

Correction on: "Population genetics models with skewed fertilities: a forward and backward analysis".

Thierry Huillet, Martin Moehle

▶ To cite this version:

Thierry Huillet, Martin Moehle. Correction on: "Population genetics models with skewed fertilities: a forward and backward analysis". Stochastic Models 28, Issue3, 527-532, 2012., 2012, 28 (3), pp.527-532. 10.1080/15326349.2012.700799. hal-00646215

HAL Id: hal-00646215 https://hal.science/hal-00646215

Submitted on 29 Nov 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ERRATUM ON 'POPULATION GENETICS MODELS WITH SKEWED FERTILITIES: A FORWARD AND BACKWARD ANALYSIS'

Thierry Huillet¹ and Martin $M\ddot{o}hle^2$

November 18, 2011

Abstract

The article 'Population genetics models with skewed fertilities: a forward and backward analysis', *Stochastic Models* **27**, 521–554 (2011) contains on top of page 536 a formula for the joint factorial moments of the offspring numbers μ_1, \ldots, μ_N , which is wrong in that generality. It is clarified for which compound Poisson models this formula holds true. It turns out that the only compound Poisson models for which this formula holds true are skewed generalized Wright–Fisher models and skewed generalized Dirichlet models. An erratum is provided correcting the results in Section 4 of the mentioned article from Proposition 4.2 on. The main conclusion (Theorem 3.2) that many symmetric compound Poisson population models are in the domain of attraction of the Kingman coalescent, remains valid, however, its proof turns out to be more involved. A key analytic tool in the proof is the saddle point method. In particular, the correct time-scaling (effective population size) is provided.

Keywords: Bell polynomials; compound Poisson model; Dirichlet model; Kingman coalescent; saddle point method; Wright–Fisher model

2010 Mathematics Subject Classification: Primary 60J10; Secondary 60K35, 92D10, 92D25

1 Introduction

Let us briefly recall the definitions of conditional branching process models and compound Poisson models. Conditional branching process models are population models with fixed population size $N \in \mathbb{N} := \{1, 2, ...\}$ and non-overlapping generations. They are defined in terms of a sequence $(\xi_n)_{n \in \mathbb{N}}$ of independent non-negative integer-valued random variables satisfying $\mathbb{P}(\xi_1 + \cdots + \xi_N = N) > 0$. If, for $i \in \{1, ..., N\}$, $\mu_{N,i}$ denotes the number of offspring of the *i*th individual alive in some fixed generation, then the random variables $\mu_{N,1}, \ldots, \mu_{N,N}$ have (by definition) joint distribution

$$\mathbb{P}(\mu_{N,1} = j_1, \dots, \mu_{N,N} = j_N) = \frac{\mathbb{P}(\xi_1 = j_1) \cdots \mathbb{P}(\xi_N = j_N)}{\mathbb{P}(\xi_1 + \dots + \xi_N = N)}$$

 $j_1, \ldots, j_N \in \mathbb{N}_0 := \{0, 1, \ldots\}$ with $j_1 + \cdots + j_N = N$. For convenience we will often drop the index N and simply write μ_i instead of $\mu_{N,i}$. For more information on conditional branching process models we refer the reader to [7, Section 3].

¹Laboratoire de Physique Théorique et Modélisation, CNRS-UMR 8089 et Université de Cergy-Pontoise, 2 Avenue Adolphe Chauvin, 95302 Cergy-Pontoise, France, E-mail: thierry.huillet@u-cergy.fr

²Mathematisches Institut, Eberhard Karls Universität Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany, E-mail: martin.moehle@uni-tuebingen.de

We now turn to the definition of compound Poisson population models. Let $\theta_1, \theta_2, \ldots$ be strictly positive real numbers and let $\phi(z) = \sum_{m=1}^{\infty} \phi_m z^m / m!$, |z| < r, be a power series with radius $r \in (0, \infty]$ of convergence and with nonnegative coefficients $\phi_m \ge 0$, $m \in \mathbb{N}$. It is also assumed that $\phi_1 > 0$. Compound Poisson models are particular conditional branching process models where each random variable ξ_n has probability generating function (pgf)

$$f_n(x) := \mathbb{E}(x^{\xi_n}) = \exp(-\theta_n(\phi(z) - \phi(zx))), \quad |x| \le 1.$$
 (1)

In (1), z is viewed as a fixed parameter satisfying |z| < r. However, in order to analyze compound Poisson models it is useful to view z as a variable and to introduce, for $\theta \in [0, \infty)$, the Taylor expansion

$$\exp(\theta\phi(z)) = \sum_{k=0}^{\infty} \frac{\sigma_k(\theta)}{k!} z^k, \qquad |z| < r$$

The coefficients $\sigma_k(\theta)$ depend on $(\phi_m)_{m\in\mathbb{N}}$ and they satisfy the recursion $\sigma_0(\theta) = 1$ and

$$\sigma_{k+1}(\theta) = \theta \sum_{l=0}^{k} \binom{k}{l} \phi_{k-l+1} \sigma_l(\theta), \qquad k \in \mathbb{N}_0, \theta \in [0, \infty).$$
(2)

The coefficients $\sigma_k(\theta)$ are mainly introduced, since, by (1), the distribution of ξ_n , $n \in \mathbb{N}$, satisfies

$$\mathbb{P}(\xi_n = k) = \sigma_k(\theta_n) \frac{z^k}{k!} \exp(-\theta_n \phi(z)), \qquad k \in \mathbb{N}_0.$$
(3)

From $\phi_1 > 0$ it follows that $\sigma_k(\theta)$ is a polynomial in θ of degree k. In the literature (see, for example, [1] or [3]) the $\sigma_k(\theta)$ are called the exponential polynomials. We have $\sigma_1(\theta) = \theta\phi_1$, $\sigma_2(\theta) = \theta\phi_2 + \theta^2\phi_1^2$, $\sigma_3(\theta) = \theta\phi_3 + 3\theta^2\phi_1\phi_2 + \theta^3\phi_1^3$, $\sigma_4(\theta) = \theta\phi_4 + \theta^2(4\phi_1\phi_3 + 3\phi_2^2) + 6\theta^3\phi_1^2\phi_2 + \theta^4\phi_1^4$, and so on. The coefficients $B_{kl}(\phi_1, \phi_2, \ldots), k \in \mathbb{N}_0, l \in \{0, \ldots, k\}$, of the polynomials $\sigma_k(\theta) = \sum_{l=0}^k B_{kl}(\phi_1, \phi_2, \ldots) \theta^l, k \in \mathbb{N}_0$, are called the Bell coefficients. It is readily checked that ξ_n has descending factorial moments

$$\mathbb{E}((\xi_n)_k) = f_n^{(k)}(1) = z^k \sum_{l=0}^k B_{kl}(\phi'(z), \phi''(z), \ldots) \theta_n^l, \qquad n \in \mathbb{N}, k \in \mathbb{N}_0,$$

i.e. $\mathbb{E}(\xi_n) = \theta_n z \phi'(z)$, $\mathbb{E}((\xi_n)_2) = \theta_n z^2 \phi''(z) + \theta_n^2 z^2 (\phi'(z))^2$ and so on. The descending factorial moments therefore satisfy the recursion

$$\mathbb{E}((\xi_n)_{k+1}) = \theta_n \sum_{l=0}^k \binom{k}{l} z^{k-l+1} \phi^{(k-l+1)}(z) \mathbb{E}((\xi_n)_l), \qquad n \in \mathbb{N}, k \in \mathbb{N}_0.$$

Throughout the article, for $x \in \mathbb{R}$ and $k \in \mathbb{N}_0$, the notations $(x)_k := x(x-1)\cdots(x-k+1)$ and $[x]_k := x(x+1)\cdots(x+k-1)$ are used for the descending and ascending factorials respectively, with the convention that $(x)_0 := [x]_0 := 1$.

2 Results

We derive the correct formula for the joint factorial moments of the offspring random variables μ_1, \ldots, μ_N for an arbitrary compound Poisson population model. It is known (see, for example, [7, p. 535]) that $\mu = (\mu_1, \ldots, \mu_N)$ has distribution

$$\mathbb{P}(\mu=j) = \frac{N!}{\sigma_N(\Theta_N)} \prod_{n=1}^N \frac{\sigma_{j_n}(\theta_n)}{j_n!}, \qquad j = (j_1, \dots, j_N) \in \Delta(N), \tag{4}$$

where $\Theta_N := \theta_1 + \cdots + \theta_N$ and $\Delta(N)$ denotes the discrete N-simplex consisting of all $j = (j_1, \ldots, j_N) \in \mathbb{N}_0^N$ satisfying $j_1 + \cdots + j_N = N$. Note that the distribution of μ is not necessarily exchangeable. It follows that μ has joint factorial moments

$$\mathbb{E}((\mu_1)_{k_1}\cdots(\mu_N)_{k_N}) = \sum_{\substack{j=(j_1,\dots,j_N)\in\Delta(N)\\ j=(j_1,\dots,j_N)\in\Delta(N)}} (j_1)_{k_1}\cdots(j_N)_{k_N}\mathbb{P}(\mu=j)$$

$$= \frac{N!}{\sigma_N(\Theta_N)}\sum_{\substack{j_1,\dots,j_N\in\mathbb{N}\\ j_1+\dots+j_N=N}} (j_1)_{k_1}\cdots(j_N)_{k_N}\frac{\sigma_{j_1}(\theta_1)\cdots\sigma_{j_N}(\theta_N)}{j_1!\cdots j_N!}$$

$$= \frac{N!}{\sigma_N(\Theta_N)}\sum_{\substack{j_1\geq k_1,\dots,j_N\geq k_N\\ j_1+\dots+j_N=N}} \frac{\sigma_{j_1}(\theta_1)\cdots\sigma_{j_N}(\theta_N)}{(j_1-k_1)!\cdots(j_N-k_N)!}, \quad (5)$$

 $N \in \mathbb{N}, k_1, \ldots, k_N \in \mathbb{N}_0, \theta_1, \ldots, \theta_N \in (0, \infty)$, which is not a simple expression. However, it is known that for some compound Poisson models, for example for the Wright–Fisher model, the alternative and simpler formula

$$\mathbb{E}((\mu_1)_{k_1}\cdots(\mu_N)_{k_N}) = (N)_k \frac{\sigma_{k_1}(\theta_1)\cdots\sigma_{k_N}(\theta_N)}{\sigma_k(\theta_1+\cdots+\theta_N)}$$
(6)

holds for all $k_1, \ldots, k_N \in \mathbb{N}_0$ and all $\theta_1, \ldots, \theta_N \in (0, \infty)$, where $k := k_1 + \cdots + k_N$. In the following it is clarified which compound Poisson models satisfy (6). Our results are based on the following basic but fundamental lemma, which essentially coincides with [9, Lemma 2.1], however, we prefer here to state it slightly different.

Lemma 2.1 A compound Poisson model (with given fixed power series ϕ) satisfies the factorial moment formula (6) for all population sizes $N \in \mathbb{N}$, all $k_1, \ldots, k_N \in \mathbb{N}_0$ and all $\theta_1, \ldots, \theta_N \in (0, \infty)$ if and only if the coefficients $\sigma_k(\theta)$, $k \in \mathbb{N}_0$, $\theta \in (0, \infty)$, satisfy the relation

$$\frac{\sigma_{k+1}(\theta)}{\sigma_k(\theta)} + \frac{\sigma_{k'+1}(\theta')}{\sigma_{k'}(\theta')} = \frac{\sigma_{k+k'+1}(\theta+\theta')}{\sigma_{k+k'}(\theta+\theta')}, \qquad k,k' \in \mathbb{N}_0, \theta, \theta' \in (0,\infty).$$
(7)

Proof. Suppose first that (7) holds. We essentially follow the proof of [9, Lemma 2.1]. By induction on $N \in \mathbb{N}$ it follows that

$$\sum_{j=1}^{N} \frac{\sigma_{k_j+1}(\theta_j)}{\sigma_{k_j}(\theta_j)} = \frac{\sigma_{k+1}(\theta_1 + \dots + \theta_N)}{\sigma_k(\theta_1 + \dots + \theta_N)}, \qquad k_1, \dots, k_N \in \mathbb{N}_0, \theta_1, \dots, \theta_N \in (0, \infty).$$
(8)

Let us now verify (6) by backward induction on $k = k_1 + \cdots + k_N$. For k > N both sides of (6) are equal to zero. For k = N we have

$$\mathbb{E}((\mu_1)_{k_1}\cdots(\mu_N)_{k_N}) = k_1!\cdots k_N!\mathbb{P}(\mu_1 = k_1,\ldots,\mu_N = k_N) = N!\frac{\sigma_{k_1}(\theta_1)\cdots\sigma_{k_N}(\theta_N)}{\sigma_k(\theta_1+\cdots+\theta_N)},$$

which is (6) for k = N. The backward induction step from k + 1 to $k \ (\in \{0, ..., N - 1\})$ works as follows. From $\mu_1 + \cdots + \mu_N = N$ and by induction it follows that

$$(N-k)\mathbb{E}((\mu_{1})_{k_{1}}\cdots(\mu_{N})_{k_{N}}) = \mathbb{E}((\mu_{1})_{k_{1}}\cdots(\mu_{N})_{k_{N}}\sum_{j=1}^{N}(\mu_{j}-k_{j}))$$

$$= \sum_{j=1}^{N}\mathbb{E}((\mu_{1})_{k_{1}}\cdots(\mu_{j})_{k_{j}+1}\cdots(\mu_{N})_{k_{N}}) = \sum_{j=1}^{N}(N)_{k+1}\frac{\sigma_{k_{1}}(\theta_{1})\cdots\sigma_{k_{j}+1}(\theta_{j})\cdots\sigma_{k_{N}}(\theta_{N})}{\sigma_{k+1}(\theta_{1}+\cdots+\theta_{N})}$$

$$= (N)_{k+1}\frac{\sigma_{k_{1}}(\theta_{1})\cdots\sigma_{k_{N}}(\theta_{N})}{\sigma_{k+1}(\theta_{1}+\cdots+\theta_{N})}\sum_{j=1}^{N}\frac{\sigma_{k_{j}+1}(\theta_{j})}{\sigma_{k_{j}}(\theta_{j})} = (N)_{k+1}\frac{\sigma_{k_{1}}(\theta_{1})\cdots\sigma_{k_{N}}(\theta_{N})}{\sigma_{k}(\theta_{1}+\cdots+\theta_{N})}, \quad (9)$$

where the last equality holds by (8). Division of (9) by N - k shows that (6) holds for k which completes the backward induction.

Conversely, if (6) holds for all $N \in \mathbb{N}, k_1, \ldots, k_N \in \mathbb{N}_0$ and all $\theta_1, \ldots, \theta_N \in (0, \infty)$, then

$$(N)_{k+1} \frac{\sigma_{k_1}(\theta_1) \cdots \sigma_{k_N}(\theta_N)}{\sigma_k(\theta_1 + \dots + \theta_N)} = (N-k)\mathbb{E}((\mu_1)_{k_1} \cdots (\mu_N)_{k_N})$$
$$= (N)_{k+1} \frac{\sigma_{k_1}(\theta_1) \cdots \sigma_{k_N}(\theta_N)}{\sigma_{k+1}(\theta_1 + \dots + \theta_N)} \sum_{j=1}^N \frac{\sigma_{k_j+1}(\theta_j)}{\sigma_{k_j}(\theta_j)},$$

where the last equation is obtained by doing the same calculations as above. Thus, (8) holds. Choosing N = 2 in (8), it follows that (7) holds which completes the proof.

Relation (7) puts strong constrains on the coefficients $\sigma_k(\theta)$, $k \in \mathbb{N}_0$, $\theta \in (0, \infty)$, and, consequently, on the power series ϕ of the compound Poisson model. Lemma 2.1 of [9] and the remarks thereafter show that (7) holds if and only if $\phi(z) = \phi_1 z$ or

$$\phi(z) = -\frac{\phi_1^2}{\phi_2} \log\left(1 - \frac{\phi_2}{\phi_1}z\right), \quad |z| < \frac{\phi_1}{\phi_2}.$$
(10)

The only compound Poisson models satisfying the factorial moment formula (6) for all $N \in \mathbb{N}$, all $k_1, \ldots, k_N \in \mathbb{N}_0$, and all $\theta_1, \ldots, \theta_N \in (0, \infty)$ are therefore generalized Wright–Fisher models (with ϕ of the form $\phi(z) = \phi_1 z$ for some constant $\phi_1 \in (0, \infty)$) and generalized Dirichlet models with ϕ of the form (10) for some constants $\phi_1, \phi_2 \in (0, \infty)$. For all other compound Poisson models the formula (6) does not hold for all $N \in \mathbb{N}$, all $k_1, \ldots, k_N \in \mathbb{N}_0$ and all $\theta_1, \ldots, \theta_N \in (0, \infty)$, a typical counter example being the compound Poisson model with power series ϕ of the form $\phi(z) = e^z - 1$, or, equivalently, $\phi_m = 1$ for all $m \in \mathbb{N}$.

3 Corrections and further results

We comment which parts of [7] need correction. The results in [7, Sections 5 and 6] are correct, since the models studied there satisfy (6). The statements in [7, Section 7] are as well correct, since the Kimura model does not belong to the compound Poisson class. Thus, only [7, Section 4] from [7, Proposition 4.2] on needs correction. From the results of the previous Section 2 it follows that [7, Proposition 4.2] and [7, Theorem 4.3] and the remarks thereafter are valid for generalized Wright–Fisher models and generalized Dirichlet models with power series ϕ as given at the end of Section 2. However, for general compound Poisson models, Proposition 4.2 and Theorem 4.3 of [7] and the remark thereafter are wrong. Proposition 4.2 of [7] should be replaced by the following proposition.

Proposition 3.1 If, for each $n \in \mathbb{N}$, the random variable ξ_n has a pgf of the form [7, Eq. (13)], then the backward process \hat{X} of the associated skewed conditional branching process model has transition probabilities

$$\widehat{P}_{i,j} = \frac{1}{\binom{N}{i}} \sum_{1 \le n_1 < \dots < n_j \le N} \sum_{\substack{k_1, \dots, k_j \in \mathbb{N} \\ k:=k_1 + \dots + k_j \le N}} \frac{(N)_k}{k_1! \cdots k_j!} \cdot \frac{(\prod_{i=1}^j \sigma_{k_i}(\theta_{n_i})) \sigma_{N-k}(\Theta_N - \sum_{i=1}^j \theta_{n_i})}{\sigma_N(\Theta_N)} \sum_{\substack{l_1, \dots, l_j \in \mathbb{N} \\ l_1 + \dots + l_j = i}} \binom{k_1}{l_1} \cdots \binom{k_j}{l_j}, \quad i, j \in S, (11)$$

with the convention that $\widehat{P}_{i,0} = \delta_{i0}$, $i \in S$. Here $\Theta_N := \theta_1 + \cdots + \theta_N$, $S := \{0, \ldots, N\}$, and the coefficients $\sigma_k(\theta)$, $k \in \mathbb{N}_0$, $\theta \in [0, \infty)$, are recursively defined via (2). In particular,

$$\widehat{P}_{i,1} = \sum_{n=1}^{N} \sum_{k=i}^{N} \binom{N-i}{k-i} \frac{\sigma_k(\theta_n) \sigma_{N-k}(\Theta_N - \theta_n)}{\sigma_N(\Theta_N)}, \quad i \in \{1, \dots, N\}.$$
(12)

Proof. For $j \in \{1, \ldots, N\}$, pairwise distinct $n_1, \ldots, n_j \in \{1, \ldots, N\}$, and $k_1, \ldots, k_j \in \mathbb{N}_0$ with $k := k_1 + \cdots + k_j \leq N$ we have

$$\mathbb{P}(\mu_{n_1} = k_1, \dots, \mu_{n_j} = k_j) = \frac{\left(\prod_{i=1}^j \mathbb{P}(\xi_{n_i} = k_i)\right) \mathbb{P}(\sum_{m \in [N] \setminus \{n_1, \dots, n_j\}} \xi_m = N - k)}{\mathbb{P}(\xi_1 + \dots + \xi_N = N)}$$
$$= \frac{(N)_k}{k_1! \cdots k_j!} \frac{\sigma_{k_1}(\theta_{n_1}) \cdots \sigma_{k_j}(\theta_{n_j}) \sigma_{N-k}(\Theta_N - \sum_{i=1}^j \theta_{n_i})}{\sigma_N(\Theta_N)}$$

and, therefore, for $l_1, \ldots, l_j \in \mathbb{N}_0$,

$$\mathbb{E}\left(\prod_{i=1}^{j} \binom{\mu_{n_i}}{l_i}\right) = \sum_{k_1,\dots,k_j} \left(\prod_{i=1}^{j} \binom{k_i}{l_i}\right) \frac{(N)_k}{k_1!\cdots k_j!} \frac{\sigma_{k_1}(\theta_{n_1})\cdots\sigma_{k_j}(\theta_{n_j})\sigma_{N-k}(\Theta_N - \sum_{i=1}^{j} \theta_{n_i})}{\sigma_N(\Theta_N)},$$

where the sum \sum_{k_1,\ldots,k_j} extends over all $k_1,\ldots,k_j \in \mathbb{N}_0$ satisfying $k := k_1 + \cdots + k_j \leq N$. Thus (11) follows from [7, Eq. (4)]. For j = 1, (11) reduces to (12). Alternatively, (12) follows as well via

$$\widehat{P}_{i,1} = \sum_{n=1}^{N} \frac{\mathbb{E}((\mu_n)_i)}{(N)_i} = \sum_{n=1}^{N} \sum_{k=i}^{N} \frac{(k)_i}{(N)_i} \mathbb{P}(\mu_n = k)$$

$$= \sum_{n=1}^{N} \sum_{k=i}^{N} \frac{(k)_i}{(N)_i} {N \choose k} \frac{\sigma_k(\theta_n) \sigma_{N-k}(\Theta_N - \theta_n)}{\sigma_N(\Theta_N)}$$

$$= \sum_{n=1}^{N} \sum_{k=i}^{N} {N-i \choose k-i} \frac{\sigma_k(\theta_n) \sigma_{N-k}(\Theta_N - \theta_n)}{\sigma_N(\Theta_N)}.$$

Remark. Proposition 3.1 in particular yields the coalescence probability

$$c_{N} := \widehat{P}_{2,1} = \sum_{n=1}^{N} \sum_{k=2}^{N} {\binom{N-2}{k-2}} \frac{\sigma_{k}(\theta_{n}) \, \sigma_{N-k}(\Theta_{N} - \theta_{n})}{\sigma_{N}(\Theta_{N})}, \qquad N \ge 2$$
(13)

and the probability

$$d_{N} := \widehat{P}_{3,1} = \sum_{n=1}^{N} \sum_{k=3}^{N} \binom{N-3}{k-3} \frac{\sigma_{k}(\theta_{n}) \sigma_{N-k}(\Theta_{N} - \theta_{n})}{\sigma_{N}(\Theta_{N})}, \qquad N \ge 3$$
(14)

that three individuals share a common parent. For the unbiased case, when all the parameters $\theta_n = \theta$ are equal to some constant $\theta \in (0, \infty)$, (13) and (14) reduce to

$$c_N = N \sum_{k=2}^{N} {\binom{N-2}{k-2}} \frac{\sigma_k(\theta) \, \sigma_{N-k}((N-1)\theta)}{\sigma_N(N\theta)}$$
(15)

and

$$d_N = N \sum_{k=3}^{N} {N-3 \choose k-3} \frac{\sigma_k(\theta) \,\sigma_{N-k}((N-1)\theta)}{\sigma_N(N\theta)}.$$
(16)

It is clear that Theorem 4.3 of [7] does not hold for arbitrary symmetric compound Poisson models. Theorem 4.3 of [7] should be replaced by the following Theorem 3.2, which clarifies that many symmetric compound Poisson models are in the domain of attraction of the Kingman coalescent.

Theorem 3.2 Fix $\theta \in (0, \infty)$ and suppose that the equation $\theta z \phi'(z) = 1$ has a real solution $z(\theta) \in (0, r)$, where $r \in (0, \infty]$ denotes the radius of convergence of ϕ . Then, $\mu_{N,1} \to X$ in distribution as $N \to \infty$, where X is a nonnegative integer valued random variable with distribution

$$\mathbb{P}(X=k) = \sigma_k(\theta) \frac{(z(\theta))^k}{k!} e^{-\theta\phi(z(\theta))}, \qquad k \in \mathbb{N}_0.$$
(17)

Moreover, the associated symmetric compound Poisson model is in the domain of attraction of the Kingman coalescent in the sense of [7, Definition 2.1 (a)]. The effective population size $N_e := 1/c_N$ satisfies $N_e \sim \varrho N$ as $N \to \infty$, where $\varrho := 1/\mathbb{E}((X)_2) = 1/(1 + \theta(z(\theta))^2 \phi''(z(\theta))) \in (0, 1].$ **Remarks.** 1. If $\phi'(r-) = \infty$, then for all $\theta \in (0, \infty)$ the equation $\theta z \phi'(z) = 1$ has a real solution $z(\theta) \in (0, r)$. If $\phi'(r-) < \infty$, then a real solution $z(\theta) \in (0, r)$ of the equation $\theta z \phi'(z) = 1$ exists if and only if $\theta r \phi'(r-) > 1$. A concrete model satisfying $\phi'(r-) < \infty$ is provided in Example 4.8 in Section 4. The solution $z(\theta) \in (0, r)$ (if it exists) is unique since the map $z \mapsto z \phi'(z)$ is strictly monotone increasing on (0, r).

2. The following proof even shows the convergence of all moments $\mathbb{E}(\mu_{N,1}^p) \to \mathbb{E}(X^p)$ as $N \to \infty$, p > 0. The distribution of the limiting variable X coincides with the distribution (3) of ξ_1 with the parameter z in (3) replaced by $z(\theta)$. Note that X has mean $\mathbb{E}(X) = \theta z(\theta) \phi'(z(\theta)) = 1$. Thus, conditioning ξ_1 on the event that $\xi_1 + \cdots + \xi_N = N$ and afterwards taking $N \to \infty$, has altogether the effect that we 'nearly' recover the distribution of ξ_1 . We only loose the information about the mean of ξ_1 .

3. Closed expressions for the solution $z = z(\theta)$ of the equation $u(z) := z\phi'(z) = 1/\theta$ seem to be not available in general. By the inversion formula of Lagrange (see, for example, [4, Section 3.8]), for $k \in \mathbb{N}$ and $x \in (0, \infty)$,

$$[x^{k}]u^{-1}(x) = \frac{[z^{k-1}](z/u(z))^{k}}{k} = \frac{[z^{k-1}](1/\phi'(z))^{k}}{k} = B_{k-1,k}(\psi_{0},\psi_{1},\ldots),$$

where $\psi_n := \psi^{(n)}(0), n \in \mathbb{N}_0$, with $\psi(z) := 1/\phi'(z)$. Choosing $x := 1/\theta$ and noting that $z(\theta) = u^{-1}(x)$ yields the formal expansion

$$z(\theta) = \sum_{k=1}^{\infty} \frac{[z^{k-1}](\psi(z))^k}{k} \, \theta^{-k} = \sum_{k=1}^{\infty} B_{k-1,k}(\psi_0, \psi_1, \ldots) \theta^{-k}.$$

Note however that, depending on ϕ and θ , this series does not necessarily need to converge, so we can only speak about a formal expansion here.

Proof. Fix $k, l \in \mathbb{N}_0$ and $\theta \in (0, \infty)$. Let us verify that

$$\frac{\sigma_{n-k}((n-l)\theta)}{(n-k)!} \sim \frac{(a(\theta))^n}{\sqrt{2\pi n}} b_{kl}(\theta), \qquad n \to \infty, \tag{18}$$

where

$$a(\theta) := \frac{e^{\theta\phi(z(\theta))}}{z(\theta)} \quad \text{and} \quad b_{kl}(\theta) := \frac{(z(\theta))^k e^{-l\theta\phi(z(\theta))}}{\sqrt{1 + \theta(z(\theta))^2 \phi''(z(\theta))}}.$$
(19)

We proceed similarly as in the proof of Theorem 2.1 of [3]. However, note that in [3], asymptotic expansions for $\sigma_n(\theta)$ are provided whereas we are essentially interested in asymptotic expansions of $\sigma_n(n\theta)$. By Cauchy's integral formula, $\sigma_n(\theta)/n! = (2\pi i)^{-1} \int_C z^{-(n+1)} e^{\theta\phi(z)} dz$, $n \in \mathbb{N}_0$, where C is some contour around the origin. Replacing n by n - k and θ by $(n - l)\theta$ it follows that

$$\begin{aligned} \frac{\sigma_{n-k}((n-l)\theta)}{(n-k)!} &= \frac{1}{2\pi i} \int_C \frac{e^{(n-l)\theta\phi(z)}}{z^{n-k+1}} \, dz \\ &= \frac{1}{2\pi i} \int_C z^{k-1} e^{-l\theta\phi(z)} e^{n(\theta\phi(z)-\log z)} \, dz = \frac{1}{2\pi i} \int_C h(z) e^{ng(z)} \, dz, \end{aligned}$$

where $h(z) := z^{k-1}e^{-l\theta\phi(z)}$ and $g(z) := \theta\phi(z) - \log z$. Note that $g'(z) = \theta\phi'(z) - 1/z$ and that $g''(z) = \theta\phi''(z) + 1/z^2$. In particular, g' has a single real zero at the point $z(\theta)$ solving

the equation $\theta z(\theta)\phi'(z(\theta)) = 1$. Note that g''(z) > 0 for all $z \in \mathbb{R}$ with |z| < r. In order to derive the asymptotics of the integral $\int_C h(z)e^{ng(z)} dz$ we use the saddle point method (see, for example, [2] or [5] for general references) and choose the contour C to be the circle around the origin with radius $z(\theta)$ such that it passes through the zero $z(\theta)$ of the derivative g'. Note that $g'(z(\theta)) = 0$, so $g(z(\theta)e^{it})$ has Taylor expansion

$$g(z(\theta)e^{it}) = \sum_{j=0}^{\infty} \frac{g^{(j)}(z(\theta))}{j!} (z(\theta))^j (e^{it} - 1)^j$$

= $g(z(\theta)) + \frac{g''(z(\theta))}{2} (z(\theta))^2 (e^{it} - 1)^2 + \frac{g'''(z(\theta))}{3!} (z(\theta))^3 (e^{it} - 1)^3 + O(t^4)$

leading to the expansions $\operatorname{Re}(g(z(\theta)e^{it})) = g(z(\theta)) - (z(\theta))^2 g''(z(\theta))t^2/2 + O(t^4)$ and $\operatorname{Im}(g(z(\theta)e^{it})) = O(t^3)$. The saddle point method yields the asymptotics

$$\int_C h(z) e^{ng(z)} dz \sim i \sqrt{\frac{2\pi}{ng''(z(\theta))}} e^{ng(z(\theta))} h(z(\theta)).$$

Dividing this expression by $2\pi i$ and writing z instead of $z(\theta)$ for convenience yields

$$\frac{\sigma_{n-k}((n-l)\theta)}{(n-k)!} \sim \frac{e^{ng(z)}h(z)}{\sqrt{2\pi ng''(z)}} = \frac{1}{\sqrt{2\pi n}}(a(\theta))^n b_{kl}(\theta)$$

with $a(\theta) := e^{g(z)} = e^{\theta\phi(z)}/z$ and $b_{kl}(\theta) := h(z)/\sqrt{g''(z)} = z^{k-1}e^{-l\theta\phi(z)}/\sqrt{z^{-2} + \theta\phi''(z)} = z^k e^{-l\theta\phi(z)}/\sqrt{1 + \theta z^2 \phi''(z)}$. Thus, (18) is established. Note that for k = l = 0 we have

$$\frac{\sigma_n(n\theta)}{n!} = \frac{1}{2\pi i} \int_C \frac{e^{ng(z)}}{z} dz = \frac{1}{2\pi i} \int_{-\pi}^{\pi} \frac{e^{ng(z(\theta)e^{it})}}{z(\theta)e^{it}} iz(\theta)e^{it} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ng(z(\theta)e^{it})} dt.$$

Taking the real part yields

$$\frac{\sigma_n(n\theta)}{n!} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \operatorname{Re}(e^{ng(z(\theta)e^{it})}) dt \sim \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{n\operatorname{Re}(g(z(\theta)e^{it}))} dt,$$
(20)

where the last asymptotics is based on the Laplace method as follows. Choose a sequence $(\delta_n)_{n \in \mathbb{N}}$ of positive real numbers satisfying $n\delta_n^2 \to \infty$ and $n\delta_n^3 \to 0$, for example, $\delta_n := n^{-\alpha}$ for some fixed $\alpha \in (1/3, 1/2)$. Decomposing the first integral in (20) into the two parts

$$I_1 := \int_{-\delta_n}^{\delta_n} \operatorname{Re}(e^{ng(z(\theta)e^{it})}) dt \quad \text{and} \quad I_2 := \int_{\{\delta_n < |t| \le \pi\}} \operatorname{Re}(e^{ng(z(\theta)e^{it})}) dt$$

we can approximate I_1 and show that I_2 is negligible (in comparison to I_1) for large n. Obviously, $1 - x^2/2 \leq \cos x \leq 1$ for all $x \in \mathbb{R}$. Choosing $x := n \operatorname{Im}(g(z(\theta)e^{it}))$ and using $\operatorname{Im}(g(z(\theta)e^{it})) = O(t^3)$ and $nt^3 \to 0$ as $n \to \infty$ uniformly for all $|t| \leq \delta_n$ it follows that

$$\lim_{n \to \infty} \sup_{|t| \le \delta_n} |\cos(n \operatorname{Im}(g(z(\theta)e^{it}))) - 1| = 0.$$

Thus, as $n \to \infty$, the map $t \mapsto \cos(n \operatorname{Im}(g(z(\theta)e^{it})))$ converges uniformly on $[-\delta_n, \delta_n]$ to the constant map $t \mapsto 1$, which implies that

$$I_1 = \int_{-\delta_n}^{\delta_n} \cos(n \operatorname{Im}(g(z(\theta)e^{it}))) e^{n\operatorname{Re}(g(z(\theta)e^{it}))} dt \sim \int_{-\delta_n}^{\delta_n} e^{n\operatorname{Re}(g(z(\theta)e^{it}))} dt.$$

The second integral I_2 is negligible (in comparison to I_1) since

$$|I_2| \leq \int_{\{\delta_n < |t| \leq \pi\}} e^{n \operatorname{Re}(g(z(\theta)e^{it}))} dt,$$

 $\operatorname{Re}(g(z(\theta)e^{it})) = g(z(\theta)) - (z(\theta))^2 g''(z(\theta))t^2/2 + O(t^4) \text{ and } nt^2 \ge n\delta_n^2 \to \infty \text{ uniformly for } |t| > \delta_n.$ Thus, (20) is established. For all $N \in \mathbb{N}$ and all $k \in \mathbb{N}_0$ it follows from (18) that

$$\mathbb{P}(\mu_{N,1}=k) = \binom{N}{k} \frac{\sigma_k(\theta) \,\sigma_{N-k}((N-1)\theta)}{\sigma_N(N\theta)} = \frac{\sigma_k(\theta)}{k!} \frac{\frac{\sigma_{N-k}((N-l)\theta)}{(N-k)!}}{\frac{\sigma_N(N\theta)}{N!}}$$
$$\sim \frac{\sigma_k(\theta)}{k!} \frac{(a(\theta))^n b_{k1}(\theta)/\sqrt{2\pi n}}{(a(\theta))^n b_{00}(\theta)/\sqrt{2\pi n}} = \frac{\sigma_k(\theta)}{k!} \frac{b_{k1}(\theta)}{b_{00}(\theta)} = \sigma_k(\theta) \frac{(z(\theta))^k}{k!} e^{-\theta\phi(z(\theta))}.$$

Thus, $\mu_{N,1} \to X$ in distribution as $N \to \infty$, where X has distribution (17). In the following, for arbitrary but fixed p > 0, the convergence $\mathbb{E}(\mu_{N,1}^p) \to \mathbb{E}(X^p)$ as $N \to \infty$ of the p-th moments is established. For all $k, l \in \mathbb{N}_0$ and all $N \in \mathbb{N}$ we have

$$\frac{\sigma_{N-k}((N-l)\theta)}{(N-k)!} \leq \frac{\sigma_{N-k}(N\theta)}{(N-k)!} = \frac{1}{2\pi i} \int_C z^{k-1} e^{Ng(z)} dz$$
$$= \frac{1}{2\pi i} \int_{-\pi}^{\pi} (z(\theta)e^{it})^{k-1} e^{Ng(z(\theta)e^{it})} iz(\theta)e^{it} dt = \frac{(z(\theta))^k}{2\pi} \int_{-\pi}^{\pi} e^{ikt} e^{Ng(z(\theta)e^{it})} dt.$$

Taking the complex absolute value it follows for all $k, l \in \mathbb{N}$ and all $N \in \mathbb{N}$ that

$$\frac{\sigma_{N-k}((N-l)\theta)}{(N-k)!} \leq \frac{(z(\theta))^k}{2\pi} \int_{-\pi}^{\pi} |e^{Ng(z(\theta)e^{it})}| \, dt = \frac{(z(\theta))^k}{2\pi} \int_{-\pi}^{\pi} e^{N\operatorname{Re}(g(z(\theta)e^{it}))} \, dt$$

Since, by (20), $(2\pi)^{-1} \int_{-\pi}^{\pi} e^{N \operatorname{Re}(g(z(\theta)e^{it}))} dt \sim \sigma_N(N\theta)/N!$, it follows that there exists a constant $N_0 \in \mathbb{N}$ (which may depend on θ and ϕ but not on k and l) such that

$$\frac{\sigma_{N-k}((N-l)\theta)}{(N-k)!} \leq 2(z(\theta))^k \frac{\sigma_N(N\theta)}{N!}$$

for all $k, l \in \mathbb{N}_0$ and all $N \ge N_0$. In particular, for all $N \ge N_0$ and all $k \in \mathbb{N}_0$,

$$\mathbb{P}(\mu_{N,1}=k) = \frac{\sigma_k(\theta)}{k!} \frac{\frac{\sigma_{N-k}((N-1)\theta)}{(N-k)!}}{\frac{\sigma_N(N\theta)}{N!}} \le \frac{\sigma_k(\theta)}{k!} 2(z(\theta))^k = \kappa(\theta)\mathbb{P}(X=k),$$

where $\kappa(\theta) := 2e^{\theta\phi(z(\theta))} \in (0,\infty)$. For arbitrary but fixed p > 0 the map $k \mapsto k^p \mathbb{P}(X = k)$, $k \in \mathbb{N}_0$, is integrable with respect to the counting measure $\varepsilon_{\mathbb{N}_0}$ on \mathbb{N}_0 , since $\int k^p \mathbb{P}(X = k)$.

 $k \geq \varepsilon_{N_0}(dk) = \sum_{k=0}^{\infty} k^p \mathbb{P}(X=k) = \mathbb{E}(X^p) < \infty$. By dominated convergence we therefore have convergence $\mathbb{E}(\mu_{N,1}^p) \to \mathbb{E}(X^p)$ of all moments. In particular, $(N-1)c_N = \mathbb{E}((\mu_{N,1})_2) \to \mathbb{E}((X)_2) > 0$ and $(N-1)(N-2)d_N = \mathbb{E}((\mu_{N,1})_3) \to \mathbb{E}((X)_3)$ as $N \to \infty$. Thus, $d_N/c_N = O(1/N) \to 0$ as $N \to \infty$, which ensures (see [10] or [8, Theorem 4 (b)]) that the considered symmetric compound Poisson model is in the domain of attraction of the Kingman coalescent. \Box

4 Examples

We start with the two most popular examples, the Wright–Fisher model and the Dirichlet model. Note that (7) holds for these two models. These two examples have the advantage that most calculations can be done explicitly. For example, we will verify the asymptotic results stated in Theorem 3.2 directly.

Example 4.1 (Wright–Fisher model) For the standard symmetric Wright–Fisher model, $\phi(z) = z$. Therefore, ξ_n has a Poisson distribution with parameter θz and $\sigma_k(\theta) = \theta^k$, $k \in \mathbb{N}_0$, $\theta \in (0, \infty)$. For $k \in \mathbb{N}_0$ and $\theta \in (0, \infty)$ it follows that

$$\mathbb{P}(\mu_{N,1}=k) = \binom{N}{k} \frac{\sigma_k(\theta) \,\sigma_{N-k}((N-1)\theta)}{\sigma_N(N\theta)} = \binom{N}{k} \left(\frac{1}{N}\right)^k \left(1-\frac{1}{N}\right)^{N-k}$$

Thus, $\mu_{N,1}$ has a binomial distribution with parameters N and 1/N. In particular, for arbitrary but fixed $p \in \mathbb{N}$, $\mathbb{E}((\mu_{N,1})_p) = (N)_p/N^p \to 1 = \mathbb{E}((X)_p)$ as $N \to \infty$, where X has a Poisson distribution with parameter 1. The solution $z(\theta)$ of the equation $\theta z(\theta)\phi'(z(\theta)) = 1$ is $z(\theta) = 1/\theta$. Note that (18) holds with $a(\theta) := e\theta$ and $b_{kl}(\theta) := \theta^{-k}e^{-l}$, $k, l \in \mathbb{N}_0, \theta \in (0, \infty)$. The effective population size $N_e = 1/c_N = N$ coincides with the actual population size N, a well known result.

Example 4.2 (Dirichlet model) For the symmetric Dirichlet model, $\phi(z) = -\log(1-z)$, |z| < 1. Therefore, ξ_1 has a negative binomial distribution with parameters θ and 1-z. In particular, $\mathbb{E}(\xi_1) = \frac{\theta z \phi'(z)}{\theta z} = \frac{\theta z}{(1-z)}$. Moreover, $\sigma_k(\theta) = [\theta]_k := \theta(\theta+1) \cdots (\theta+k-1) = \Gamma(k+\theta)/\Gamma(\theta), \ k \in \mathbb{N}_0, \ \theta \in (0,\infty)$. Thus, for $k \in \mathbb{N}_0$ and $\theta \in (0,\infty)$,

$$\mathbb{P}(\mu_{N,1}=k) = \binom{N}{k} \frac{[\theta]_k[(N-1)\theta]_{N-k}}{[N\theta]_N} = \frac{[\theta]_k}{k!} (N)_k \frac{\Gamma(N\theta - \theta + N - k)\Gamma(N\theta)}{\Gamma(N\theta - \theta)\Gamma(N\theta + N)}$$

Since $\Gamma(x+c) \sim x^c \Gamma(x)$ as $x \to \infty$ for any $c \in \mathbb{R}$, it follows that

$$\mathbb{P}(\mu_{N,1} = k) \sim \frac{[\theta]_k}{k!} (N)_k \frac{(N\theta + N)^{-\theta - k} \Gamma(N\theta + N) \Gamma(N\theta)}{(N\theta)^{-\theta} \Gamma(N\theta) \Gamma(N\theta + N)} \\ = \frac{[\theta]_k}{k!} \frac{(N)_k}{(N\theta + N)^k} \left(\frac{\theta + 1}{\theta}\right)^{-\theta} \rightarrow \frac{[\theta]_k}{k!} \left(\frac{1}{\theta + 1}\right)^k \left(\frac{\theta}{\theta + 1}\right)^{\theta}.$$

Thus $\mu_{N,1} \to X$ in distribution as $N \to \infty$ where X has a negative binomial distribution with parameters θ and $\theta/(\theta + 1) \in (0, 1)$. Moreover, for arbitrary but fixed $p \in \mathbb{N}$, $\mathbb{E}((\mu_{N,1})_p) = (N)_p[\theta]_p/[N\theta]_p \to [\theta]_p/\theta^p = \mathbb{E}((X)_p)$ as $N \to \infty$. Thus, $\mathbb{E}(\mu_{N,1}^p) \to \mathbb{E}(X^p)$ as $N \to \infty$ for all $p \in \mathbb{N}$ in agreement with the general results derived in the proof of Theorem 3.2. The solution $z(\theta)$ of the equation $\theta z(\theta) \phi'(z(\theta)) = 1$ is $z(\theta) = 1/(\theta + 1)$. Note that (18) holds with $a(\theta) := (\theta + 1)^{\theta+1}/\theta^{\theta}$ and $b_{kl}(\theta) := (\theta/(\theta + 1))^{l\theta+1/2}(1/(\theta + 1))^k$, $k, l \in \mathbb{N}_0, \theta \in (0, \infty)$. The symmetric Dirichlet model has effective population size $N_e = 1/c_N = (N-1)/\mathbb{E}((\mu_{N,1})_2) = (N\theta + 1)/(\theta + 1) \sim N\varrho$ with $\varrho = \theta/(\theta + 1) = 1/\mathbb{E}((X)_2)$, in agreement with Theorem 3.2.

In the following examples are studied which do not satisfy (7). We start with a model which involves the absolute Lah numbers. This example was the starting point for writing this erratum, since we recognized that the formula (6), and, hence, the relation (7), are not satisfied for this model.

Example 4.3 Suppose that $\phi_m = m!$ for all $m \in \mathbb{N}$ or, equivalently, that $\phi(z) = z/(1-z)$, |z| < 1. Note that $\phi'(z) = 1/(1-z)^2$ and that $\phi''(z) = 2/(1-z)^3 = 2\phi'(z)/(1-z)$, |z| < 1. Then (Comtet [4, Section 3.3, p. 135, Theorem B]), $\sigma_n(\theta) = \sum_{k=1}^n L(n,k)\theta^k$, $n \in \mathbb{N}$, $\theta \in \mathbb{R}$, where $L(n,k) := B_{nk}(1!, 2!, \ldots) = n!(n-1)!/(k!(k-1)!(n-k)!)$, $n \in \mathbb{N}$, $k \in \{1, \ldots, n\}$, denote the absolute Lah numbers. The solution $z = z(\theta) \in (0,1)$ of the equation $1 = \theta z \phi'(z) = \theta z/(1-z)^2$ is $z(\theta) = 1 + \theta/2 - \sqrt{\theta(\theta+4)}/2$. By Theorem 3.2, the model is in the domain of attraction of the Kingman coalescent and the effective population size N_e satisfies $N_e \sim \rho N$ with $\rho := 1/(1 + \theta z^2 \phi''(z)) = (1-z)/(1+z) < 1$, since

$$1 + \theta z^2 \phi''(z) = 1 + \theta z^2 \phi'(z) \frac{2}{1-z} = 1 + \frac{2z}{1-z} = \frac{1+z}{1-z}$$

Note that (18) holds with $a(\theta) = e^{\theta \phi(z)}/z = e^{\theta z/(1-z)}/z = e^{1-z}/z$ and $b_{kl}(\theta) := z^k e^{-l\theta \phi(z)}/\sqrt{1+\theta z^2 \phi''(z)} = z^k e^{-l(1-z)}/\sqrt{(1+z)/(1-z)}, \ k, l \in \mathbb{N}_0, \ \theta \in (0,\infty).$

Let us generalize Example 4.3 as follows.

Example 4.4 Suppose that (see [1, p. 402, Eqs. (50) and (51)])

$$\phi_m = (m-1)! \binom{am}{m-1} b^{m-1}, \qquad m \in \mathbb{N},$$

where it is assumed that the real parameters a and b are either both negative or b > 0 and $a \ge 1$ (such that all the coefficients ϕ_m are nonnegative). Note that [1, Eq. (43)] the power series $\phi(z) = \sum_{m=1}^{\infty} \phi_m z^m / m!$ is the solution of the functional equation $\phi(z) = zf(\phi(z))$ with $f(x) := (1 + bx)^a$ and that ϕ is related to the generalized binomial series B_a (see, for example, [6, p. 200]) via $\phi(z) = (B_a(bz) - 1)/b$. It is readily checked that ϕ has radius of convergence

$$r := \lim_{m \to \infty} \frac{(m+1)\phi_m}{\phi_{m+1}} = \frac{1}{ab} \left(1 - \frac{1}{a}\right)^{a-1} \in (0, \infty)$$

and that $\phi(r-) = 1/(b(a-1))$ (= ∞ for a = 1), since, for $a \neq 1$, $\phi(r-) = 1/(b(a-1))$ is the only positive solution of the equation $\phi(r-) = rf(\phi(r-))$. Differentiating both sides of the functional equation $\phi(z) = zf(\phi(z))$ yields the derivatives

$$\phi'(z) = \frac{f(\phi(z))}{1 - zf'(\phi(z))} \quad \text{and} \quad \phi''(z) = \frac{2f'(\phi(z))\phi'(z) + zf''(\phi(z))(\phi'(z))^2}{1 - zf'(\phi(z))}, \quad |z| < r.$$

For a = 1 (and b > 0) we have $\phi_m = m!b^{m-1}$, $m \in \mathbb{N}$, and $\phi(z) = z/(1-bz)$, |z| < 1/b. For a = b = 1 we are back in the previous example involving the absolute Lah numbers. For a = 2 we have $\phi(z) = 4z/(1 + \sqrt{1-4bz})^2$, |z| < 1/(4b). For a = -1 (and b < 0) we have $\phi_m = (2m-2)!/(m-1)!(-b)^{m-1}$, $m \in \mathbb{N}$, and $\phi(z) = (\sqrt{1+4bz}-1)/(2b) = 2z/(1+\sqrt{1+4bz})$, |z| < 1/(-4b). For general parameters a and b, to the best of the authors knowledge, there seems to be no special function related to the power series ϕ . Nevertheless, an explicit expression for the solution $z = z(\theta)$ of the equation $\theta z \phi'(z) = 1$ is obtained as follows. Let $x = x(\theta)$ be the solution in the open interval (0, 1) of the quadratic equation $\theta x = b(1-x)(a-x)$, i.e.

$$x(\theta) = \frac{\theta + b + ab - \sqrt{(\theta + b + ab)^2 - 4ab^2}}{2b} \in (0, 1).$$
 (21)

We have $0 < (1-x)/\theta < 1/(b(a-1)) = \phi(r-)$ (= ∞ for a = 1). Thus, $z = z(\theta) := \phi^{-1}((1-x)/\theta) \in (0,r)$ is well defined. From $(1-x)/\theta = \phi(z) = zf(\phi(z)) = zf((1-x)/\theta)$ it follows that

$$z = \frac{1-x}{\theta} \frac{1}{f(\frac{1-x}{\theta})} = \frac{1-x}{\theta} \frac{1}{(1+b\frac{1-x}{\theta})^a} = \frac{1-x}{\theta} \frac{1}{(1+\frac{x}{a-x})^a}$$
$$= \frac{1-x}{\theta} \left(\frac{a-x}{a}\right)^a = \frac{1-x}{\theta} \frac{a-x}{a} \left(\frac{a-x}{a}\right)^{a-1} = \frac{x}{ab} \left(\frac{a-x}{a}\right)^{a-1}.$$
 (22)

Moreover,

$$f'(\phi(z)) = f'\left(\frac{1-x}{\theta}\right) = ab\left(1+b\frac{1-x}{\theta}\right)^{a-1} = ab\left(1+\frac{x}{a-x}\right)^{a-1} = ab\left(\frac{a}{a-x}\right)^{a-1} = \frac{x}{z}$$

and, hence,

$$\theta z \phi'(z) = \frac{\theta z f(\phi(z))}{1 - z f'(\phi(z))} = \frac{\theta \phi(z)}{1 - z f'(\phi(z))} = \frac{1 - x}{1 - z \frac{x}{z}} = 1$$

Thus, $z = z(\theta)$ satisfies $\theta z \phi'(z) = 1$. Theorem 3.2 is therefore applicable. Straightforward computations show that the limiting random variable X in Theorem 3.2 has second factorial moment $\mathbb{E}((X)_2) = 1 + \theta z^2 \phi''(z) = (a - x^2)/(a(1-x)^2)$. By Theorem 3.2, the model is in the domain of attraction of the Kingman coalescent and the effective population size N_e satisfies $N_e \sim \rho N$ as $N \to \infty$ with $\rho = 1/\mathbb{E}((X)_2) = (a(1-x)^2)/(a-x^2) \in (0,1)$. Note that the asymptotics (18) holds with $a(\theta) = e^{\theta \phi(z)}/z = e^{1-x}/z$ and

$$b_{kl}(\theta) = \frac{z^k e^{-l\theta\phi(z)}}{\sqrt{1+\theta z^2 \phi''(z)}} = z^k e^{-l(1-x)} \sqrt{\frac{a(1-x)^2}{a-x^2}}$$

with $x = x(\theta)$ and $z = z(\theta)$ defined in (21) and (22). At first glance the solution $x(\theta)$ of the quadratic equation $\theta x = b(1-x)(a-x)$ seems to come 'from nowhere'. In the following an intuitive argument is provided showing how $x(\theta)$ comes into play. It is known [1, Eq. (51)] that $\sigma_n(\theta) = \sum_{i=1}^n B(n, j)\theta^j$ with

$$B(n,j) = \frac{(n-1)!}{(j-1)!} {an \choose n-j} b^{n-j}.$$
 (23)

The fraction

$$\frac{B(n,n-j+1)(n\theta)^{n-j+1}}{B(n,n-j)(n\theta)^{n-j}} = \frac{n\theta j}{(n-j)(an-j+1)b}$$

is equal to 1 for $j = j_n$ with

$$j_n := \frac{n(\theta + b + ab) + b - \sqrt{(n(\theta + b + ab) + b)^2 - 4b(bn + abn^2)}}{2b} \sim n\frac{\theta + b + ab - \sqrt{(\theta + b + ab)^2 - 4ab^2}}{2b} = nx(\theta).$$
(24)

Intuitively, when n is large, the contribution to the sum $\sigma_n(n\theta) = \sum_{j=0}^{n-1} B(n, n-j)(n\theta)^{n-j}$ is essentially entirely originated from indices j having the property that j/n belongs to a (small) neighborhood of $x(\theta)$. The choice of the neighborhood is rather unimportant. It is hence not surprising that $x(\theta)$ plays a crucial role in finding the asymptotics of $\sigma_n(n\theta)$. Since the Bell numbers (23) are known explicitly, one may carry out the Laplace method in detail leading to an alternative proof of (18). We leave the details to the reader.

Example 4.5 Suppose that $\phi_m = 1$ for all $m \in \mathbb{N}$ or, equivalently, that $\phi(z) = e^z - 1, z \in \mathbb{C}$. Then (see, for example, Comtet [4, Section 3.3, p. 135, Theorem B]) $\sigma_n(\theta) = \sum_{k=1}^n S(n,k)\theta^k$, $n \in \mathbb{N}, \theta \in \mathbb{R}$, where the S(n,k) denote the Stirling numbers of the second kind. In this case the solution $z = z(\theta)$ of the equation $1 = \theta z \phi'(z) = \theta z e^z$ cannot be expressed in closed form anymore. By Theorem 3.2, $\mu_{N,1} \to X$ in distribution, where X is a random variable with distribution (17). Of course, by Theorem 3.2, the model is in the domain of attraction of the Kingman coalescent and the effective population size N_e satisfies $N_e \sim \rho N$ with $\rho := 1/(1 + \theta z^2 \phi''(z)) = 1/(1 + \theta z^2 e^z) = 1/(1 + z) < 1$.

Example 4.6 As in [7, Example 4.6], let $\alpha \in (0,1)$ and assume that $\phi_m = -\Gamma(m - \alpha)/\Gamma(-\alpha) = (-1)^{m+1}(\alpha)_m$, $m \in \mathbb{N}$, or, equivalently, that $\phi(z) = 1 - (1-z)^{\alpha}$, |z| < 1. Note that $\phi'(z) = \alpha(1-z)^{\alpha-1}$ and that $\phi''(z) = \alpha(1-\alpha)(1-z)^{\alpha-2} = \phi'(z)(1-\alpha)/(1-z)$. Closed forms for the positive real solution $z = z(\theta)$ of the equation $\theta z \phi'(z) = 1$ are only available for particular values of the parameter α . For example, $z(\theta) = 2\theta^{-2}(\sqrt{\theta^2 + 1} - 1)$ for $\alpha = 1/2$. By Theorem 3.2, the model is in the domain of attraction of the Kingman coalescent and the effective population size N_e satisfies $N_e \sim \rho N$ with $\rho := 1/(1+\theta z^2 \phi''(z)) = (1-z)/(1-\alpha z) \leq 1$, since

$$1 + \theta z^2 \phi''(z) = 1 + \theta z^2 \phi'(z) \frac{1 - \alpha}{1 - z} = 1 + z \frac{1 - \alpha}{1 - z} = \frac{1 - \alpha z}{1 - z}$$

which corrects the wrong asymptotic result for N_e provided at the bottom of p. 538 of [7].

Example 4.7 As in [7, Example 4.7] let $\alpha \in (0, \infty)$ and assume that $\phi_m = \Gamma(m+\alpha)/\Gamma(\alpha) = (-1)^m (-\alpha)_m$, $m \in \mathbb{N}$, or, equivalently, that $\phi(z) = (1-z)^{-\alpha} - 1$, |z| < 1. Note that $\phi'(z) = \alpha(1-z)^{-\alpha-1}$ and that $\phi''(z) = \alpha(\alpha+1)(1-z)^{-\alpha-2} = \phi'(z)(\alpha+1)/(1-z)$. For $\alpha = 1$ we are back in Example 4.3. Again, closed forms for the solution z of the equation $\theta z \phi'(z) = 1$ are only available for particular values of α . Theorem 3.2 is again applicable and the effective population size N_e satisfies $N_e \sim \rho N$ with $\rho := (1-z)/(1+\alpha z)$, since

$$1 + \theta z^2 \phi''(z) = 1 + \theta z^2 \phi'(z) \frac{\alpha + 1}{1 - z} = 1 + z \frac{\alpha + 1}{1 - z} = \frac{1 + \alpha z}{1 - z},$$

which again corrects the wrong asymptotic result for N_e provided on top of p. 539 of [7].

All examples considered so far satisfy $\phi'(r-) = \infty$ such that for all $\theta \in (0, \infty)$ the equation $\theta z \phi'(z) = 1$ has a unique real solution $z(\theta) \in (0, r)$. We finally provide an example satisfying $\phi'(r-) < \infty$.

Example 4.8 (Polylog model) Fix $\alpha \in (0, \infty)$ and suppose that ϕ is the polylog function, i.e. $\phi(z) := \sum_{m=1}^{\infty} m^{-\alpha} z^m$, |z| < 1. For $\alpha = 1$ the polylog model coincides with the Dirichlet model (Example 4.2). For $\alpha \to 0$ we are back in Example 4.3 whereas for $\alpha \to \infty$ we approach the Wright–Fisher model (Example 4.1).

If $\alpha \leq 2$, then $\phi'(1-) = \infty$, so Theorem 3.2 is applicable for all $\theta \in (0, \infty)$. There seems to be no closed expression available for the solution $z = z(\theta)$ of the equation $z\phi'(z) = 1/\theta$. The case $\alpha = 2$ is a nice exception where it is easily seen that $z(\theta) = 1 - e^{-1/\theta}, \ \theta \in (0, \infty)$.

If $\alpha > 2$, then $\phi'(1-) = \zeta(\alpha - 1) < \infty$. In this case the equation $\theta z \phi'(z) = 1$ admits a solution $z = z(\theta) \in (0, 1)$ if and only if $\theta > \theta_c$ with critical value $\theta_c := 1/\phi'(1-) < 1$. Thus, Theorem 3.2 is not applicable for $\theta \leq \theta_c$. We leave it open for future work to discuss the limiting behavior of the random variable $\mu_{N,1}$ in this case.

Acknowledgement. The authors thank Elmar Teufl for fruitful discussions concerning the Laplace method and the saddle point method.

References

- [1] BERESTYCKI, N. AND PITMAN, J. (2007) Gibbs distributions for random partitions generated by a fragmentation process. J. Stat. Phys. 127, 381–418. MR2314353
- [2] BRUIJN, N. G. (1981) Asymptotic Methods in Analysis. Dover, New-York. MR0671583
- [3] CHARALAMBIDES, CH. A. AND KYRIAKOUSSIS, A. (1985) An asymptotic formula for the exponential polynomials and a central limit theorem for their coefficients. *Discrete Math.* 54, 259–270. MR0790587
- [4] COMTET, L. (1974) Advanced Combinatorics. D. Reidel Publishing Company, Dordrecht, Holland. MR0460128
- [5] FLAJOLET, P. AND SEDGEWICK, R. (2009) Analytic Combinatorics. Cambridge University Press, Cambridge. MR2483235
- [6] GRAHAM, R. L., KNUTH, D. E., AND PATASHNIK, O. (1994) Concrete Mathematics. A Foundation for Computer Science, Second Edition, Addison Wesley, Reading. MR1397498
- [7] HUILLET, T. AND MÖHLE, M. (2011) Population genetics models with skewed fertilities: a forward and backward analysis. *Stochastic Models* 27, 521–554. MR number not yet available
- [8] MÖHLE, M. (2000) Total variation distances and rates of convergence for ancestral coalescent processes in exchangeable population models. Adv. Appl. Probab. 32, 983– 993. MR1808909
- [9] MÖHLE, M. (2011) Coalescent processes derived from some compound Poisson population models. *Electron. Comm. Probab.* 16, 567–582. MR number not yet available
- [10] MÖHLE, M. AND SAGITOV, S. (2001) A classification of coalescent processes for haploid exchangeable population models. Ann. Probab. 29, 1547–1562. MR1880231