m \geq 1$ for linear affine urns.
- Analytic combinatorics - PDEs and an alternative derivation of the linear affine urns.

Outlook

We obtain higher order PDEs using Morcrette’s approach also for
- $r$-colors
- unbalanced models
- ordered samples

Linear affine models: We can characterize and analyze $r$-colors models (joint work with Hosam Mahmoud).
Thanks for your attention!

Danke für Ihre Aufmerksamkeit!

Merci de votre attention!