ON THE AREA UNDER LATTICE PATHS ASSOCIATED WITH TRIANGULAR DIMINISHING URN MODELS.

MARKUS KUBA AND ALOIS PANHOLZER

ABSTRACT. This work is devoted to the analysis of the area under certain lattice paths. The lattice paths of interest are associated to a class of 2×2 triangular Pólya-Eggenberger urn models with ball replacement matrix $M = \left(\begin{smallmatrix} -a & 0 \\ c & -d \end{smallmatrix}\right)$, with $a, d \in \mathbb{N}$ and $c = p \cdot a$, $p \in \mathbb{N}_0$. We study the random variable counting the area under sample paths associated to these urn models, where we obtain a precise recursive description of its integer moments. This description allows us to derive exact formulæ for the expectation and the variance and, in principle, also for higher moments and, most nobably, it yields asymptotic expansions of all integer moments leading to a complete characterization of the limiting distributions appearing for the area under sample paths associated with these urn models. As a special instance we obtain limiting distributions for the area under sample paths of the *pills problem* urn model, originally proposed by Knuth and McCarthy, which corresponds to the special case $a = c = d = 1$. Furthermore we also obtain limiting distributions for the well known *sampling without replacement* urn, $a = d = 1$ and $c = 0$, and generalizations of it to $a, d \in \mathbb{N}$.

1. INTRODUCTION

1.1. **Pólya-Eggenberger urn models.** Pólya-Eggenberger urn models are defined as follows. We start with an urn containing n white balls and m black balls. The evolution of the urn occurs in discrete time steps. At every step a ball is drawn at random from the urn. The color of the ball is inspected and then the ball is returned to the urn. According to the observed color of the ball there are added/removed balls due to the following rules. If a white ball has been drawn, we put into the urn α white balls and β black balls, but if a black ball has been drawn, we put into the urn γ white balls and δ black balls. The values $\alpha, \beta, \gamma, \delta \in \mathbb{Z}$ are fixed integer values and the urn model is specified by the 2×2 ball replacement matrix $M = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$. This definition extends naturally also to higher dimensions.

Urn models are simple, useful mathematical tools for describing many evolutionary processes in diverse fields of application such as in the analysis of algorithms and data structures, in statistics and in genetics. Due to their importance in applications, there is a huge literature on the stochastic behavior of urn models; see for example Johnson and Kotz [12], Kotz and Balakrishnan [14] or Mahmoud [18]. Recently, a few different approaches have been proposed, which yield deep and far-reaching results for very general urn models; see the recent works of Flajolet et al. [4, 5], Janson [10, 11], or Pouyanne [19, 20] and the references therein.

Most papers in the literature impose the so-called *tenability condition* on the ball replacement matrix, so that the process can be continued ad infinitum, or no balls of a given color being completely removed. However, in some applications, there are urn models with a very different nature, which we will refer to as *diminishing urn models*. Examples of such urn models are, e.g., the pills problem urn and the sampling without replacement urn, which are considered here, but also the OK Corral urn and the Cannibal urn; see, e.g., [9] for a detailed description. Such kind of urn models can be described as follows. We consider Pólya-Eggenberger urn models specified by a transition matrix $M = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$, but in addition there is a set of absorbing states $\mathcal{A} \subseteq \mathcal{S}$ contained in a state space $\mathcal{S} \subseteq \mathbb{N} \times \mathbb{N}$. The urn contains m black balls and n white balls at the beginning, with $(m, n) \in \mathcal{S}$. Then the urn evolves by successive draws at discrete instances according to the transition matrix until an absorbing state $s = (j, k) \in \mathcal{A}$ is reached, namely, until the urn contains exactly j black balls and k white balls. Then the urn process stops. We only call an urn model "diminishing urn model" if it is guaranteed that from any initial state $(m, n) \in \mathcal{S}$ (starting with m black balls and n white balls) we will reach an absorbing state $s \in \mathcal{A}$ after a finite number of draws. Furthermore we always assume that the state space $\mathcal{S}$ is chosen in a

2000 *Mathematics Subject Classification.* 60C05, 60F05.
Key words and phrases. Pólya-Eggenberger urns models, area distribution, diminishing urns, pills problem, generalized sampling urns.
This work was supported by the Austrian Science Foundation FWF, grant S9608-N23.

suitable way, i.e., it must be guaranteed that, when starting at $(m, n) \in \mathcal{S} \setminus \mathcal{A}$ and we make a draw leading to m_1 black and n_1 white balls, we stay in our state space: $(m_1, n_1) \in \mathcal{S}$.

1.2. **Urn models and lattice paths.** A (unweighted) lattice path is the drawing of a sum of vectors from $\mathbb{Z} \times \mathbb{Z}$ in $\mathbb{Z} \times \mathbb{Z}$, where the vectors belong to a finite fixed set V, and where a certain point $s \in \mathbb{Z} \times \mathbb{Z}$ is chosen as the origin of the path (often the point $(0, 0) \in \mathbb{Z} \times \mathbb{Z}$ is chosen). If either all vectors are in $\mathbb{N} \times \mathbb{Z}$ or all vectors are in $(-\mathbb{N}) \times \mathbb{Z}$, the path is called directed (the path is going "to the right" or "to the left"). A study of lattice paths by applying methods from analytic combinatorics has been carried our recently by Banderier and Flajolet [1].

It is well known that the urn histories of Pólya-Eggenberger urn models with ball replacement matrix $M = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$ can be interpreted as weighted lattice paths in $\mathbb{Z} \times \mathbb{Z}$, see, e.g., [4]. This description of the evolution of an urn is given as follows. If the urn contains m black balls and n white balls and we select a white ball (which happens with probability $\frac{n}{m+n}$), then this corresponds to a step from (m, n) to $(m + \alpha, n + \beta)$, to which the weight $\frac{n}{m+n}$ is associated; and if we select a black ball (which happens with probability $\frac{m}{m+n}$), this corresponds to a step from (m, n) to $(m + \gamma, n + \delta)$ (with weight $\frac{m}{m+n}$). The weight of a path after t successive draws consists of the product of the weights of every step. Of course, the weight of a path $P = \big((m_0, n_0), (m_1, n_1), \dots, (m_\ell, n_\ell)\big)$ corresponds to the probability that the urn starts with m_0 black and n_0 white balls and contains after the t-th draw exactly m_t black and n_t white balls.

1.3. **Area under lattice paths associated with urn models.** The study of the area under lattice paths, measured either in a continuous way or in a discrete way as the number of lattice points below the path, has a long history; we refer to the work of Banderier and Gittenberger [2] and the references therein. Further we want to point to connections between the area under lattice paths and the area under a Browian excursion, see Louchard [17]. Considering the natural description of the evolution of diminishing urns as weighted lattice paths this naturally leads to the following question: For a given diminishing urn model with replacement matrix M, state space S and absorbing states $\mathcal{A}$, what can be said about the (discrete) area[1] below the sample paths associated with the diminishing urn? Such questions relate two widely studied topics in combinatorics and probability theory, namely lattice path enumeration and Pólya-Eggenberger urn models.

The aim of this paper is to study the distributional behaviour of the area under lattice paths associated with a whole class of triangular diminishing Pólya-Eggenberger urn models. We consider diminishing urn models with ball replacement matrix M given by

$$M = \begin{pmatrix} -a & 0 \\ c & -d \end{pmatrix}, \quad \text{with} \;\; a, d \in \mathbb{N} \;\; and \;\; c = p \cdot a, \quad p \in \mathbb{N}_0, ^2 \tag{1}$$

state space $\mathcal{S} := \{(d \cdot m, a \cdot n) | \, m, n \in \mathbb{N}_0\}$, and the set of absorbing states $\mathcal{A} := \{(0, a \cdot n) | \, n \in \mathbb{N}_0\}$. The steps of weighted lattice paths associated with this diminishing urn model are illustrated in Figure 1.

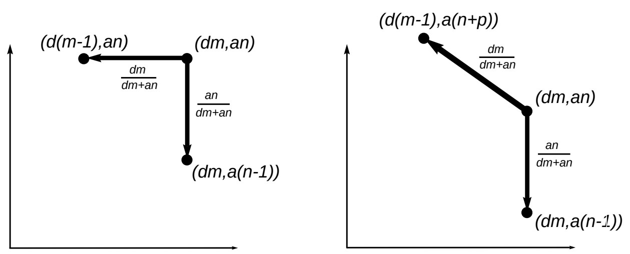

FIGURE 1. The steps associated with $M = \left(\begin{smallmatrix} -a & 0 \\ c & -d \end{smallmatrix}\right)$ for $c = 0$ and $c = p \cdot a > 0$, respectively.

Let us consider now a weighted lattice path $P = \big((m_0, n_0), (m_1, n_1), \dots, (m_\ell, n_\ell)\big)$, with $(m_0, n_0) = (dm, an)$ and $(m_\ell, n_\ell) = (0, ak)$, associated with the urn model with ball replacement matrix M given by (1), which starts at $(dm, an) \in \mathcal{S}$ and ends at the absorbing state $(0, ak) \in \mathcal{A}$. We define then the

[1] It is natural to measure the discrete area as the number of points of the state space $\mathcal{S} \subseteq \mathbb{N} \times \mathbb{N}$, which are below a certain sample path.

[2] Throughout this work we use the notation $\mathbb{N} := \{1, 2, 3, \dots\}$ and $\mathbb{N}_0 := \{0, 1, 2, \dots\}$.

discrete area under such a weighted lattice path P as the number points of the state space $\mathcal{S}$ (with a positive x-coordinate), which are below the path P. To be more precise the discrete area under the lattice path P will be defined as follows:

$$|\{(d\tilde{m}, a\tilde{n}) \in \mathcal{S} \mid 0 < \tilde{m} \leq m \text{ and } a\tilde{n} < m_t, \text{ for all } (m_t, n_t) \in P \text{ with } m_t = d\tilde{m}\}|.$$

Thus the discrete area can be interpreted as the number of rectangles with side lengths d and a, which fit below the lattice path P. The continuous area under the lattice path P will be defined naturally as the area of the subset of $\mathbb{R}^2$ enclosed by the lines $x = 0$, $y = 0$, $x = dm$, and the curve obtained by connecting consecutive points of P by straight lines. These notions of area are illustrated in Figure 2.

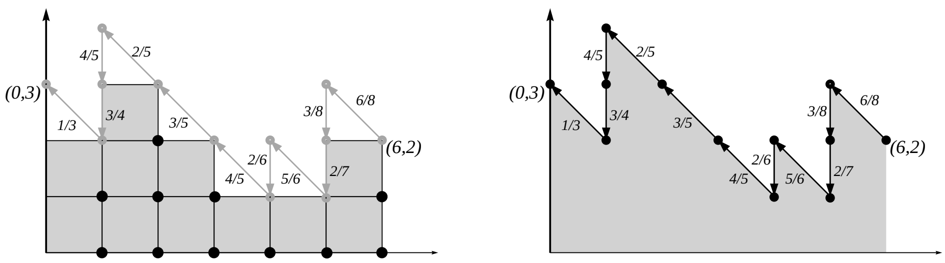

FIGURE 2. An example of a weighted path from $(6, 2)$ to the absorbing state $(0, 3)$ for the pills problem $M = \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix}$ and the vertical absorbing axis $\mathcal{A} = \{(0, n) : n \geq 0\}$. The illustrated path has weight $\frac{6}{8}\frac{3}{8}\frac{2}{7}\frac{5}{6}\frac{2}{6}\frac{4}{5}\frac{3}{5}\frac{2}{5}\frac{4}{5}\frac{3}{4}\frac{1}{3} = \frac{3}{3500}$, discrete area 11, and continuous area 14.

We introduce now the random variable $A_{an,dm}$, which counts the discrete area under weighted lattice paths associated with the urn model $M = \begin{pmatrix} -a & 0 \\ c & -d \end{pmatrix}$, with $a, d \in \mathbb{N}$ and $c = p \cdot a$, $p \in \mathbb{N}_0$, starting with $a \cdot n$ white and $d \cdot m$ black balls. Moreover, we denote with $C_{an,dm}$ the corresponding random variable counting the continuous area under weighted lattice paths.

This paper is now devoted to a study of the limiting distribution behaviour of the random variables $A_{an,dm}$ and $C_{an,dm}$ for diminishing urn models, where the urns evolve according to a ball replacement matrix M given by (1). We will see that different limiting distributions arise according to the growth of m and n, and we are able to give a full characterization of the limiting behaviour of the random variables considered, for $\max(m, n) \to \infty$. It is evident that the distributions of the discrete area $A_{an,dm}$ and the continuous area $C_{an,dm}$ are related by the linear equation

$$A_{an,dm} = \frac{C_{an,dm} - \frac{mcd}{2}}{ad}.$$

Hence, it is justified that we restrict our analysis solely to the discrete area, since all the obtained results can be transferred easily to corresponding results for the continuous area.

Note that when starting with $a \cdot n + q$ white balls, where $1 \leq q < a$, it turn out that the urn model is no longer well defined. It may happen that at some stage only q white balls are left, but when choosing a white ball we are forced to remove a white balls. The same problem occurs when the parameter c is not a multiple of the parameter a in the definition of the ball replacement matrix. Thus it is natural to restrict our considerations to the instances given in the definition of (1).

1.4. **Motivation.** Besides our theoretical interest in combining studies concerning the area under lattice paths and studies concerning diminishing urn models, our analysis of the class of diminishing urns with a ball replacement matrix given by (1) is motivated by two particular urn models contained in this class, namely the pills problem urn model and the sampling without replacement urn. Our studies concerning the area of lattice paths associated to these urn models will give more insight into the behaviour of these fundamental urns.

The pills problem urn is given by (1) for the particular choice $a = d = c = 1$, i.e., by the ball replacement matrix $M = \begin{pmatrix} -1 & 0 \\ 1 & -1 \end{pmatrix}$. The so-called "pills problem" was originally proposed by Knuth and McCarthy in

[13], p. 264, and later solved by Hesterberg [8]. It can be formulated as computing the expected number of white balls remaining in the urn when all black balls are removed starting with m black and n white balls. The problem has been revisited by Brennan and Prodinger in [3]. Recently, a distributional analysis of the parameter "number of white balls remaining in the urn when all black balls are removed starting with m black and n white balls" for the pills problem urn model has been given by Hwang et al [9] using a generating functions approach, leading to a full characterization of the limiting distributions appearing (depending on the growth of m and n). Furthermore, a general study of the behaviour of this parameter for diminishing urn models with a ball replacement matrix given by (1) was conducted in [15].

The sampling without replacement urn is a simple fundamental urn model corresponding to the ball replacement matrix $M = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$.

1.5. **Notation and plan of the paper.** We denote by $X_n \xrightarrow{\mathcal{L}} X$ the weak convergence, i.e., the convergence in distribution, of the sequence of random variables X_n to a random variable X. We use the notation $X \stackrel{\mathcal{L}}{=} Y$ for the equality in distribution of random variables X and Y. We denote with $\left[{n \atop k}\right]$ the signless Stirling numbers of the first kind and with $\left\{{n \atop k}\right\}$ the Stirling numbers of the second kind. Furthermore, we use the abbreviation $x^{\underline{\ell}} := x(x-1)\cdots(x-\ell+1)$ and $x^{\overline{\ell}} := x(x+1)\cdots(x+\ell-1)$ for the falling and rising factorials, respectively. The floor function $\lfloor x \rfloor$ of a real number x is defined as the largest integer q smaller or equal to x, i.e., $\lfloor x \rfloor := \max\{q \in \mathbb{Z} \mid q \le x\}$. Moreover, we frequently use the shorthand notation $e_{n,m}^{[s]} := \mathbb{E}(A_{an,dm}^s)$.

Let $P(x) = \sum_{i=0}^n a_n x^n$, with $a_n \neq 0$, denote a polynomial $P(x) \in \mathbb{R}[x]$. We use the notation $\deg P(x) = n$ for the degree of $P(x)$, and $\operatorname{lc} P(x) = [x^n]P(x) = a_n$ for the leading coefficient of $P(x)$.

The main result of this work is a full characterization of the limiting distribution behaviour of $A_{an,dm}$ depending on the ball replacement matrix M and the initial state (dm, an), for $\max(m,n) \to \infty$, which is given in the next section. In Section 3 we study the structure of the moments of $A_{an,dm}$ and obtain a precise recursive description of the moments.. In Sections 4-5 we provide proofs of the limiting distribution results by using the method of moments, i.e., by applying the Theorem of Fréchet and Shohat (the second central limit theorem), see, e.g., [16].

2. RESULTS

Theorem 1. *The expectation and the variance of the random variable $A_{an,dm}$, counting the discrete area under lattice paths associated with diminishing urn models with a ball replacement matrix defined in equation* (1)*, are given by the following exact formulæ.*

$$\mathbb{E}(A_{an,dm}) = \frac{nm}{1+\frac{a}{d}} + \frac{cm(m-1)}{2a(1+\frac{a}{d})},$$

$$\begin{aligned}\mathbb{V}(A_{an,dm}) = {} & \frac{a^2 d}{(a+d)^2(2a+d)} mn^2 + \left[\frac{ad^2(2a+2c+d)}{(a+d)^2(2a+d)(a+2d)} m^2 + \frac{ad(a^2+ad-2cd+d^2)}{(a+d)^2(2a+d)(a+2d)} m\right] n \\ & + \frac{cd^2(2a+2c+d)}{3(a+d)^2(2a+d)(a+2d)} m^3 + \frac{cd(a^2+ac-ad-2cd)}{2(a+d)^2(2a+d)(a+2d)} m^2 \\ & - \frac{cd(3a^2+3ac+ad-2cd+2d^2)}{6(a+d)^2(2a+d)(a+2d)} m.\end{aligned}$$

Theorem 2. *Depending on the growth of m and n, we obtain for arbitrary sequences (m,n), with $\max(m,n) \to \infty$, the following complete characterization of the limiting distributions appearing for the discrete area $A_{an,dm}$ under lattice paths associated with diminishing urn models with a ball replacement matrix defined in equation* (1)*.*

$I.$: *Instance $c = 0$: we have do distinguish between the following three cases.*

(a) *For sequences (m,n), such that $m \in \mathbb{N}$ is fixed and $n \to \infty$, the normalized random variable $A_{an,dm}/n$ converges in distribution to a random variable $X_m = X_m(a,d)$, which is characterized by the (sequence of) distributional equations*

$$\frac{A_{an,dm}}{n} \xrightarrow{\mathcal{L}} X_m, \quad \text{with} \quad X_m \stackrel{\mathcal{L}}{=} Y_m \cdot (1 + X_{m-1}), \quad \text{for } m \ge 1, \quad X_0 = 0,$$

where $Y_m \stackrel{\mathcal{L}}{=} B\big(\frac{dm}{a},1\big)$, being independent of $X_0,\ X_1,\ \dots$, with $B(\alpha,\beta)$ denoting a Beta-distributed random variable with parameters α and β. Equivalently, X_m can be characterized as follows:

$$X_m \stackrel{\mathcal{L}}{=} \sum_{k=1}^{m}\prod_{\ell=0}^{k-1} Y_{m-\ell},$$

with $Y_m \stackrel{\mathcal{L}}{=} B\big(\frac{dm}{a},1\big)$, and the random variables Y_m being mutually independent.

(b) *For sequences (m,n), such that $\min(m,n)\to\infty$, the centered and normalized random variable $A^*_{an,dm}$ is asymptotically Gaussian distributed:*

$$A^*_{an,dm} := \frac{A_{an,dm}-\mathbb{E}(A_{an,dm})}{\sqrt{\mathbb{V}(A_{an,dm})}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1),$$

where $\mathcal{N}(0,1)$ denotes the standard normal distribution.

(c) *For sequences (m,n), such that n is fixed and $m\to\infty$, the normalized random variable $A_{an,dm}/m$ converges in distribution to a random variable $W_n = W_n(a,d)$, which is characterized by the (sequence of) distributional equations*

$$\frac{A_{an,dm}}{m} \xrightarrow{\mathcal{L}} W_n, \quad \text{with} \quad W_n \stackrel{\mathcal{L}}{=} W_{n-1}Z_n + n\big(1-Z_n\big), \quad \text{for } n\ge 1, \quad W_0=0,$$

where $Z_n \stackrel{\mathcal{L}}{=} B\big(\frac{an}{d},1\big)$, being independent of $W_0,\ W_1,\ \dots$, with $B(\alpha,\beta)$ denoting a Beta-distributed random variable with parameters α and β. Equivalently, the random variable $W_n = W_n(a,d)$ is related to the random variable $X_m = X_m(a,d)$ defined above due to the equation $W_n(a,d) \stackrel{\mathcal{L}}{=} n - X_n(d,a)$, and consequently characterized via

$$W_n(a,d) \stackrel{\mathcal{L}}{=} n - \sum_{k=1}^{n}\prod_{\ell=0}^{k-1} Z_{n-\ell},$$

where the random variables $Z_n \stackrel{\mathcal{L}}{=} B\big(\frac{an}{d},1\big)$ are mutually independent.

$II.$: *Instance $c\neq 0$: we have do distinguish between the following two cases.*

(a) *For sequences (m,n), such that $m\in\mathbb{N}$ is fixed and $n\to\infty$, we have the same limiting behaviour of $A_{an,dm}$ as for the case $I.(a)$.*

(b) *For sequences (m,n), such that $m\to\infty$ and arbitrary n (in particular, n might be fixed or $n=n(m)$ can arbitrarily grow with m), the centered and normalized random variable $A^*_{an,dm}$ is asymptotically Gaussian distributed:*

$$A^*_{an,dm} := \frac{A_{an,dm}-\mathbb{E}(A_{an,dm})}{\sqrt{\mathbb{V}(A_{an,dm})}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1).$$

Remark 1. In all three cases we have convergence of all integer moments. In the case where m is fixed and n is tending to infinity, and for the instance $c=0$ also the case where n is fixed and m is tending to infinity, we have alternative characterizations of the limiting distributions appearing in terms of recurrence relations for the moments of $X_m(a,d)$ and $W_n(a,d)$, which are equivalent to the stated distributional equations, see Section 4.

3. A RECURSIVE APPROACH FOR THE MOMENT STRUCTURE

3.1. **A recurrence for the discrete area.** The starting point of our analysis is the following recursive distributional equation for the random variable $A_{an,dm}$, which counts the discrete area under weighted lattice paths associated to diminishing urn models with a ball replacement matrix defined in (1). This equation is obtained immediately by considering the urn after the first draw when starting at state (dm,an). We obtain then that the random variable $A_{an,dm}$ satisfies the distributional equation

$$A_{an,dm} \stackrel{\mathcal{L}}{=} \mathbb{I}_{n,m}A_{a(n-1),dm} + (1-\mathbb{I}_{n,m})(A'_{a(n+p),d(m-1)}+n), \quad \text{with } A_{an,0}=0, \tag{2}$$

where $\mathbb{I}_{n,m}$ denotes the indicator variable of the event “choosing a white ball”,

$$\mathbb{P}\{\mathbb{I}_{n,m}=1\} = \frac{an}{an+dm}, \quad \mathbb{P}\{\mathbb{I}_{n,m}=0\} = \frac{dm}{an+dm},$$

with $\mathbb{I}_{n,m}$ being independent of $(A_{an,dm})_{n,m\ge 0}$ and $(A'_{an,dm})_{n,m\ge 0}$, which are independent copies of each other.

3.2. **Recurrence relations for the moments.** In the next step we consider the s-th moments of $A_{an,dm}$, which will be denoted by $e_{n,m}^{[s]} := \mathbb{E}(A_{an,dm}^s)$. Using the distributional recurrence (2) we obtain immediately the following recurrence for these s-th moments, $s \ge 1$:

$$e_{n,m}^{[s]} = \frac{an}{an+dm} e_{n-1,m}^{[s]} + \frac{dm}{an+dm} \sum_{\ell=0}^{s} \binom{s}{\ell} n^\ell e_{n+p,m-1}^{[s-\ell]}, \quad \text{for } n \ge 0,\ m \ge 1, \tag{3}$$

with initial values $e_{n,m}^{[0]} = 1$, for $n, m \ge 0$, and $e_{n,0}^{[s]} = 0$, for $n \ge 0$ and $s \ge 1$.

Our aim is an asymptotic study of the moments $e_{n,m}^{[s]}$, which are given as solutions of recurrence (3). As preliminary steps we will prove results concerning the exact solutions for $e_{n,m}^{[s]}$, considered as functions depending on n, m and s. The following proposition states that $e_{n,m}^{[s]}$ has a rather simple dependence on n, namely $e_{n,m}^{[s]}$ can be considered as a polynomial in n of degree at most s, with certain recursively computable coefficients.

Proposition 1. *The moments $e_{n,m}^{[s]} = \mathbb{E}(A_{an,dm}^s)$ of the random variable $A_{an,dm}$ satisfy the expansion*

$$e_{n,m}^{[s]} = \sum_{\ell=0}^{s} \varphi_{s,\ell,m} n^\ell,$$

with coefficients $\varphi_{s,\ell,m}$ that can be computed recursively as follows, where the initial values are given by $\varphi_{s,\ell,0} = 0$, for $0 \le \ell \le s$, $s \ge 1$, and $\varphi_{0,0,m} = 1$, for $m \ge 0$.

- *For $1 \le \ell \le s$ and $m \ge 1$ the values $\varphi_{s,\ell,m}$ satisfy the following recurrence:*

$$\varphi_{s,\ell,m} = \frac{1}{a\ell \binom{m+\frac{a\ell}{d}}{m}} \sum_{k=1}^{m} \binom{k+\frac{a\ell}{d}-1}{k} \psi_{s,\ell,k}, \tag{4a}$$

with

$$\psi_{s,\ell,m} := a \sum_{k=\ell+1}^{s} \binom{k}{\ell-1} (-1)^{k-\ell-1} \varphi_{s,k,m} + dm \sum_{k=\ell+1}^{s} \binom{k}{\ell} p^{k-\ell} \varphi_{s,k,m-1}$$
$$+ dm \sum_{i=1}^{\ell} \binom{s}{i} \sum_{k=\ell-i}^{s-i} \binom{k}{\ell-i} \varphi_{s-i,k,m-1} p^{k-\ell+i}.$$

- *For $\ell = 0$ and $m \ge 1$ the values $\varphi_{s,\ell,m}$ satisfy the following recurrence:*

$$\varphi_{s,0,m} = \sum_{k=0}^{m-1} \psi_{s,0,k}, \quad \textit{with} \quad \psi_{s,0,m} := \sum_{i=1}^{s} \varphi_{s,i,m} p^i. \tag{4b}$$

Remark 2. If $c = p \cdot a = 0$ it trivially holds $e_{0,m}^{[s]} = 0$, for $s \ge 1$, which implies that $\varphi_{s,0,m} = 0$, for $c = 0$ and $s \ge 1$. Of course, one also obtains this result by considering the recurrences (4) for the instance $c = 0$.

Proof. First we remark that due to $e_{n,m}^{[0]} = 1$, for $m, n \ge 0$, Proposition 1 holds for $s = 0$ leading to values $\varphi_{0,0,m} = 1$, for $m \ge 0$. Furthermore, due to $e_{n,0}^{[s]} = 0$, for all $s \ge 1$ and $n \ge 0$, we obtain that Proposition 1 also holds for $e_{n,0}^{[s]}$, with $s \ge 1$, leading to values $\varphi_{s,\ell,0} = 0$, for $0 \le \ell \le s$ and $s \ge 1$.

In order to prove the stated expansion of $e_{n,m}^{[s]}$ for $s \ge 1$, $m \ge 1$ and $n \ge 0$ we start with the *Ansatz*: $e_{n,m}^{[s]} = \sum_{k=0}^{s} \varphi_{s,k,m} n^k$ and plug it into recurrence (3). This leads to the following equation:

$$(an+dm) \sum_{k=0}^{s} \varphi_{s,k,m} n^k = an \sum_{k=0}^{s} \varphi_{s,k,m} (n-1)^k + dm \sum_{\ell=0}^{s} \binom{s}{\ell} n^\ell \sum_{k=0}^{s-\ell} \varphi_{s-\ell,k,m-1} (n+p)^k. \tag{5}$$

By comparing the coefficients of n^ℓ, for $0 \le \ell \le s+1$, in equation (5) we obtain the following system of $s+2$ equations:

$$\varphi_{s,s,m} = \varphi_{s,s,m}, \quad \text{for } \ell = s+1, \tag{6a}$$

$$a\varphi_{s,\ell-1,m} + dm\varphi_{s,\ell,m} = a\sum_{k=\ell-1}^{s}(-1)^{k-\ell+1}\varphi_{s,k,m}\binom{k}{\ell-1} + dm\sum_{k=\ell}^{s}\binom{k}{\ell}p^{k-\ell}\varphi_{s,k,m-1} + dm\sum_{i=1}^{\ell}\binom{s}{i}\sum_{k=\ell-i}^{s-i}\binom{k}{\ell-i}p^{k-\ell+1}\varphi_{s-i,k,m-1}, \quad \text{for } 1 \le \ell \le s, \tag{6b}$$

$$\varphi_{s,0,m} = \sum_{k=0}^{s} p^k \varphi_{s,k,m-1}, \quad \text{for } \ell = 0, \tag{6c}$$

where we already determined the initial values $\varphi_{s,i,0} = 0$, for $0 \le i \le s$ and $s \ge 1$.

Considering (6b) it turns out that the term $\varphi_{s,\ell-1,m}$ on the left hand side of (6b) cancels with the first summand of $\sum_{k=\ell-1}^{s}(-1)^{k-\ell+1}\varphi_{s,k,m}\binom{k}{\ell-1}$ on the right hand side of (6b). This allows to rewrite equations (6b) and (6c) as follows:

$$\varphi_{s,\ell,m} = \frac{dm}{dm+a\ell}\varphi_{s,\ell,m-1} + \frac{1}{dm+a\ell}\psi_{s,\ell,m}, \quad \text{for } 1 \le \ell \le s, \tag{7a}$$

$$\text{with} \quad \psi_{s,\ell,m} := a\sum_{k=\ell+1}^{s}\binom{k}{\ell-1}(-1)^{k-\ell-1}\varphi_{s,k,m} + dm\sum_{k=\ell+1}^{s}\binom{k}{\ell}p^{k-\ell}\varphi_{s,k,m-1} + dm\sum_{i=1}^{\ell}\binom{s}{i}\sum_{k=\ell-i}^{s-i}\binom{k}{\ell-i}\varphi_{s-i,k,m-1}p^{k-\ell+i},$$

and

$$\varphi_{s,0,m} = \varphi_{s,0,m-1} + \psi_{s,0,m-1}, \quad \text{with} \quad \psi_{s,0,m} := \sum_{i=1}^{s}\varphi_{s,i,m}p^i. \tag{7b}$$

Equations (7a) and (7b) can be considered as linear first order recurrences with respect to the parameter m, whose solutions are given by the following explicit expressions:

$$\varphi_{s,\ell,m} = \sum_{k=0}^{m-1}\frac{\binom{m}{k}\psi_{s,\ell,m-k}}{\binom{m+\frac{a\ell}{d}}{k}(d(m-k)+a\ell)} = \sum_{k=1}^{m}\frac{m!\psi_{s,\ell,k}}{k!d(m+\frac{a\ell}{d})^{\underline{m-k+1}}} = \frac{1}{a\ell\binom{m+\frac{a\ell}{d}}{m}}\sum_{k=1}^{m}\binom{k+\frac{a\ell}{d}-1}{k}\psi_{s,\ell,k}, \quad \text{for } 1 \le \ell \le s \text{ and } m \ge 1, \tag{8a}$$

$$\varphi_{s,0,m} = \sum_{k=0}^{m-1}\psi_{s,0,k}, \quad \text{for } s \ge 1 \text{ and } m \ge 1. \tag{8b}$$

When considering $\psi_{s,\ell,k}$ as defined in equations (7a) and (7b) it is apparent that in order to compute the values $\varphi_{s,\ell,m}$ as given in equations (8a) and (8b) one only requires values $\varphi_{r,i,k}$, with $0 \le k \le m$ and either $0 \le r < s$ or ($r = s$ and $\ell + 1 \le i \le s$). Hence we obtain by a simple induction argument that the system of recurrences (8), together with the initial values $\varphi_{0,0,m} = 1$, for $m \ge 0$ and $\varphi_{s,\ell,0} = 0$, for $0 \le \ell \le s$ and $s \ge 1$, leads to a unique solution for all values $\varphi_{s,\ell,m}$. This shows that the Ansatz $e_{n,m}^{[s]} = \sum_{\ell=0}^{s}\varphi_{s,\ell,m}n^\ell$ is indeed justified, and that the values $\varphi_{s,\ell,m}$, are determined recursively by (8). These recurrences are given in Proposition 1.

□

3.3. **The structure of the values $\varphi_{s,\ell,m}$.** We continue now our studies on the structure of the moments $e_{n,m}^{[s]}$, which are given as solutions of recurrence (3) and consider the values $\varphi_{s,\ell,m}$ appearing in Proposition 1.

To do this we require the following two lemmas concerning the sums appearing in (4).

Lemma 1. *For $r \ge 1$ and $\ell \ge 1$ the sum*

$$S(m) := \frac{1}{a\ell\binom{m+\frac{a\ell}{d}}{m}}\sum_{k=1}^{m}\binom{k+\frac{a\ell}{d}-1}{k}k^r$$

can be written as a polynomial in m *of degree* r*,* $\deg S(m) = r$*, whose constant term vanishes:*

$$S(m) = \sum_{i=1}^{r} \alpha_i m^i, \quad \text{with} \quad \alpha_i = \sum_{j=i}^{r} \frac{\left\{ {r \atop j} \right\} \left[{j \atop i} \right] (-1)^{j-i}}{dj + a\ell}.$$

The leading coefficient $\alpha_r = \operatorname{lc} S(m)$ *is given by* $\alpha_r = \frac{1}{dr + a\ell}$.

Proof. Since $r \geq 1$ we can write

$$k^r = \sum_{j=1}^{r} \left\{ {r \atop j} \right\} k^{\underline{j}},$$

and consider the sums

$$S_j(m) := \frac{1}{a\ell \binom{m + \frac{a\ell}{d}}{m}} \sum_{k=1}^{m} \binom{k + \frac{a\ell}{d} - 1}{k} k^{\underline{j}}, \quad \text{for } j \geq 1.$$

We obtain

$$\begin{aligned}
\sum_{k=1}^{m} \binom{k + \frac{a\ell}{d} - 1}{k} k^{\underline{j}} &= \sum_{k=j}^{m} \frac{\left(k + \frac{a\ell}{d} - 1\right)^{\underline{k}}}{(k-j)!} = \sum_{k=0}^{m-j} \frac{\left(k + j + \frac{a\ell}{d} - 1\right)^{\underline{k+j}}}{k!} \\
&= \sum_{k=0}^{m-j} \frac{\left(k + j + \frac{a\ell}{d} - 1\right)^{\underline{k}} \left(j + \frac{a\ell}{d} - 1\right)^{\underline{j}}}{k!} \\
&= j! \binom{j + \frac{a\ell}{d} - 1}{j} \sum_{k=0}^{m-j} \binom{k + j + \frac{a\ell}{d} - 1}{k},
\end{aligned}$$

and further

$$S_j(m) = \frac{j! \binom{j + \frac{a\ell}{d} - 1}{j}}{a\ell \binom{m + \frac{a\ell}{d}}{m}} \sum_{k=0}^{m-j} \binom{k + j + \frac{a\ell}{d} - 1}{k} = \frac{j! \binom{j + \frac{a\ell}{d} - 1}{j}}{a\ell \binom{m + \frac{a\ell}{d}}{m}} \binom{m + \frac{a\ell}{d}}{j + \frac{a\ell}{d}} = \frac{m^{\underline{j}}}{dj + a\ell}.$$

Thus we get

$$S(m) = \sum_{j=1}^{r} \left\{ {r \atop j} \right\} S_j(m) = \sum_{j=1}^{r} \left\{ {r \atop j} \right\} \frac{m^{\underline{j}}}{dj + a\ell} = \sum_{i=1}^{r} m^i \sum_{j=i}^{r} \frac{\left\{ {r \atop j} \right\} \left[{j \atop i} \right] (-1)^{j-i}}{dj + a\ell},$$

which proves Lemma 1, where we used

$$m^{\underline{j}} = \sum_{i=1}^{j} \left[{j \atop i} \right] (-1)^{j-i} m^i, \quad \text{for } j \geq 1.$$

The leading coefficient of $S(m)$ is immediately obtained by evaluating the appearing expression at $i = r$. □

The next lemma is very well known, so we state it without proof.

Lemma 2. *For* $r \geq 0$ *the sum*

$$S(m) := \sum_{k=0}^{m-1} k^r$$

can be written as a polynomial in m *of degree* $r+1$*,* $\deg S(m) = r + 1$*, whose constant term vanishes:*

$$S(m) = \sum_{i=1}^{r+1} \alpha_i m^i, \quad \text{with} \quad \alpha_i = \sum_{j=i}^{r+1} \left\{ {r \atop j-1} \right\} \left[{j \atop i} \right] \frac{(-1)^{j-i}}{j}.$$

The leading coefficient $\alpha_{r+1} = \operatorname{lc} S(m)$ *is given by* $\alpha_{r+1} = \frac{1}{r+1}$.

We are going to describe now the values $\varphi_{s,\ell,m}$, when we consider them as functions depending on m. It turns out that the functions $\varphi_{s,\ell}(m) := \varphi_{s,\ell,m}$ are, for $0 \leq \ell \leq s$, polynomials in m.

Proposition 2. *The functions* $\varphi_{s,\ell}(m) := \varphi_{s,\ell,m}$ *are, for* $0 \leq \ell \leq s$*, polynomials in* m*. The degree of these polynomials depends on whether* $c = 0$ *or* $c \neq 0$*:*

- *The case $c = 0$: We have the initial values $\varphi_{0,0}(m) = 1$ and $\varphi_{s,0}(m) = 0$, for $s \ge 1$. Furthermore, for $s \ge 1$ and $1 \le \ell \le s$, the function $\varphi_{s,\ell}(m)$ is a polynomial in m of degree $\le s$ whose constant term vanishes:*

$$\varphi_{s,\ell}(m) = \sum_{j=1}^{s} \vartheta_{s,\ell,j} m^j,$$

where the coefficients $\vartheta_{s,\ell,j}$ are certain computable constants satisfying a system of recurrences. In particular, the leading coefficients $\vartheta_{s,\ell,s}$ satisfy, for $1 \le \ell \le s$, the following recurrence relation:

$$\vartheta_{s,\ell,s} = \frac{s\vartheta_{s-1,\ell-1,s-1}}{s+\frac{a\ell}{d}} + \frac{a}{d}\sum_{k=\ell+1}^{s}\binom{k}{\ell-1}(-1)^{k-\ell-1}\frac{\vartheta_{s,k,s}}{s+\frac{a\ell}{d}}, \quad \text{for } 2 \le \ell \le s-1,$$

$$\vartheta_{s,1,s} = \frac{a}{d}\sum_{k=2}^{s}(-1)^k\frac{\vartheta_{s,k,s}}{s+\frac{a}{d}}, \quad \vartheta_{s,s,s} = \frac{1}{(1+\frac{a}{d})^s}.$$

- *The case $c \ne 0$: We have the initial value $\varphi_{0,0}(m) = 1$. Furthermore, for $s \ge 1$ and $0 \le \ell \le s$, the function $\varphi_{s,\ell}(m)$ is a polynomial in m of degree $\le 2s-\ell$ whose constant term vanishes:*

$$\varphi_{s,\ell}(m) = \sum_{j=1}^{2s-\ell} \vartheta_{s,\ell,j} m^j,$$

where the coefficients $\vartheta_{s,\ell,j}$ are certain computable constants satisfying a system of recurrences. In particular, the leading coefficients $\vartheta_{s,\ell,2s-\ell}$ satisfy, for $0 \le \ell \le s$ and $s \ge 1$, the following recurrence relation:

$$\vartheta_{s,\ell,2s-\ell} = \frac{p(\ell+1)\vartheta_{s,\ell+1,2s-\ell-1} + s\vartheta_{s-1,\ell-1,2s-\ell-1}}{2s-\ell+\frac{a\ell}{d}}, \quad \text{for } 1 \le \ell \le s-1,$$

$$\vartheta_{s,0,2s} = \frac{p\vartheta_{s,1,2s-1}}{2s}, \quad \vartheta_{s,s,s} = \frac{1}{(1+\frac{a}{d})^s}.$$

Proof. We distinguish between the cases $c = 0$ and $c \ne 0$.

• The case $c = 0$. The initial values $\varphi_{0,0}(m) = \varphi_{0,0,m} = 1$ and $\varphi_{s,0}(m) = \varphi_{s,0,m} = 0$, for $s \ge 1$, have already been computed.

Next we consider the functions $\varphi_{s,s}(m) = \varphi_{s,s,m}$, for $s \ge 1$, where recurrence (4a) leads to

$$\varphi_{s,s}(m) = \frac{1}{as\binom{m+\frac{as}{d}}{m}}\sum_{k=1}^{m}\binom{k+\frac{as}{d}-1}{k}\psi_{s,s}(k), \tag{9}$$

with

$$\psi_{s,s}(m) = dm\sum_{i=1}^{s}\binom{s}{i}\varphi_{s-i,s-i}(m-1). \tag{10}$$

To show that $\varphi_{s,s}(m)$ is, for $s \ge 0$, a polynomial in m of degree $\le s$, i.e., $\deg \varphi_{s,s}(m) \le s$, one can use an easy induction argument: for $s = 0$ this trivially holds. We assume now, for $s \ge 1$, that we have shown this statement for all j, with $0 \le j < s$. Then $\varphi_{s-i,s-i}(k-1)$ is a polynomial in $k-1$, and thus also in k, of degree $\le s-i$, for $1 \le i \le s$. This implies that $\psi_{s,s}(m)$ is a polynomial in m of degree $\le s$, which has a vanishing constant term: $[m^0]\psi_{s,s}(m) = 0$, due to the factor m in (10). With Lemma 1 we obtain then that $\varphi_{s,s}(m)$ is a polynomial in m of degree $\le s$ whose constant term vanishes.

Considering (10) it is easily seen that the leading coefficient of $\psi_{s,s}(m)$ is, for $s \ge 1$, given as follows:

$$[m^s]\psi_{s,s}(m) = ds\vartheta_{s-1,s-1,s-1},$$

which leads due to Lemma 1 to the following recurrence for $[m^s]\varphi_{s,s}(m) = \vartheta_{s,s,s}$:

$$\vartheta_{s,s,s} = \frac{ds\vartheta_{s-1,s-1,s-1}}{ds+as}, \quad s \ge 1, \qquad \vartheta_{0,0,0} = 1.$$

This immediately gives

$$\vartheta_{s,s,s} = \frac{1}{\left(1+\frac{a}{d}\right)^s},$$

which is stated in Proposition (2).

To show that $\varphi_{s,\ell}(m) = \varphi_{s,\ell,m}$ is, for $1 \le \ell < s$, a polynomial in m of degree $\le s$ we again use induction; we assume now that Proposition 2 holds for all $\varphi_{t,j}(m)$, with $0 \le j \le t < s$ or $(t = s$ and $\ell < j \le s$.

Due to $c = pa = 0$ the recurrence (4a) leads to

$$\varphi_{s,\ell}(m) = \frac{1}{a\ell\binom{m+\frac{a\ell}{d}}{m}} \sum_{k=1}^{m} \binom{k+\frac{a\ell}{d}-1}{k} \psi_{s,\ell}(k), \tag{11}$$

with

$$\psi_{s,\ell}(m) = a \sum_{k=\ell+1}^{s} \binom{k}{\ell-1} (-1)^{k-\ell-1} \varphi_{s,k}(m) + dm \sum_{i=1}^{\ell} \binom{s}{i} \varphi_{s-i,\ell-i}(m-1). \tag{12}$$

Using this induction hypothesis we obtain that $\psi_{s,\ell}(m)$ as given by (12) is a polynomial in m of degree $\le s$ whose constant term vanishes. With Lemma 1 this also shows that $\varphi_{s,\ell}(m)$ is a polynomial in m of degree $\le s$ whose constant term vanishes. When considering the leading coefficient of $\psi_{s,\ell}(m)$, which is given as follows:

$$[m^s]\psi_{s,\ell}(m) = ds\vartheta_{s-1,\ell-1,s-1} + a \sum_{k=\ell+1}^{s} \binom{k}{\ell-1} (-1)^{k-\ell-1} \vartheta_{s,k,s}, \quad \text{for } \ell \ge 2,$$

$$[m^s]\psi_{s,1}(m) = a \sum_{k=2}^{s} (-1)^k \vartheta_{s,k,s},$$

we obtain by applying Lemma 1 the recurrences for the leading coefficient $\vartheta_{s,\ell,s}$ as given in Proposition 2 and the proof of this proposition for the instance $c = 0$ is completed.

• The case $c \neq 0$. The initial values $\varphi_{0,0}(m) = \varphi_{0,0,m} = 1$ have already been computed.

The first part of the proposition concerning the functions $\varphi_{s,s}(m)$, $s \ge 1$, and the leading coefficients $\vartheta_{s,s,s}$ follows from the corresponding statements for $c = 0$, which has been proven above, since in the proof we do not require the assumption $c = 0$ and thus it also holds for $c \neq 0$.

To show that $\varphi_{s,\ell}(m) = \varphi_{s,\ell,m}$ is, for $0 \le \ell < s$, a polynomial in m of degree $\le 2s - \ell$ we use induction; we assume now that Proposition 2 holds for all $\varphi_{t,j}(m)$, with $0 \le j \le t < s$ or $(t = s$ and $\ell < j \le s$.

Since $\varphi_{s,\ell}(m)$ satisfies different recurrences for the instances $\ell > 0$ and $\ell = 0$ we have to consider both cases separately. For $0 < \ell < s$ recurrence (4a) gives

$$\varphi_{s,\ell}(m) = \frac{1}{a\ell\binom{m+\frac{a\ell}{d}}{m}} \sum_{k=1}^{m} \binom{k+\frac{a\ell}{d}-1}{k} \psi_{s,\ell}(k), \tag{13}$$

with

$$\begin{aligned}\psi_{s,\ell}(m) := a \sum_{k=\ell+1}^{s} \binom{k}{\ell-1} (-1)^{k-\ell-1} \varphi_{s,k}(m) + dm \sum_{k=\ell+1}^{s} \binom{k}{\ell} p^{k-\ell} \varphi_{s,k}(m-1) \\ + dm \sum_{i=1}^{\ell} \binom{s}{i} \sum_{k=\ell-i}^{s-i} \binom{k}{\ell-i} \varphi_{s-i,k}(m-1) p^{k-\ell+i}.\end{aligned} \tag{14}$$

Using the induction hypothesis we easily get that the first term of $\psi_{s,\ell}(m)$ as given by (14) is a polynomial in m of degree $\le 2s - \ell - 1$, whereas the second and third term are polynomials in m of degree $\le 2s - \ell$. Thus $\psi_{s,\ell}(m)$ is a polynomial in m of degree $\le 2s - \ell$ and, again by using the induction hypothesis and equation (14), that the constant term of $\psi_{s,\ell}(m)$ vanishes. Thus by applying Lemma 1 we get that also $\varphi_{s,\ell}(m)$ is, for $1 \le \ell < s$ a polynomial in m of degree $\le 2s - \ell$ whose constant term vanishes.

Considering (14) we obtain easily that the leading coefficient $[m^{2s-\ell}]\psi_{s,\ell}(m)$ of $\psi_{s,\ell}(m)$ is given as follows:

$$[m^{2s-\ell}]\psi_{s,\ell}(m) = d(\ell+1)p\vartheta_{s,\ell+1,2s-\ell-1} + ds\vartheta_{s-1,\ell-1,2s-\ell-1},$$

which shows, by applying Lemma 1, the corresponding recurrences for the leading coefficient $\vartheta_{s,\ell,2s-\ell}$, $1 \le \ell \le s-1$, as given in Proposition 2.

It remains to consider the case $\ell = 0$. Due to (4b) the functions $\varphi_{s,0}(m)$ satisfy the recurrence

$$\varphi_{s,0}(m) = \sum_{k=0}^{m-1} \psi_{s,0}(k), \quad \text{with} \quad \psi_{s,0}(m) = \sum_{i=1}^{s} \varphi_{s,i}(m) p^i. \tag{15}$$

Using the induction hypothesis we obtain that $\psi_{s,0}(m)$ is a polynomial in m of degree $\leq 2s-1$. Applying Lemma 2 we get then that $\varphi_{s,0}(m)$ is, for $s \geq 1$, a polynomial in m of degree $\leq 2s$ whose constant term vanishes.

Furthermore, we easily obtain that the leading coefficient $[m^{2s-1}]\psi_{s,0}(m)$ of $\psi_{s,0}(m)$ is given by $p\vartheta_{s,1,2s-1}$, which also shows, by applying Lemma 2, the corresponding recurrence for the leading coefficient $\vartheta_{s,0,2s}$, $s \geq 1$, as given in Proposition 2. This completes the proof of Proposition 2.

□

3.4. **The structure of the s-th moments of $A_{an,dm}$.** We summarize now our findings of the previous subsections concerning the explicit structure of $e_{n,m}^{[s]}$, i.e., the s-th moments of $A_{an,dm}$.

Proposition 3. *The s-th moments $e_{n,m}^{[s]} = \mathbb{E}(A_{an,dm}^s)$ of the discrete area $A_{an,dm}$ under lattice paths associated with diminishing urn models with a ball replacement matrix defined in* (1) *are polynomials in n and m, where the degree depends on whether $c = 0$ or $c \neq 0$.*

The functions $\varphi_{s,\ell}(m) := \varphi_{s,\ell,m}$ are, for $0 \leq \ell \leq s$, polynomials in m. The degree of these polynomials depends on whether $c = 0$ or $c \neq 0$:

- *The case $c = 0$: We have*

$$e_{n,m}^{[0]} = 1, \qquad e_{n,m}^{[s]} = \sum_{\ell=1}^{s} \sum_{j=1}^{s} \vartheta_{s,\ell,j} n^\ell m^j,$$

  where the coefficients $\vartheta_{s,\ell,j}$ are certain computable constants satisfying a system of recurrences. The recurrence for the leading coefficients $\vartheta_{s,\ell,s}$, with $1 \leq \ell \leq s$, is given in Proposition 2.
- *The case $c \neq 0$: We have*

$$e_{n,m}^{[0]} = 1, \qquad e_{n,m}^{[s]} = \sum_{\ell=0}^{s} \sum_{j=1}^{2s-\ell} \vartheta_{s,\ell,j} n^\ell m^j, \quad \text{for } s \geq 1,$$

  where the coefficients $\vartheta_{s,\ell,j}$ are certain computable constants satisfying a system of recurrences. The recurrence for the leading coefficients $\vartheta_{s,\ell,2s-\ell}$, with $0 \leq \ell \leq s$ and $s \geq 1$, is given in Proposition 2.

Finally, the exact values for the expectation and the variance, as given in Theorem 1, are obtained by straightforward computations, which are omitted here.

4. PROOFS OF THE LIMITING DISTRIBUTIONS OBTAINED WITHOUT CENTERING

In this section we give the proof of those limiting distribution results in Theorem 2 for the discrete area $A_{an,dm}$ under lattice paths associated with diminishing urn models with a ball replacement matrix defined in (1), which can be obtained without centering.

4.1. **The case n tending to infinity, with m fixed and arbitrary c.** In the case of arbitrary but fixed $m \in \mathbb{N}$ and $n \to \infty$ we proceed by applying Proposition 1. As a consequence of this proposition the moments satisfy the asymptotic expansion

$$\mathbb{E}(A_{an,dm}^s) = e_{n,m}^{[s]} = \sum_{\ell=0}^{s} \varphi_{s,\ell,m} n^\ell = n^s \varphi_{s,s,m}\big(1 + \mathcal{O}(n^{-1})\big),$$

with values $\varphi_{s,s,m}$ given recursively by

$$\varphi_{s,s,m} = \frac{1}{as\binom{m+\frac{as}{d}}{m}} \sum_{k=1}^{m} \binom{k+\frac{as}{d}-1}{k} \left[dk \sum_{i=1}^{s} \binom{s}{i} \varphi_{s-i,s-i,k-1} \right], \quad \text{for } s \geq 1, \quad \varphi_{0,0,m} = 1. \tag{16}$$

Furthermore, we obtain after normalization the expansion

$$\mathbb{E}\Big(\big(\frac{A_{an,dm}}{n}\big)^s\Big) = \frac{e_{n,m}^{[s]}}{n^s} = \varphi_{s,s,m}\big(1+\mathcal{O}(n^{-1})\big).$$

Hence the s-th moments of the scaled random variable $A_{an,dm}/n$ converge, for m fixed and $n \to \infty$, to the values $\varphi_{s,s,m}$ as given by (16). Due to obvious geometric reasons $A_{an,dm}$ is bounded by $n + (n+p) + (n+2p) + \cdots + (n+(m-1)p) = nm + pm(m-1)/2$. Thus $A_{an,dm} \le 2mn$, for m fixed and n large enough, which shows

$$\varphi_{s,s,m} = \lim_{n\to\infty} \mathbb{E}\big(\big(\frac{A_{an,dm}}{n}\big)^s\big) \le (2m)^s.$$

Thus it follows by an application of Carlemans criterion that the sequence of moments $(\varphi_{s,s,m})_{s\ge1}$ as given by (16) uniquely describes a random variable X_m. Therefore we can apply the Theorem of Fréchet and Shohat, which proves that $A_{an,dm}/n$ converges weakly to a random variable X_m characterized by its moments:

$$\frac{A_{an,dm}}{n} \xrightarrow{\mathcal{L}} X_m, \quad \text{with} \quad \mathbb{E}(X_m^s) = \varphi_{s,s,m}.$$

In order to show the characterization of X_m via the distributional equation given in Theorem 2 one simply checks by straightforward calculations that the resulting recurrences for the sequence of the s-th moments of X_m as defined in this theorem match with recurrence (16), which proves the distributional equation

$$X_m \stackrel{\mathcal{L}}{=} Y_m \cdot (1 + X_{m-1}), \quad \text{for } m \ge 1, \quad X_0 = 0,$$

where $Y_m \stackrel{\mathcal{L}}{=} B\big(\frac{dm}{a}, 1\big)$ being independent of X_0, X_1, $\dots$, and where $B(\alpha,\beta)$ denotes a Beta-distributed random variable with parameters α and β. By unwinding this equation we obtain the explicit form stated in Theorem 2. Note that on the level of the moments this correponds to the unwinding of the recurrence relation (16) for the s-th moment $\varphi_{s,s,m}$ of X_m.

Remark 3. In the following we also give a (non-rigorous) probabilistic argument for the distributional equation of X_m as proven before. To do this we rewrite the distribution equation of $A_{an,dm}$ as given in (2) in the following form:

$$A_{an,dm} \stackrel{\mathcal{L}}{=} A_{a(Y_{n,m}+p),d(m-1)} + Y_{n,m}, \quad \text{with } A_{an,0} = 0,$$

where

$$\mathbb{P}\{Y_{n,m} = k\} = \frac{\binom{k-1+\frac{dm}{a}}{k}}{\binom{n+\frac{dm}{a}}{n}}, \quad \text{for } 0 \le k \le n.$$

Note that the random variable $Y_{n,m}$ counts the contribution to $A_{an,dm}$ stemming from the m-th column of the grid $\mathbb{N}^2$. It can be checked easily that, for m fixed and n tending to infinity, the normalized random variable $Y_{n,m}/n$ converges to a Beta distributed random variable with parameters dm/a and 1. Since we already know that $A_{an,dm}/n$ converges weakly to a random variable X_m, we expect that, by formally taking the limits in equation

$$\frac{A_{an,dm}}{n} \stackrel{\mathcal{L}}{=} \frac{A_{a(Y_{n,m}+p),d(m-1)}}{Y_{n,m}+p} \cdot \frac{Y_{n,m}+p}{n} + \frac{Y_{n,m}}{n},$$

the random variable X_m satisfies the recursive distributional equation stated in Theorem 2. We expect that this formal argument can be made rigorous, but we do not pursue this direction, since above we have already proven the distributional equation using the recursive description of the moments.

4.2. **The case m tending to infinity, with n fixed and $c = 0$.** For urns with $c = 0$ we use, for the case n fixed and $m \to \infty$, Proposition 3:

$$\mathbb{E}(A_{an,dm}^s) = e_{n,m}^{[s]} = \sum_{\ell=1}^{s} n^\ell \sum_{j=1}^{s} \vartheta_{s,\ell,j} m^j,$$

which gives

$$\mathbb{E}\Big(\big(\frac{A_{an,dm}}{m}\big)^s\Big) = \sum_{\ell=0}^{s} n^\ell \vartheta_{s,\ell,s} + \mathcal{O}\big(\frac{1}{m}\big).$$

Hence the s-th moments of the normalized random variable $A_{an,dm}/m$ converge, for n fixed and $m \to \infty$, to $\sum_{\ell=0}^{s} n^\ell \vartheta_{s,\ell,s}$, where the values $\vartheta_{s,\ell,s}$ are certain computable constants satisfying some recurrence equations;

see Proposition 2. For $c = 0$ we trivially obtain $A_{an,dm} \le mn$ and thus the estimates $\mathbb{E}((A_{an,dm}/m)^s) \le n^s$, for $s \ge 1$, that allows to apply Carlemans criterion to show that $A_{an,dm}/m$ converges weakly to a random variable W_n, which is characterized by the sequence of its s-th moments:

$$\frac{A_{an,dm}}{m} \xrightarrow{\mathcal{L}} W_n, \quad \mathbb{E}(W_n^s) = \sum_{\ell=0}^{s} n^\ell \vartheta_{s,\ell,s}.$$

In order to show the characterization of W_n via the distributional equation given in Theorem 2 one simply checks again that the resulting recurrences for the sequence of the s-th moments of W_n as defined in this theorem match with the recurrence relation for the moments $\vartheta_{s,\ell,s}$ stated in Proposition 2, which proves the stated result.

Remark 4. An alternative approach to characterize the limiting distribution of W_n by using more combinatorial arguments and a certain symmetry relation can be given also. We use there that, for the instance $c = 0$, any path starting at state (dm, an) and ending at an absorbing state partitions the $m \cdot n$ rectangles (with length a and d, respectively) filling the area $(0, 0) \le (x, y) \le (dm, an)$ into two parts: one part counts of the number of rectangles below the path (corresponding to $A_{an,dm}$), and another part counts the number of rectangles to the left of the path (corresponding to a new random variable $F_{dm,an}$). In terms of the random variables $A_{an,dm}$ and $F_{dm,an}$ we have the relation

$$A_{an,dm} + F_{dm,an} \stackrel{\mathcal{L}}{=} mn \quad \text{or equivalently} \quad A_{an,dm} \stackrel{\mathcal{L}}{=} mn - F_{dm,an}. \tag{17}$$

Note that by a symmetry argument the random variable $F_{dm,an}$ also counts the discrete area below sample paths, starting at state (an, dm) and ending at an absorbing state, of urns associated with a ball replacement matrix $\left(\begin{smallmatrix} -d & 0 \\ 0 & -a \end{smallmatrix}\right)$. Hence, for n fixed and $m \to \infty$, we can use our earlier results (stated for m fixed and $n \to \infty$) obtained in Subsection 4.1 and get, as a consequence, the stated distribution law

$$\frac{A_{an,dm}}{m} \xrightarrow{\mathcal{L}} W_m(a, d), \quad \text{with } W_m(a, d) \stackrel{\mathcal{L}}{=} n - X_n(d, a).$$

Remark 5. As before, we also give a (non-rigorous) probabilistic argument for the distributional equation of W_n given in Theorem 2. We rewrite the distribution equation of $A_{an,dm}$ given in (2) in the following form:

$$A_{an,dm} \stackrel{\mathcal{L}}{=} A_{a(n-1),d(m-Z_{n,m})} + nZ_{n,m}, \quad \text{with } A_{an,0} = 0,$$

where

$$\mathbb{P}\{Z_{n,m} = m - k\} = \frac{\binom{k-1+\frac{an}{d}}{k}}{\binom{m+\frac{an}{d}}{m}}, \quad \text{for } 0 \le k \le m.$$

Note that the random variable $Z_{n,m}$ counts the contribution to $A_{an,dm}$ stemming from the m-th row of the grid $\mathbb{N}^2$. It can be checked easily that, for n fixed and $m \to \infty$, the normalized random variable $(m - Z_{n,m})/m$ converges to a Beta distributed random variable with parameters an/d and 1. Hence, by formally taking the limits of the equation below,

$$\frac{A_{an,dm}}{m} \stackrel{\mathcal{L}}{=} \frac{A_{a(n-1),d(m-Z_{n,m})}}{m - Z_{n,m}} \cdot \frac{m - Z_{n,m}}{m} + \frac{nZ_{n,m}}{m},$$

we expect the recursive distributional equation for W_n as given in Theorem 2. It should be possible to make this formal argument rigorous but again, we do not pursue this direction, since we have already proven the distributional equation using the recursive description of the moments.

5. Proof of the normal limit laws

Here we give the proof of the Gaussian limiting distribution results for the discrete area $A_{an,dm}$ stated in Theorem 2. Again we will apply the method of moments to show these results, but in order to handle the massive cancellations for the centered moments of $A_{an,dm}$ appearing in the instance $c \neq 0$ and $m \to \infty$ and the instance $c = 0$ and $m, n \to \infty$ we have to study them in detail. Thus we introduce the centered random variable

$$\hat{A}_{an,dm} := A_{an,dm} - \mathbb{E}(A_{an,dm}) = A_{an,dm} - \frac{nm}{1 + \frac{a}{d}} - \frac{cm(m-1)}{2a(1 + \frac{a}{d})}, \tag{18}$$

where we used the explicit formula for $\mathbb{E}(A_{an,dm})$ stated in Theorem 1, and their s-th moments (i.e., the s-th centered moments of the discrete area $A_{an,dm}$):

$$\hat{e}_{n,m}^{[s]} := \mathbb{E}(\hat{A}_{an,dm}^{s}) = \mathbb{E}\Big(\big(A_{an,dm} - \mathbb{E}(A_{an,dm})\big)^{s}\Big). \tag{19}$$

In Subsection 5.1 we give a recursive description for $\hat{e}_{n,m}$, which is used to show the explicit structure of $\hat{e}_{n,m}$ stated in Subsection 5.2-5.3. Of course, due to Proposition 3 it is obvious that the centered moments $\hat{e}_{n,m}$ of $A_{an,dm}$, when considering them as functions depending on m and n, are polynomials in m and n, but it is of great importance for the approach used to show suitable bounds on the degree of the polynomials. Furthermore we obtain a recurrence for the leading coefficients, which is studied in Subsection 5.4 leading to an explicit formula for these leading coefficients. In Subsection 5.5-5.6 we determine asymptotic results for the s-th centered moments of $A_{an,dm}$, which leads, after applying the Theorem of Fréchet and Shohat, to the Gaussian limiting distribution results stated in Theorem 2.

5.1. **A recurrence for the centered moments of $A_{an,dm}$.** We start with the distributional equation (2) for the discrete area $A_{an,dm}$ under lattice paths associated with diminishing urns and get by using the explicit formula for the expectation as given in Theorem 1 the following distributional equation for the centered random variable $\hat{A}_{an,dm} := A_{an,dm} - \mathbb{E}(A_{an,dm})$:

$$\hat{A}_{an,dm} \stackrel{\mathcal{L}}{=} \mathbb{I}_{n,m}\Big(\hat{A}_{a(n-1),m} - \frac{md}{a+d}\Big) + (1-\mathbb{I}_{n,m})\Big(\hat{A}'_{an,d(m-1)} + \frac{na}{a+d}\Big), \tag{20}$$

where $\mathbb{I}_{n,m}$ denotes the indicator variable of the event "choosing a white ball",

$$\mathbb{P}\{\mathbb{I}_{n,m} = 1\} = \frac{an}{an+dm}, \quad \mathbb{P}\{\mathbb{I}_{n,m} = 0\} = \frac{dm}{an+dm},$$

with $\mathbb{I}_{n,m}$ being independent of $(\hat{A}_{an,dm})_{n,m\ge 0}$ and $(\hat{A}'_{an,dm})_{n,m\ge 0}$, which are independent copies of each other.

The distributional equation (20) immediately leads to the following recurrences for the s-th moments $\hat{e}_{n,m}^{[s]} := \mathbb{E}\big(\hat{A}_{an,dm}^{s}\big)$:

$$\hat{e}_{n,m}^{[s]} = \frac{an}{an+dm}\sum_{\ell=0}^{s}\binom{s}{\ell}\hat{e}_{n-1,m}^{[s-\ell]}\frac{(-m)^{\ell}d^{\ell}}{(a+d)^{\ell}} + \frac{dm}{an+dm}\sum_{\ell=0}^{s}\binom{s}{\ell}\hat{e}_{n+p,m-1}^{[s-\ell]}\frac{n^{\ell}a^{\ell}}{(a+d)^{\ell}}, \tag{21}$$

with initial conditions $\hat{e}_{n,0}^{[s]} = 0$, for $n \ge 0$ and $s \ge 1$, and $\hat{e}_{n,m}^{[0]} = 1$, for $m, n \in \mathbb{N}_0$.

5.2. **The structure of the centered moments of $A_{an,dm}$: a crude analysis.** We start here to study the s-th centered moments $\hat{e}_{n,m}^{[s]}$ of $A_{an,dm}$, which are given recursively via equation (21). The following lemma states that, when considering $\hat{e}_{n,m}^{[s]}$ as a function depending on n and m, $\hat{e}_{n,m}^{[s]}$ is a polynomial in n of degree at most s, with certain recursively computable coefficients. Of course, in principle this also follows directly from the corresponding result, i.e., Proposition (1), for the ordinary s-th moments of $A_{an,dm}$, but it will be important to have a suitable description of the coefficients appearing and this is provided also in this lemma.

Lemma 3. *The centered moments $\hat{e}_{n,m}^{[s]} = \mathbb{E}\big((A_{an,dm} - \mathbb{E}(A_{an,dm})^{s}\big)$ of the random variable $A_{an,dm}$ satisfy the expansion*

$$\hat{e}_{n,m}^{[s]} = \sum_{\ell=0}^{s}\hat{\varphi}_{s,\ell,m}n^{\ell},$$

with coefficients $\hat{\varphi}_{s,\ell,m}$ that can be computed recursively as follows, where the initial values are given by $\hat{\varphi}_{s,\ell,0} = 0$, for $0 \le \ell \le s$, $s \ge 1$, and $\hat{\varphi}_{0,0,m} = 1$, for $m \ge 0$.

- *For $1 \le \ell \le s$ and $m \ge 1$ the values $\hat{\varphi}_{s,\ell,m}$ satisfy the following recurrence:*

$$\hat{\varphi}_{s,\ell,m} = \frac{1}{a\ell\binom{m+\frac{a\ell}{d}}{m}}\sum_{k=1}^{m}\binom{k+\frac{a\ell}{d}-1}{k}\hat{\psi}_{s,\ell,k}, \tag{22a}$$

with

$$\hat{\psi}_{s,\ell,m} := a\sum_{k=\ell+1}^{s}\binom{k}{\ell-1}(-1)^{k-\ell+1}\hat{\varphi}_{s,k,m}$$

$$+ a \sum_{i=\ell-1}^{s-1} \sum_{k=\ell-1}^{i} \binom{s}{i} \binom{k}{\ell-1} \frac{(-m)^{s-i} d^{s-i}}{(a+d)^{s-i}} (-1)^{k-\ell+1} \hat{\varphi}_{i,k,m}$$
$$+ dm \sum_{i=s-\ell}^{s-1} \sum_{k=i+\ell-s}^{i} \binom{s}{i} \binom{k}{i+\ell-s} \frac{a^{s-i}}{(a+d)^{s-i}} (\frac{c}{a})^{s-\ell+k-i} \hat{\varphi}_{i,k,m-1}$$
$$+ dm \sum_{k=\ell+1}^{s} \binom{k}{\ell} (\frac{c}{a})^{k-\ell} \hat{\varphi}_{s,k,m-1}.$$

- *For $\ell = 0$ and $m \geq 1$ the values $\hat{\varphi}_{s,\ell,m}$ satisfy the following recurrence:*

$$\hat{\varphi}_{s,0,m} = \sum_{k=0}^{m-1} \hat{\psi}_{s,0,k}, \quad \textit{with} \quad \hat{\psi}_{s,0,m} := \sum_{i=1}^{s} (\frac{c}{a})^{i} \hat{\varphi}_{s,i,m}. \tag{22b}$$

Proof. The proof of this lemma is completely analogous to the proof of Proposition 1 and thus we will only sketch the computations.

In order to prove the stated expansion of $\hat{e}_{n,m}^{[s]}$ for $s \geq 1$, $m \geq 1$ and $n \geq 0$ one can use again the *Ansatz*: $\hat{e}_{n,m}^{[s]} = \sum_{k=0}^{s} \hat{\varphi}_{s,k,m} n^k$ and plug it into recurrence (21). This leads to the following equation:

$$(an + dm) \sum_{k=0}^{s} \hat{\varphi}_{s,k,m} n^k = an \sum_{\ell=0}^{s} \binom{s}{\ell} \big(-\frac{d}{a+d} m\big)^{\ell} \sum_{k=0}^{s-\ell} \hat{\varphi}_{s-\ell,k,m} (n-1)^k$$
$$+ dm \sum_{\ell=0}^{s} \binom{s}{\ell} \big(\frac{a}{a+d} n\big)^{\ell} \sum_{k=0}^{s-\ell} \hat{\varphi}_{s-\ell,k,m-1} (n+p)^k. \tag{23}$$

By comparing the coefficients of n^ℓ, for $0 \leq \ell \leq s+1$, in equation (23) we eventually obtain by simple manipulations and using $c = pa$ a system of $s+1$ equations (the equation obtained for $\ell = s+1$ cancels out) for $\hat{\varphi}_{s,\ell,m}$, $m \geq 1$:

$$(dm + a\ell) \hat{\varphi}_{s,\ell,m} = dm \hat{\varphi}_{s,\ell,m-1} + \hat{\psi}_{s,\ell,m}, \quad \text{for } 1 \leq \ell \leq s,$$
$$\text{with} \quad \hat{\psi}_{s,\ell,m} := a \sum_{k=\ell+1}^{s} \binom{k}{\ell-1} (-1)^{k-\ell+1} \hat{\varphi}_{s,k,m}$$
$$+ a \sum_{i=\ell-1}^{s-1} \sum_{k=\ell-1}^{i} \binom{s}{i} \binom{k}{\ell-1} \frac{(-m)^{s-i} d^{s-i}}{(a+d)^{s-i}} (-1)^{k-\ell+1} \hat{\varphi}_{i,k,m}$$
$$+ dm \sum_{i=s-\ell}^{s-1} \sum_{k=i+\ell-s}^{i} \binom{s}{i} \binom{k}{i+\ell-s} \frac{a^{s-i}}{(a+d)^{s-i}} (\frac{c}{a})^{s-\ell+k-i} \hat{\varphi}_{i,k,m-1}$$
$$+ dm \sum_{k=\ell+1}^{s} \binom{k}{\ell} (\frac{c}{a})^{k-\ell} \hat{\varphi}_{s,k,m-1},$$
$$\hat{\varphi}_{s,0,m} = \hat{\varphi}_{s,0,m-1} + \hat{\psi}_{s,0,m-1}, \quad \text{with} \quad \hat{\psi}_{s,0,m} = \sum_{i=1}^{s} (\frac{c}{a})^{i} \hat{\varphi}_{s,i,m}.$$

These recurrences can be considered as linear first order recurrences for the parameter m. Solving them justifies the Ansatz $\hat{e}_{n,m}^{[s]} = \sum_{k=0}^{s} \hat{\varphi}_{s,k,m} n^k$ and leads to the results stated in Lemma 3. □

We remark that, as a consequence of $\hat{e}_{n,m}^{[1]} = 0$, one obtains the values $\hat{\varphi}_{1,1,m} = \hat{\varphi}_{1,0,m} = 0$.

5.3. **The structure of the centered moments of $A_{an,dm}$: a refined analysis.** Next we are going to describe the values $\hat{\varphi}_{s,\ell,m}$ appearing in Lemma 3, when we consider them as functions depending on m.

Lemma 4. *The functions $\hat{\varphi}_{s,\ell}(m) := \hat{\varphi}_{s,\ell,m}$ are, for $0 \le \ell \le s$, polynomials in m. We have $\hat{\varphi}_{0,0}(m) = 1$, whereas for $s \ge 1$ the constant term of $\hat{\varphi}_{s,\ell}(m)$ vanishes. Furthermore, for all $0 \le \ell \le s$, the degree of the polynomials $\hat{\varphi}_{s,\ell}(m)$ is bounded as follows:*

$$\deg \hat{\varphi}_{s,\ell}(m) \le \lfloor \frac{3s}{2} \rfloor - \ell. \tag{24}$$

Moreover, for s even, let us denote by

$$\gamma_{s,\ell} := \operatorname{lc} \hat{\varphi}_{s,\ell}(m) = [m^{\frac{3s}{2}-\ell}]\hat{\varphi}_{s,\ell}(m)$$

the leading coefficient of the polynomial $\hat{\varphi}_{s,\ell}(m)$. Then these leading coefficients $\gamma_{s,\ell}$ satisfy, for $0 \le \ell \le s$ and $s \ge 2$ even, the recurrence

$$\gamma_{s,\ell} = \frac{1}{\frac{3ds}{2} + (a-d)\ell}\bigg(\frac{cd(\ell+1)}{a}\gamma_{s,\ell+1} + a\binom{s}{2}\big(\frac{a}{a+d}\big)^2\gamma_{s-2,\ell-1} + d\binom{s}{2}\big(\frac{a}{a+d}\big)^2\gamma_{s-2,\ell-2}\bigg), \tag{25}$$

with initial value $\gamma_{0,0} = 1$, and $\gamma_{s,\ell} = 0$, for $\ell < 0$ or $\ell > s$.

Proof. The first part of the lemma, which states that the functions $\hat{\varphi}_{s,\ell}(m)$ are polynomials with respect to the parameter m, whose constant term vanish, for all $s \ge 1$, is an immediate consequence of the polynomial structure of the ordinary moments $e^{[s]}_{n,m} = \mathbb{E}(A^s_{an,dm})$ given in Proposition 3. Of course, one could show this result also by considering recurrence (22) for $\hat{\varphi}_{s,\ell}(m)$ and using induction in a way analogous to the proof of Proposition 2.

It remains to prove the degree bound for the polynomials $\hat{\varphi}_{s,\ell}(m)$, $0 \le \ell \le s$, stated in Lemma 4. Since $\hat{\varphi}_{0,0}(m) = \hat{\varphi}_{0,0,m} = 1$ the lemma is true for $s = 0$. To show that the degree bound for $\varphi_{s,\ell}(m)$ also holds for $s \ge 1$ we use induction on s and ℓ; we assume now that Lemma 4 holds for all $\varphi_{t,j}(m)$, with $0 \le j \le t < s$ or $(t = s$ and $\ell < j \le s$.

We distinguish now between the cases s odd and s even and further between $1 \le \ell \le s$ and $\ell = 0$.

(i) The case s odd and $1 \le \ell \le s$.

We use the induction hypothesis and consider the functions $\hat{\psi}_{s,\ell}(m) := \hat{\psi}_{s,\ell,m}$ defined in Lemma (3):

$$\begin{aligned}\hat{\psi}_{s,\ell}(m) &:= a\sum_{k=\ell+1}^{s}\binom{k}{\ell-1}(-1)^{k-\ell+1}\hat{\varphi}_{s,k}(m) \\ &+ a\sum_{i=\ell-1}^{s-1}\sum_{k=\ell-1}^{i}\binom{s}{i}\binom{k}{\ell-1}\frac{(-m)^{s-i}d^{s-i}}{(a+d)^{s-i}}(-1)^{k-\ell+1}\hat{\varphi}_{i,k}(m) \\ &+ dm\sum_{i=s-\ell}^{s-1}\sum_{k=i+\ell-s}^{i}\binom{s}{i}\binom{k}{i+\ell-s}\frac{a^{s-i}}{(a+d)^{s-i}}\big(\frac{c}{a}\big)^{s-\ell+k-i}\hat{\varphi}_{i,k}(m-1) \\ &+ dm\sum_{k=\ell+1}^{s}\binom{k}{\ell}\big(\frac{c}{a}\big)^{k-\ell}\hat{\varphi}_{s,k}(m-1).\end{aligned} \tag{26}$$

We obtain then the following degree bounds on the summands of the four sums appearing in equation (26):

- $\deg \hat{\varphi}_{s,k}(m) \le \lfloor \frac{3s}{2} \rfloor - k = \frac{3s-1}{2} - k \le \frac{3s-3}{2} - \ell$, in the range $\ell+1 \le k \le s$, (1. sum)
- $\deg(m^{s-i}\hat{\varphi}_{i,k,m}) \le \lfloor \frac{3i}{2} \rfloor - k + s - i \le \frac{3s+1}{2} - i$, in the range $\ell - 1 \le i \le s-1$ and $\ell - 1 \le k \le i$, (2. sum)
- $\deg(m\hat{\varphi}_{i,k,m-1}) \le \lfloor \frac{3i}{2} \rfloor - k + 1 = \lfloor \frac{i}{2} \rfloor + 1 + i - k \le \frac{s-1}{2} + 1 + s - \ell \le \frac{3s+1}{2} - \ell$, in the range $s - \ell \le i \le s-1$ and $i + \ell - s \le k \le i$, (3. sum)
- $\deg(m\hat{\varphi}_{s,k,m-1}) \le \lfloor \frac{3s}{2} \rfloor - k + 1 = \frac{3s+1}{2} - k \le \frac{3s-1}{2} - \ell$, in the range $\ell + 1 \le k \le s$. (4. sum)

We will take now a closer look on the main contributions of $\hat{\psi}_{s,\ell}(m)$ stemming from the 2. sum and the 3. sum, i.e., the summands in the 2. sum and 3. sum with $i = s-1$ and $k = \ell - 1$. When combining these contributions we obtain:

$$-\frac{ads}{a+d}m\hat{\varphi}_{s-1,\ell-1}(m) + \frac{ads}{a+d}m\hat{\varphi}_{s-1,\ell-1}(m-1) = \frac{ads}{a+d}m\big(\hat{\varphi}_{s-1,i-1}(m) - \hat{\varphi}_{s-1,i-1}(m-1)\big). \tag{27}$$

Hence, the dominant terms in (27) cancel out and we obtain the degree bound

$$\deg\Big(\frac{ads}{a+d}m\big(\hat{\varphi}_{s-1,\ell-1}(m)-\hat{\varphi}_{s-1,\ell-1}(m-1)\big)\Big)\le 1+\lfloor\frac{3(s-1)}{2}\rfloor-\ell+1-1=\frac{3s-1}{2}-\ell.$$

It can be checked easily that all the remaining terms of the 2. sum and 3. sum satisfy the degree bound $\frac{3s-1}{2}-\ell$. This, together with the degree bounds on the terms of the 1. sum and 4. sum, implies the degree bound $\hat{\psi}_{s,\ell}(m)\le\frac{3s-1}{2}-\ell$ for the polynomial $\hat{\psi}_{s,\ell}(m)$. Finally applying Lemma 1 to the sum

$$\hat{\varphi}_{s,\ell}(m)=\frac{1}{a\ell\binom{m+\frac{a\ell}{d}}{m}}\sum_{k=1}^{m}\binom{k+\frac{a\ell}{d}-1}{k}\hat{\psi}_{s,\ell}(k), \tag{28}$$

shows that the polynomial $\hat{\varphi}_{s,\ell}(m)$ also satisfies the degree bound $\hat{\varphi}_{s,\ell}(m)\le\frac{3s-1}{2}-\ell=\lfloor\frac{3s}{2}\rfloor-\ell$.

(ii) The case s odd and $\ell=0$.

Again we use the induction hypothesis and consider the functions $\hat{\psi}_{s,0}(m):=\hat{\psi}_{s,0,m}$ defined in Lemma (3):

$$\hat{\psi}_{s,0}(m)=\sum_{i=1}^{s}\big(\frac{c}{a}\big)^i\hat{\varphi}_{s,i}(m). \tag{29}$$

We obtain then the following degree bound on the summands of $\hat{\psi}_{s,0}(m)$:

$$\deg\hat{\varphi}_{s,i}(m)\le\lfloor\frac{3s}{2}\rfloor-i=\frac{3s-1}{2}-i\le\frac{3s-3}{2},\quad\text{in the range }\ 1\le i\le s.$$

This implies the degree bound $\hat{\psi}_{s,0}(m)\le\frac{3s-3}{2}$ for the polynomial $\hat{\psi}_{s,0}(m)$. Applying Lemma 2 to the sum

$$\hat{\varphi}_{s,0}(m)=\sum_{k=0}^{m-1}\hat{\psi}_{s,0}(k) \tag{30}$$

shows then that the polynomial $\hat{\varphi}_{s,\ell}(m)$ satisfies the degree bound $\hat{\varphi}_{s,0}(m)\le\frac{3s-1}{2}=\lfloor\frac{3s}{2}\rfloor$.

(iii) The case s even and $1\le\ell\le s$.

Using the induction hypothesis and considering the functions $\hat{\psi}_{s,\ell}(m):=\hat{\psi}_{s,\ell,m}$ we obtain the following degree bounds on the summands of the four sums appearing in equation (26):

- $\deg\hat{\varphi}_{s,k}(m)\le\lfloor\frac{3s}{2}\rfloor-k=\frac{3s}{2}-k\le\frac{3s-2}{2}-\ell$, in the range $\ell+1\le k\le s$, (1. sum)
- $\deg(m^{s-i}\hat{\varphi}_{i,k}(m)\le\lfloor\frac{3i}{2}\rfloor-k+s-i\le\lfloor\frac{i}{2}\rfloor+s-k$, in the range $\ell-1\le i\le s-1$ and $\ell-1\le k\le i$, (2. sum)
- $\deg(m\hat{\varphi}_{i,k}(m-1)\le\lfloor\frac{3i}{2}\rfloor-k+1=\lfloor\frac{i}{2}\rfloor+1+i-k$, in the range $s-\ell\le i\le s-1$ and $i+\ell-s\le k\le i$, (3. sum)
- $\deg(m\hat{\varphi}_{s,k}(m-1)\le\lfloor\frac{3s}{2}\rfloor-k+1=\frac{3s}{2}-k+1$, in the range $\ell+1\le k\le s$. (4. sum)

It can be checked easily that all the terms of the 1. sum, 2. sum, 3. sum and 4. sum satisfy the degree bound $\frac{3s}{2}-\ell$. This implies the degree bound $\hat{\psi}_{s,\ell}(m)\le\frac{3s}{2}-\ell$ for the polynomial $\hat{\psi}_{s,\ell}(m)$. Applying Lemma 1 to (28) shows that the polynomial $\hat{\varphi}_{s,\ell}(m)$ also satisfies the degree bound $\hat{\varphi}_{s,\ell}(m)\le\frac{3s}{2}-\ell=\lfloor\frac{3s}{2}\rfloor-\ell$.

Furthermore we will obtain a recursive description of the leading coefficients $\gamma_{s,\ell}=\operatorname{lc}\hat{\varphi}_{s,\ell}=[m^{\frac{3s}{2}-\ell}]\hat{\varphi}_{s,\ell}(m)$ when determining the leading coefficient $\operatorname{lc}\hat{\psi}_{s,\ell}(m)=[m^{\frac{3s}{2}-\ell}]\hat{\psi}_{s,\ell}(m)$ of $\hat{\psi}_{s,\ell}(m)$, which we are carrying out now. We first remark that as in the case of s odd the dominant terms in the 2. sum and 3. sum with $i=s-1$ and $k=\ell-1$ cancel out and do not contribute to the leading coefficient of $\hat{\psi}_{s,\ell}(m)$, since we obtain for these summands the degree bound

$$\deg\Big(\frac{ads}{a+d}m\big(\hat{\varphi}_{s-1,\ell-1}(m)-\hat{\varphi}_{s-1,\ell-1}(m-1)\big)\Big)\le\frac{3s-2}{2}-\ell.$$

When considering the remaining summands of $\hat{\psi}_{s,\ell}(m)$ we obtain that only the following three summands contribute to the leading coefficient of $\hat{\psi}_{s,\ell}(m)$, whereas the degree of all other terms is bounded by $\frac{3s-2}{2}-\ell$:

- (2. sum) The term $i = s-2$ and $k = i-1$ contributes as follows:

$$\mathrm{lc}\left(a\binom{s}{s-2}\binom{\ell-1}{\ell-1}\frac{(-m)^{s-(s-2)}d^{s-(s-2)}}{(a+d)^{s-(s-2)}}(-1)^{\ell-1-\ell+1}\hat{\varphi}_{s-2,\ell-1}(m)\right)$$
$$= a\binom{s}{2}\frac{d^2}{(a+d)^2}\gamma_{s-2,\ell-1}.$$

- (3. sum) The term $i = s-2$ and $k = i-2$ contributes as follows:

$$\mathrm{lc}\left(dm\binom{s}{s-2}\binom{\ell-1}{\ell-1}\frac{a^{s-(s-2)}}{(a+d)^{s-(s-2)}}(\frac{c}{a})^{s-\ell+\ell-2-s+2}\hat{\varphi}_{s-2,\ell-2}(m-1)\right)$$
$$= d\binom{s}{2}\frac{a^2}{(a+d)^2}\gamma_{s-2,\ell-2}.$$

- (4. sum) The term $k = \ell+1$ contributes as follows:

$$\mathrm{lc}\left(dm\binom{\ell+1}{\ell}(\frac{c}{a})^{\ell+1-\ell}\hat{\varphi}_{s,\ell+1}(m-1)\right) = d(\ell+1)\frac{c}{a}\,\mathrm{lc}(\hat{\varphi}_{s,\ell+1}(m-1) = \frac{cd(\ell+1)}{a}\gamma_{s,\ell+1}.$$

Collecting all these contributions gives

$$\mathrm{lc}\,\hat{\psi}_{s,\ell}(m) = \frac{cd(\ell+1)}{a}\gamma_{s,\ell+1} + a\binom{s}{2}\frac{d^2}{(a+d)^2}\gamma_{s-2,\ell-1} + d\binom{s}{2}\frac{a^2}{(a+d)^2}\gamma_{s-2,\ell-2}.$$

Applying Lemma 1 shows then, for $1 \le \ell \le s$, the recurrence for $\gamma_{s,\ell}$ stated in Lemma 4.

(iv) The case s even and $\ell = 0$.

Using the induction hypothesis and considering the functions $\hat{\psi}_{s,0}(m) := \hat{\psi}_{s,0,m}$ we obtain the following degree bound on the summands of $\hat{\psi}_{s,0}(m)$ as given in equation (29):

$$\deg\hat{\varphi}_{s,i}(m) \le \lfloor\frac{3s}{2}\rfloor - i = \frac{3s}{2} - i, \quad \text{in the range } 1 \le i \le s.$$

This implies the degree bound $\hat{\psi}_{s,0}(m) \le \frac{3s}{2} - 1$ for the polynomial $\hat{\psi}_{s,0}(m)$. Applying Lemma 2 to the sum (30) shows then that the polynomial $\hat{\varphi}_{s,\ell}(m)$ satisfies the degree bound $\hat{\varphi}_{s,0}(m) \le \frac{3s}{2} = \lfloor\frac{3s}{2}\rfloor$.

When considering the summands of $\hat{\psi}_{s,0}(m)$ it is easily seen that only the term $i = 1$ contributes to the leading coefficient $\mathrm{lc}\,\hat{\psi}_{s,0}(m) = [m^{\frac{3s}{2}-1}]\hat{\psi}_{s,0}(m)$ of $\hat{\psi}_{s,0}(m)$ and we get

$$\mathrm{lc}\,\hat{\psi}_{s,0}(m) = \mathrm{lc}\left(\frac{c}{a}\hat{\varphi}_{s,1}(m)\right) = \frac{c}{a}\gamma_{s,1}.$$

Applying Lemma 2 shows then also the instance $\ell = 0$ of the recurrence for $\gamma_{s,\ell}$ stated in Lemma 4. □

5.4. **Computing the leading coefficients.** We are going to study now, for s even, the recurrence (25) for the leading coefficients $\gamma_{s,\ell} := \mathrm{lc}\,\hat{\varphi}_{s,\ell}(m)$ of the polynomials $\varphi_{s,\ell}(m)$, which will be an important step in our characterization of the limiting distribution of $A_{an,dm}$. To do this it is advantageous to introduce the numbers

$$\tilde{\gamma}_{s,\ell} := \gamma_{s,s-\ell}$$

and study the corresponding recurrence for $\tilde{\gamma}_{s,\ell}$ (which follows immediately from (25)) that holds for $0 \le \ell \le s$ and $s \ge 2$ even:

$$\tilde{\gamma}_{s,\ell} = \frac{1}{\frac{3ds}{2} + (a-d)(s-\ell)} \times \left(\frac{cd(s-\ell+1)}{a}\tilde{\gamma}_{s,\ell-1} + a\binom{s}{2}(\frac{a}{a+d})^2\tilde{\gamma}_{s-2,\ell-1} + d\binom{s}{2}(\frac{a}{a+d})^2\tilde{\gamma}_{s-2,\ell}\right), \tag{31}$$

with initial value $\tilde{\gamma}_{0,0} = 1$, and $\tilde{\gamma}_{s,\ell} = 0$, for $\ell < 0$ or $\ell > s$, and $\tilde{\gamma}_{s,\ell} = 0$, for s odd.

We introduce now via

$$C(z,w) := \sum_{s\ge0}\sum_{\ell\ge0}\tilde{\gamma}_{s,\ell}\frac{z^s}{s!}w^\ell = \sum_{s\ge0}\sum_{\ell\ge0}\gamma_{s,s-\ell}\frac{z^s}{s!}w^\ell$$

a suitable bivariate generating function of the sequence $\tilde{\gamma}_{s,\ell}$. When multiplying recurrence (31) by $\frac{z^s}{s!}w^\ell$ and summing up, for the values $0 \le \ell \le s$ and $s \ge 2$ even, we obtain (after straightforward computations that are omitted here) the following linear first order partial differential equation for the generating function $C(z,w)$:

$$z(a+\frac{d}{2}-\frac{cd}{a}w)C_z(z,w)+w(\frac{cd}{a}w-a+d)C_w(z,w)-\frac{adz^2}{2(a+d)^2}(a+dw)C(z,w)=0, \qquad (32)$$

with initial condition $C(0,w)=1$.

This differential equation has a rather simple explicit solution, which is given in the following lemma.

Lemma 5. *The bivariate generating function* $C(z,w)=\sum_{s\ge 0}\sum_{\ell\ge 0}\tilde{\gamma}_{s,\ell}\frac{z^s}{s!}w^\ell$ *of the sequence* $\tilde{\gamma}_{s,k}$ *is given by*

$$C(z,w)=\exp\Big(\frac{z^2}{2}\big(\gamma_{2,2}+w\gamma_{2,1}+w^2\gamma_{2,0}\big)\Big), \qquad (33)$$

where the values $\gamma_{2,2}$*,* $\gamma_{2,1}$ *and* $\gamma_{2,0}$ *are given as follows:*

$$\gamma_{2,2}=\frac{a^2d}{(a+d)^2(2a+d)},\ \ \gamma_{2,1}=\frac{ad^2(2a+2c+d)}{(a+d)^2(2a+d)(a+2d)}\ \ \gamma_{2,0}=\frac{cd^2(2a+2c+d)}{3(a+d)^2(2a+d)(a+2d)}.$$

Proof. We will solve the partial differential equation (32), where we apply the so-called method of characteristics, see, e.g., [21] for a description of this method.

We do this by studying first the corresponding reduced partial differential equation

$$z(a+\frac{d}{2}-\frac{cd}{a}w)C_z(z,w)+w(\frac{cd}{a}w-a+d)C_w(z,w)=0. \qquad (34)$$

Let us thus consider the following system of ordinary differential equations, the so-called system of characteristic differential equations:

$$\dot{z}=z(a+\frac{d}{2}-\frac{cd}{a}w), \quad \dot{w}=w(\frac{cd}{a}w-a+d), \qquad (35)$$

where we regard here z and w as dependent variables of t, namely, $z=z(t)$, $w=w(t)$, and $\dot{z}=\frac{dz(t)}{dt}$, etc. We are searching now for first integrals of the system of characteristic differential equations (35), i.e., for functions $\xi(z,w)$, which are constant along any solution curve (a so called characteristic curve) of (35).

From the characteristic differential equations (35) we obtain the differential equation

$$z'(w)=\frac{dz}{dw}=\frac{z(w)(a+\frac{d}{2}-\frac{cd}{a}w)}{w(\frac{cd}{a}w-a+d)}, \qquad (36)$$

where we regard now $z=z(w)$ as a function dependent on w. Equation (36) can be solved easily and one obtains that the general solution of (36) is given as follows:

$$z=\frac{K}{w^{\frac{2a+d}{2(a-d)}}(cdw-a^2+da)^{\frac{3d}{2(d-a)}}},$$

with an arbitrary constant K. As a consequence we obtain the following first integral of the system of characteristic differential equations (35):

$$\xi(z,w)=zw^{\frac{2a+d}{2(a-d)}}(cdw-a^2+da)^{\frac{3d}{2(d-a)}}=K=\text{const.}$$

In order to study the inhomogeneous partial differential equation (32) we use the following transformation from (z,w)-coordinates to (ξ,η)-coordinates:

$$\xi=zw^{\frac{2a+d}{2(a-d)}}(cdw-a^2+da)^{\frac{3d}{2(d-a)}} \quad \text{and} \quad \eta=w,$$

or equivalently

$$z=z(\xi,\eta)=\xi\eta^{\frac{2a+d}{2(d-a)}}(cd\eta-a^2+da)^{\frac{3d}{2(a-d)}} \quad \text{and} \quad w=w(\xi,\eta)=\eta,$$

which leads to the following ordinary differential equation for $\tilde{C}(\xi,\eta):=C\big(z(\xi,\eta),w(\xi,\eta)\big)$:

$$\eta(\frac{cd}{a}\eta-a+d)\tilde{C}_\eta(\xi,\eta)-\frac{ad\xi^2\eta^{\frac{2a+d}{d-a}}(cd\eta-a^2+da)^{\frac{3d}{a-d}}}{2(a+d)^2}(a+d\eta)\tilde{C}(\xi,\eta)=0. \qquad (37)$$

By separating variables we obtain the general solution of (37), which leads, after applying the inverse transformation to (z,w)-coordinates, the general solution of (32), where $F(x)$ is an arbitrary differentiable function:

$$C(z,w) = F\big(zw^{\frac{2a+d}{2(a-d)}}(cdw - a^2 + da)^{\frac{3d}{2(d-a)}}\big) \cdot \exp\Big(\frac{z^2}{2}(\gamma_{2,2} + w\gamma_{2,1} + w^2\gamma_{2,0})\Big).$$

In order to characterize the unknown function $F(x)$, such that $C(z,w)$ also satisfies the initial condition, we use the fact that the function $C(z,w)$ is analytic around $z = w = 0$ for arbitrary choice of a and d (in particular for $a = d$). Moreover, according to the initial condition $C(0,w) = 1$, we have $F(0) = 1$. Hence, it must hold that $F(x) = 1$, which completes the proof of Lemma 5. □

When extracting coefficients from $C(z,w)$ we obtain via $\gamma_{s,\ell} = s![z^s w^{s-\ell}]C(z,w)$ the following explicit formula for the numbers $\gamma_{s,\ell}$:

Corollary 1. *For $0 \le \ell \le s$ and s even, the values $\gamma_{s,\ell}$ are given as follows:*

$$\gamma_{s,\ell} = \frac{s!}{2^{\frac{s}{2}}(\frac{s}{2})!} \sum_{\substack{s_1+s_2+s_3=\frac{s}{2}\\ s_2+2s_3=s-\ell}} \binom{\frac{s}{2}}{s_1, s_2, s_3} \gamma_{2,2}^{s_1}\gamma_{2,1}^{s_2}\gamma_{2,0}^{s_3},$$

with $\gamma_{2,2}$, $\gamma_{2,1}$ and $\gamma_{2,0}$ as given in Lemma 5. Moreover, for the instance $c = 0$, we have $\gamma_{s,\ell} = 0$, for $0 \le \ell < s/2$ and $s \ge 2$.

We summarize now our findings concerning the explicit structure of $\hat{e}_{n,m}^{[s]}$, i.e., the s-th centered moments of $A_{an,dm}$.

Proposition 4. *The s-th centered moments $\hat{e}_{n,m}^{[s]} = \mathbb{E}\Big(\big(A_{an,dm} - \mathbb{E}(A_{an,dm})\big)^s\Big)$ of the discrete area $A_{an,dm}$ under lattice paths associated with diminishing urn models with a ball replacement matrix defined in* (1) *are polynomials in n and m. These polynomials have the representation*

$$\hat{e}_{n,m}^{[0]} = 1, \qquad \hat{e}_{n,m}^{[s]} = \sum_{\ell=0}^{s} \sum_{j=1}^{\lfloor\frac{3s}{2}\rfloor-\ell} \hat{\vartheta}_{s,\ell,j} n^\ell m^j, \quad \text{for } s \ge 1,$$

with certain coefficients $\hat{\vartheta}_{s,\ell,j}$.

Furthermore, for s even, an explicit formula for the leading coefficients $\gamma_{s,\ell} = \hat{\vartheta}_{s,\ell,\frac{3s}{2}-\ell}$ of the leading terms of $\hat{e}_{n,m}^{[s]}$ is given in Corollary 1, and an explicit formula for the generating function $C(z,w) = \sum_{s\ge0}\sum_{\ell\ge0}\gamma_{s,s-\ell}\frac{z^s}{s!}w^{s-\ell}$ is given in Lemma 5.

5.5. **The proof of the normal limit law for the instance $c \neq 0$ and $m \to \infty$.** The Gaussian limit law, i.e., $\frac{A_{an,dm}-\mathbb{E}(A_{an,dm})}{\sqrt{\mathbb{V}(A_{an,dm})}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$, follows for the instance $c \neq 0$ and $m \to \infty$ immediately from the following lemma and applying the Theorem of Fréchet and Shohat.

Lemma 6. *For the instance $c \neq 0$, with $m \to \infty$ and arbitrary $n = n(m)$, the centered and normalized moments of $A_{an,dm}$ satisfy the following asymptotic expansions:*

$$\mathbb{E}\left(\Big(\frac{A_{an,dm} - \mathbb{E}(A_{an,dm})}{\sqrt{\mathbb{V}(A_{an,dm})}}\Big)^s\right) = \frac{\mathbb{E}(\hat{A}_{an,dm}^s)}{(\mathbb{V}(\hat{A}_{an,dm}))^{\frac{s}{2}}} = \frac{\hat{e}_{n,m}^{[s]}}{\big(\hat{e}_{n,m}^{[2]}\big)^{\frac{s}{2}}} = \begin{cases} \frac{s^2}{2^{\frac{s}{2}}(\frac{s}{2})!}\big(1 + \mathcal{O}(\frac{1}{m})\big), & \text{for } s \text{ even},\\ \mathcal{O}\big(\frac{1}{\sqrt{m}}\big), & \text{for } s \text{ odd}.\end{cases}$$

Proof. We use the explicit structure of the s-th centered moments $\hat{e}_{n,m}^{[s]}$ of $A_{an,dm}$ given in Proposition 4 and consider, for s even, the polynomial $f_s(m,n)$ consisting of the leading terms of $\hat{e}_{n,m}^{[s]}$:

$$f_s(m,n) := \sum_{\ell=0}^{s} \gamma_{s,\ell} n^\ell m^{\frac{3s}{2}-\ell}. \tag{38}$$

Note that

$$f_2(m,n) = \gamma_{2,2}mn^2 + \gamma_{2,1}m^2n + \gamma_{2,0}m^3,$$

with values $\gamma_{2,2}$, $\gamma_{2,1}$ and $\gamma_{2,0}$ given in Lemma 5.

The functions $f_s(m,n)$ can be obtained easily from the generating function $C(z,w)=\sum_{s\geq 0}\sum_{\ell\geq 0}\gamma_{s,s-\ell}\frac{z^s}{s!}w^{s-\ell}$ via

$$f_s(m,n)=\sum_{\ell=0}^{s}\gamma_{s,\ell}n^\ell m^{\frac{3s}{2}-\ell}=s!n^s m^{\frac{s}{2}}[z^s]C\big(z,\frac{m}{n}\big).$$

Using the exact formula for $C(z,w)$ as given in Lemma 5 we further obtain the following simple relation between $f_s(m,n)$ and $f_2(m,n)$:

$$\begin{aligned} f_s(m,n)&=s!n^s m^{\frac{s}{2}}[z^s]\exp\Big(\frac{z^2}{2}(\gamma_{2,2}+\frac{m}{n}\gamma_{2,1}+\frac{m^2}{n^2}\gamma_{2,0})\Big)\\ &=\frac{s^2}{2^{\frac{s}{2}}(\frac{s}{2})!}(\gamma_{2,2}mn^2+\gamma_{2,1}m^2n+\gamma_{2,0}m^3)^{\frac{s}{2}}=\frac{s^2}{2^{\frac{s}{2}}(\frac{s}{2})!}(f_2(m,n))^{\frac{s}{2}}. \end{aligned} \tag{39}$$

By simple asymptotic considerations we obtain, for the instance $c\neq 0$ and s even, the following asymptotic expansion:

$$\frac{\hat{e}^{[s]}_{n,m}}{\big(\hat{e}^{[2]}_{n,m}\big)^{\frac{s}{2}}}=\frac{f_s(m,n)(1+\mathcal{O}(\frac{1}{m}))}{(f_2(m,n)(1+\mathcal{O}(\frac{1}{m})))^{\frac{s}{2}}}=\frac{f_s(m,n)}{(f_2(m,n))^{\frac{s}{2}}}\big(1+\mathcal{O}(\frac{1}{m})\big)=\frac{s^2}{2^{\frac{s}{2}}(\frac{s}{2})!}\big(1+\mathcal{O}(\frac{1}{m})\big).$$

Moreover, for the instance $c\neq 0$ and s odd, we have due to Proposition 4 the bound

$$\hat{e}^{[s]}_{n,m}=\mathcal{O}\big(\sum_{\ell=0}^{s}n^\ell m^{\frac{3s-1}{2}-\ell}\big),$$

which leads ot the following asymptotic expansion:

$$\begin{aligned} \frac{(\hat{e}^{[s]}_{n,m})^2}{\big(\hat{e}^{[2]}_{n,m}\big)^s}&=\frac{(\mathcal{O}(\sum_{\ell=0}^{s}n^\ell m^{\frac{3s-1}{2}-\ell}))^2}{(f_2(m,n))^s(1+\mathcal{O}(\frac{1}{m}))}=\frac{\mathcal{O}(\sum_{\ell=0}^{2s}n^\ell m^{3s-1-\ell})}{(f_2(m,n))^s(1+\mathcal{O}(\frac{1}{m}))}=\mathcal{O}\Big(\frac{1}{m}\frac{\sum_{\ell=0}^{2s}n^\ell m^{3s-\ell}}{m^{2s}n^s(\frac{n}{m}+\frac{m}{n}+1)^s}\Big)\\ &=\mathcal{O}\Big(\frac{1}{m}\frac{\sum_{\ell=0}^{2s}n^\ell m^{3s-\ell}}{m^{2s}n^s\sum_{\ell=-s}^{s}\frac{m^\ell}{n^\ell}}\Big)=\mathcal{O}\Big(\frac{1}{m}\frac{\sum_{\ell=0}^{2s}n^\ell m^{3s-\ell}}{\sum_{\ell=0}^{2s}n^\ell m^{3s-\ell}}\Big)=\mathcal{O}(\frac{1}{m}). \end{aligned}$$

This proves Lemma 6. □

5.6. **The proof of the normal limit law for the instance case** $c=0$ **and** $m,n\to\infty$**.** For the instance $c=0$ we have to proceed slightly different in order to prove the normal limit law, i.e., $\frac{A_{an,dm}-\mathbb{E}(A_{an,dm})}{\sqrt{\mathbb{V}(A_{an,dm})}}\xrightarrow{\mathcal{L}}\mathcal{N}(0,1)$, since due to Corollary 1 many leading terms $\gamma_{s,\ell}$ vanish. Therefore we will distinguish between the cases $m\leq n$ and $m>n$. For $m,n\to\infty$ and $m\leq n$ we can proceed analogeous to the instance $c\neq 0$ and show asymptotic equivalents of the moments $\hat{e}^{[s]}_{n,m}$, which are stated in Lemma 7. Applying the Theorem of Fréchet and Shohat shows the Gaussian limit law for this instance. To prove the normal limit law also for $m,n\to\infty$ and $m>n$ we use the symmetry relation described in Remark 4.

Lemma 7. *For the case* $c=0$ *and* $m,n\to\infty$*, with* $m\leq n$*, the centered and normalized moments of* $A_{an,dm}$ *satisfy the following asymptotic expansions:*

$$\mathbb{E}\left(\Big(\frac{A_{an,dm}-\mathbb{E}(A_{an,dm})}{\sqrt{\mathbb{V}(A_{an,dm})}}\Big)^s\right)=\frac{\mathbb{E}(\hat{A}^s_{an,dm})}{(\mathbb{V}(\hat{A}_{an,dm}))^{\frac{s}{2}}}=\frac{\hat{e}^{[s]}_{n,m}}{\big(\hat{e}^{[2]}_{n,m}\big)^{\frac{s}{2}}}=\begin{cases}\frac{s^2}{2^{\frac{s}{2}}(\frac{s}{2})!}\big(1+\mathcal{O}(\frac{1}{m})\big), & \text{for } s \text{ even},\\ \mathcal{O}\big(\frac{1}{\sqrt{m}}\big), & \text{for } s \text{ odd}.\end{cases}$$

Proof. Again we use the explicit structure of the s-th centered moments $\hat{e}^{[s]}_{n,m}$ of $A_{an,dm}$ given in Proposition 4 and denote, for s even, by $f_s(m,n)$ the polynomial $f_s(m,n)=\sum_{\ell=0}^{s}\gamma_{s,\ell}n^\ell m^{\frac{3s}{2}-\ell}$ consisting of the leading terms of $\hat{e}^{[s]}_{n,m}$. When we assume that $m\leq n$ we easily obtain the following asymptotic expansions:

$$\hat{e}^{[s]}_{n,m}=\sum_{\ell=0}^{s}\sum_{j=1}^{\frac{3s}{2}-\ell}\hat{\vartheta}_{s,\ell,j}m^j n^\ell=\sum_{\ell=0}^{s}(\gamma_{s,\ell}m^{\frac{3s}{2}-\ell}+\mathcal{O}(m^{\frac{3s}{2}-\ell-1}))n^\ell=f_s(m,n)\big(1+\mathcal{O}(\frac{1}{m})\big),$$

for $s\geq 2$ even,

$$\hat{e}_{n,m}^{[s]} = \sum_{\ell=0}^{s} \sum_{j=1}^{\frac{3s-1}{2}-\ell} \hat{\vartheta}_{s,\ell,j} m^j n^\ell = \sum_{\ell=0}^{s} \mathcal{O}\big(m^{\frac{3s-1}{2}-\ell} n^\ell\big) = \mathcal{O}\big(m^{\frac{3s}{2}-s} n^s \frac{1}{\sqrt{m}}\big), \quad \text{for } s \text{ odd.}$$

Due to equation (39), which relates $f_s(m,n)$ with $f_2(m,n)$, we obtain thus for the instance s even the following asymptotic expansion:

$$\frac{\hat{e}_{n,m}^{[s]}}{\big(\hat{e}_{n,m}^{[2]}\big)^{\frac{s}{2}}} = \frac{f_s(m,n)}{(f_2(m,n))^{\frac{s}{2}}}\big(1+\mathcal{O}(\frac{1}{m})\big) = \frac{s^2}{2^{\frac{s}{2}}(\frac{s}{2})!}\big(1+\mathcal{O}(\frac{1}{m})\big).$$

On the other hand for the instance s odd we obtain the asymptotic expansion

$$\frac{(\hat{e}_{n,m}^{[s]})^2}{\big(\hat{e}_{n,m}^{[2]}\big)^s} = \frac{\mathcal{O}\Big(m^{3s-2s} n^{2s} \frac{1}{m}\Big)}{(f_2(m,n))^s} = \frac{\mathcal{O}\Big(m^{s} n^{2s} \frac{1}{m}\Big)}{\mathcal{O}(m^s n^{2s})} = \mathcal{O}(\frac{1}{m}),$$

which finishes the proof of Lemma 7. □

Now we turn to the case $m, n \to \infty$, with $m > n$. Here we use the symmetry relation described in Remark 4 and consider the random variable $F_{dm,an}$, which counts the discrete area under sample paths, starting at state (an, dm) and ending at an absorbing state, associated to urns with a ball replacement matrix $\left(\begin{smallmatrix}-d & 0\\ 0 & -a\end{smallmatrix}\right)$. Due to equation (17) the random variables $A_{an,dm}$ and $F_{dm,an}$ are related via $A_{an,dm} \stackrel{\mathcal{L}}{=} mn - F_{dm,an}$. Since $m > n$ (note that the rôle of m and n has been changed), we can apply Lemma 7 to $F_{dm,an}$ and obtain, for $m, n \to \infty$, a Gaussian limit law:

$$\frac{F_{dm,an} - \mathbb{E}(F_{dm,an})}{\sqrt{\mathbb{V}(F_{dm,an})}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1).$$

Due to the distributional equality

$$\frac{F_{dm,an} - \mathbb{E}(F_{dm,an})}{\sqrt{\mathbb{V}(F_{dm,an})}} \stackrel{\mathcal{L}}{=} -\frac{A_{an,dm} - \mathbb{E}(A_{an,dm})}{\sqrt{\mathbb{V}(A_{an,dm})}},$$

which follows from (17), we obtain further

$$\frac{A_{an,dm} - \mathbb{E}(A_{an,dm})}{\sqrt{\mathbb{V}(A_{an,dm})}} \xrightarrow{\mathcal{L}} -\mathcal{N}(0,1) \stackrel{\mathcal{L}}{=} \mathcal{N}(0,1),$$

which proves the Gaussian limit law also in the remaining case $m, n \to \infty$, with $m > n$.

Outlook and Conclusion

We have studied the area under lattice paths associated with a certain class of diminishing urn models. Our study aims at combining the fields of lattice path enumeration and (diminishing) Pólya-Eggenberger urn models. Best to the authors knowledge such questions have not been considered before in the literature. Several extensions of this approach, e.g., to certain classes of higher dimensional urn models, seems to be possible and are certainly also of interest. Moreover, it seems interesting to study the distribution of the area under lattice paths associated to urn models with different ball replacement matrices, such as the famous OK Corral urn model with a ball replacement matrix given by $M = \left(\begin{smallmatrix}0 & -1\\ -1 & 0\end{smallmatrix}\right)$.

References

[1] C. Banderier and P. Flajolet, Basic analytic combinatorics of directed lattice paths, *Theoretical Computer Science* 281, 37-80, 2002.

[2] C. Banderier and B. Gittenberger, Analytic combinatorics of lattice paths: enumeration and asymptotics for the average area, in: *Proceedings of the Fourth Colloquium of Mathematics and Computer Science*, *Discrete Mathematics and Theoretical Computer Science*, AG, 345–355, 2006.

[3] C. A. C. Brennan and H. Prodinger, The pills problem revisited, *Quaestiones Mathematicae* 26, 427–439, 2003.

[4] P. Flajolet, P. Dumas, and V. Puyhaubert, Some exactly solvable models of urn process theory, in: *Proceedings of Fourth Colloquium on Mathematics and Computer Science*, *Discrete Mathematics and Computer Science*, AG, 59–118, 2006.

[5] P. Flajolet, J. Gabarró, H. Pekari, Analytic Urns, *Annals of Probability* 33, 1200–1233, 2005.

[6] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics*, Addison-Wesley, Boston, 1994.

[7] D. Greene and D. E. Knuth, *Mathematics for the Analysis of Algorithms*, Birkhäuser, Boston, 1982.

[8] T. Hesterberg et al., Problems and solutions, E3429, *American Mathematical Monthly* 99, p. 684, 1992.

[9] H. K. Hwang, M. Kuba and A. Panholzer, Analysis of some exactly solvable diminishing urn models, in: *Proceedings of the 19th International Conference on Formal Power Series and Algebraic Combinatorics*, Nankai University, Tianjin, 2007. available at http://www.fpsac.cn/PDF-Proceedings/Posters/43.pdf

[10] S. Janson, Functional limit theorems for multitype branching processes and generalized Pólya urns, *Stochastic processes and applications* 110, 177–245, 2004.

[11] S. Janson, Limit theorems for triangular urn schemes, *Probability Theory and Related Fields* 134, 417–452, 2005.

[12] N. L. Johnson and S. Kotz, *Urn models and their application. An approach to modern discrete probability theory*, John Wiley, New York, 1977.

[13] D. E. Knuth and J. McCarthy, Problem E3429: Big pills and little pills, *American Mathematical Monthly* 98, p. 264, 1991.

[14] S. Kotz and N. Balakrishnan, Advances in urn models during the past two decades, in: *Advances in combinatorial methods and applications to probability and statistics*, 203–257, Statistics for Industry and Technology, Birkhäuser, Boston, 1997.

[15] M. Kuba and A. Panholzer, Limiting distributions for a class of diminishing urn models, submitted, 2007.

[16] M. Loève, *Probability Theory I*, 4th Edition, Springer, New York, 1977.

[17] G. Louchard. Kacs formula, Levys local time and Brownian excursion, *Journal of Applied Probability* 21, 479-499, 1984.

[18] H. Mahmoud, Urn models and connections to random trees: a review, *Journal of the Iranian Mathematical Society* 2, 53-114, 2003.

[19] N. Pouyanne, Classification of large Pólya-Eggenberger urns with regard to their asymptotics, in: *International Conference on the Analysis of Algorithms*, *Discrete Mathematics and Theoretical Computer Science*, AD, 275-285, 2005.

[20] N. Pouyanne, An algebraic approach to Pólya processes, *Annales de l'Institut Henri Poincaré* 44, 293–323, 2008.

[21] M. E. Taylor, *Partial differential equations. Basic theory. Texts in Applied Mathematics, 23*, Springer, New York, 1996.

MARKUS KUBA, INSTITUT FÜR DISKRETE MATHEMATIK UND GEOMETRIE, TECHNISCHE UNIVERSITÄT WIEN, WIEDNER HAUPTSTR. 8-10/104, 1040 WIEN, AUSTRIA

E-mail address: markus.kuba+e104@tuwien.ac.at

ALOIS PANHOLZER, INSTITUT FÜR DISKRETE MATHEMATIK UND GEOMETRIE, TECHNISCHE UNIVERSITÄT WIEN, WIEDNER HAUPTSTR. 8-10/104, 1040 WIEN, AUSTRIA

E-mail address: Alois.Panholzer@tuwien.ac.at