

On a Λ -mutation model and the harmonic model

CRISTIAN GIARDINÀ¹ and MARTIN MÖHLE²

May 28, 2026

Abstract

We introduce a continuous-time mutation model with two types determined by a finite measure Λ on the unit interval. The model satisfies a certain consistency property known from mathematical population genetics and includes so called harmonic models being of interest in mathematical statistical physics. We mainly focus on the situation when the number of particles is equal to some constant N . Duality results and scaling limits as $N \rightarrow \infty$ for the forward and backward processes are provided leading to a commutative diagram. The stationary distribution of the forward process is studied with an emphasis on the case when Λ is a beta distribution. The work bridges particle models from mathematical statistical physics and mutation models from mathematical population genetics.

Keywords: Consistency; duality; harmonic model; Λ -mutation model; stationary distribution

1 Introduction

In this section we define the class of models that will be studied in this paper, we provide their interpretation and we explain how the paper is structured.

1.1 Model definition

Let $N \in \mathbb{N} := \{1, 2, \dots\}$. We study a class of continuous-time Markov processes $X^{(N)} = (X_t^{(N)})_{t \geq 0}$ with state space $E_N := \{0, \dots, N\}$ and generator $L^{X^{(N)}}$ acting on functions $f : E_N \rightarrow \mathbb{R}$ via

$$L^{X^{(N)}} f(n) = \sum_{k=1}^n \varphi(k, n) (f(n-k) - f(n)) + \sum_{k=1}^{N-n} \varphi(k, N-n) (f(n+k) - f(n)), \quad n \in E_N, \quad (1)$$

where empty sums are defined as zero. The nonnegative rates $\varphi(k, n)$, $1 \leq k \leq n$, are assumed to satisfy the consistency relations

$$\varphi(k, n) = \left(1 - \frac{k}{n+1}\right) \varphi(k, n+1) + \frac{k+1}{n+1} \varphi(k+1, n+1), \quad n \in \mathbb{N}, k \in \{1, \dots, n\}. \quad (2)$$

The consistency relations (2) are equivalent (see Proposition 2.1) to the commutator equations

$$[L^{X^{(N)}}, A^{(N)}] = 0, \quad N \in \mathbb{N}, \quad (3)$$

¹Department of Mathematics, University of Modena and Reggio Emilia, Via G. Campi 213/b, 44125 Modena, Italy, E-mail address: cristian.giardina@unimore.it

²Department of Mathematics, University of Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany, E-mail address: martin.moehle@uni-tuebingen.de

where $A^{(N)}$ denotes the annihilation operator acting as $A^{(N)}f(n) = nf(n-1) + (N-n)f(n)$, $n \in E_N$. Thus, for a consistent system, the action of removing at random a particle commutes with the evolution.

Proposition 2.1 also shows that the consistency relations (2) are equivalent to the existence of a finite measure Λ on the unit interval $[0, 1]$ such that

$$\varphi(k, n) = \binom{n}{k} \int_{[0,1]} t^{k-1}(1-t)^{n-k} \Lambda(dt), \quad n \in \mathbb{N}, k \in \{1, \dots, n\}. \quad (4)$$

For this reason, we call the process $X^{(N)}$ with generator (1) the ‘ Λ -mutation model’.

1.2 Model interpretation

The model introduced above can be seen as a mutation model in mathematical population genetics. To see this it is convenient to define $\mathbb{N}_0 := \{0, 1, 2, \dots\}$ and remark that the one-dimensional process $X^{(N)}$ is obtained from a two-dimensional (two-type) process with state space \mathbb{N}_0^2 , whose generator acts on bounded test functions $f : \mathbb{N}_0^2 \rightarrow \mathbb{R}$ via

$$\begin{aligned} Lf(n_1, n_2) &= \sum_{k=1}^{n_1} \varphi(k, n_1)(f(n_1 - k, n_2 + k) - f(n_1, n_2)) \\ &\quad + \sum_{k=1}^{n_2} \varphi(k, n_2)(f(n_1 + k, n_2 - k) - f(n_1, n_2)), \quad n_1, n_2 \in \mathbb{N}_0. \end{aligned} \quad (5)$$

The generator (1) is obtained from (5) when restricting on $n_1 + n_2 = N$. Again, the consistency relations (2) are equivalent (see Proposition 2.1) to the commutator relation

$$[L, A] = 0, \quad (6)$$

where A denotes the annihilation operator acting as $Af(n_1, n_2) = n_1f(n_1 - 1, n_2) + n_2f(n_1, n_2 - 1)$. The two-dimensional process might be interpreted in terms of a population genetics model as follows. Consider a population, where each individual has one of two possible types. If the process is in state (n_1, n_2) , then $k \in \{1, \dots, n_1\}$ of the n_1 particles of type 1 mutate to type 2 at the rate $\varphi(k, n_1)$ and $k \in \{1, \dots, n_2\}$ of the n_2 particles of type 2 mutate to type 1 at the rate $\varphi(k, n_2)$. The study of this model is partly motivated from related forward and backward processes known from mathematical population genetics models, namely the Λ -Fleming–Viot process (see, for example, Griffiths [14, Eq. (5)]) and its dual counter part, the block counting process of the Λ -coalescent. Note however that the latter process has state space \mathbb{N} and absorbing state 1 (corresponding to the most recent common ancestor), whereas the dual process Y of our Λ -mutation model (see Theorem 2.6) has state space \mathbb{N}_0 and absorbing state 0. Although there does not seem to be a direct relation between our mutation model and those models from mathematical population models, they share the common structure of duality linking the forward and backward evolution.

Another interpretation of the Λ -mutation model is in the framework of mathematical statistical physics. Indeed, the process with generator (5) belongs to the class of stochastic exchange models, which are a family of Markov processes in which a quantity, such as energy or wealth or mass, is randomly exchanged between agents interacting through the edges of a graph. The

stochastic exchange models include, among others, the *averaging process* [1] and the *Kipnis–Marchioro–Presutti model* [19]. If we interpret (n_1, n_2) as the number of particles at two sites of a graph, then we see from (5) that in the Λ -mutation model the redistribution rule is a function of the number of particles at the departure site only. In particular, if the measure Λ is the uniform distribution on $[0, 1]$, then the rate $\varphi(k, n) = 1/k$ does not depend on n and the model reduces to the so called *harmonic model* studied in the mathematical statistical physics literature on integrable particle systems (see, for example, [9]). The specific choice of Λ as the uniform measure (or more generally as the $\beta(1, b)$ -distribution) yields a direct connection to integrability, in the sense of the Yang–Baxter equation [8]. This connection has been exploited to identify in closed form the stationary measure for boundary-driven one-dimensional chains [4, 3], which is rarely available in statistical physics. For more general choices of the measure Λ and the study of intertwining and dynamical properties (e.g. spectral gap) we refer the reader to [10] and [18].

The mathematical analysis of both ‘worlds’ has much in common and essentially all results presented in this paper strengthen the deep bridge between particle models from mathematical statistical physics and mutation models from mathematical population genetics.

1.3 Main contributions

The main contributions of this paper are as follows. We introduce a class of continuous-time mutation models whose transition rates satisfy a natural consistency property, and we show that this property is equivalent to the existence of a finite measure Λ on the unit interval. This representation provides a unified framework encompassing both mutation models from population genetics and stochastic exchange models from statistical physics. We further analyze the behavior of these models in the large population limit and identify a limiting Markov process on $[0, 1]$, whose generator has an explicit nonlocal form reflecting the underlying Λ -mechanism. A key contribution of the paper is the systematic identification of duality relations. We construct dual processes for both the discrete and continuous models, showing that the forward dynamics can be related to suitable death processes through explicit duality functions. These dualities are derived from the algebraic structure of the generator, in particular from its action on monomials and factorial moments, and provide a powerful tool to analyze the evolution. In particular, they allow one to transfer information between the forward process and its dual, yielding insight into moment dynamics, scaling limits, and structural properties of the model. Finally, we investigate the stationary distribution of the limiting process. We obtain a recursive characterization of its moments and a functional equation for its moment generating function. In the important special case where Λ is a beta distribution, we derive more explicit results, including a characterization of when the stationary distribution is itself a beta distribution, as well as regimes exhibiting reversibility and connections to integrable (harmonic) models.

1.4 Paper organization

This article is organized as follows. The main results are collected in Section 2 with an emphasis on consistency in Subsection 2.1 and scaling limits and duality in Subsection 2.2, leading to the commutative diagram illustrated in Figure 1. A more advanced duality result is provided in Subsection 2.3. Moreover, Section 2 contains detailed information on the model

for the particular case when Λ is a beta distribution (see Section 2.4). Proofs are provided in Section 3.

2 Results

2.1 Consistency

We start our results with equivalent conditions relating consistency to commutator equations and to a measure Λ on the unit interval.

Proposition 2.1 (Consistency). *Let $(\varphi(k, n))_{n \in \mathbb{N}, k \in \{1, \dots, n\}}$ be a triangular array of nonnegative real numbers. Then the following four statements are equivalent.*

- (i) *The array $(\varphi(k, n))_{n \in \mathbb{N}, k \in \{1, \dots, n\}}$ satisfies the consistency relations (2).*
- (ii) *There exists a finite measure Λ on the unit interval $[0, 1]$ satisfying (4).*
- (iii) *The generator L , defined via (5), satisfies the commutator relation (6).*
- (iv) *For every $N \in \mathbb{N}$, the generator $L^{X^{(N)}}$, defined via (1), satisfies the commutator relation (3).*

2.2 Scaling limits and duality

Our first main result is a scaling limit for the process $X^{(N)}$ as $N \rightarrow \infty$.

Theorem 2.2 (Scaling limit for $X^{(N)}/N$). *Assume that $X_0^{(N)}/N \rightarrow X_0$ in distribution as $N \rightarrow \infty$ for some random variable X_0 . Then, as $N \rightarrow \infty$, the space-scaled process $(X_t^{(N)}/N)_{t \geq 0}$ converges in $D_{[0,1]}([0, \infty))$ to the process $X = (X_t)_{t \geq 0}$ having state space $[0, 1]$, whose generator L^X acts on test functions $f \in C^2([0, 1])$ via*

$$L^X f(x) = \Lambda(\{0\})(1 - 2x)f'(x) + \int_{(0,1]} (f(x(1-t)) + f(x(1-t)+t) - 2f(x)) \frac{\Lambda(dt)}{t}, \quad (7)$$

where Λ is the finite measure on $[0, 1]$ uniquely determined via (4).

Remark (Two-type process). As in the case of the process $X^{(N)}$ with discrete state space, also the process X is obtained from a two-dimensional process. This is the process with state space $[0, \infty)^2$ whose generator acts on test functions $f \in C^2([0, \infty)^2)$ via

$$\begin{aligned} \mathcal{L}f(x_1, x_2) &= -\Lambda(\{0\})(x_1 - x_2) \left(\frac{\partial}{\partial x_1} - \frac{\partial}{\partial x_2} \right) f(x_1, x_2) \\ &\quad + \int_{(0,1]} (f(x_1 - x_1 t, x_2 + x_1 t) - f(x_1, x_2)) \frac{\Lambda(dt)}{t} \\ &\quad + \int_{(0,1]} (f(x_1 + x_2 t, x_2 - x_2 t) - f(x_1, x_2)) \frac{\Lambda(dt)}{t}. \end{aligned} \quad (8)$$

The generator (7) is obtained from (8) when restricting to $x_1 + x_2 = 1$. Furthermore, when Λ is the uniform measure on the interval $[0, 1]$, because of the occurrence of the $1/t$ -term in the integral in (7), we speak of the continuous harmonic model.

The following lemma clarifies the action of the generator L^X on monomials, a result that will be central to prove duality later.

Lemma 2.3 (Action of L^X on monomials). *The generator L^X acts on the n -th monomial $f_n(x) := x^n$ via*

$$L^X f_n(x) = \sum_{k=1}^n \varphi(k, n) x^{n-k} - 2\varphi(n) x^n, \quad n \in \mathbb{N}_0, \quad (9)$$

where $\varphi(k, n)$ is defined via (4) and the total rate $\varphi(n)$ is defined via

$$\varphi(n) := \sum_{k=1}^n \varphi(k, n) = dn + \kappa + \int_{(0,1)} \frac{1 - (1-t)^n}{t} \Lambda(dt), \quad n \in \mathbb{N}_0, \quad (10)$$

with drift $d := \Lambda(\{0\})$ and killing rate $\kappa := \Lambda(\{1\})$. In particular, L^X acts on the identity f_1 via $L^X f_1(x) = (1 - 2x)\Lambda([0, 1])$.

Remark (Cone property). For $n \in \mathbb{N}_0$ let P_n denote the space of all polynomials on $[0, 1]$ of degree less than or equal to n . Eq. (9) in particular implies that $L^X P_n \subseteq P_n$ for all $n \in \mathbb{N}_0$, a crucial and useful property of L^X .

Remark (The rates are determined by the total rates). Let $n \in \mathbb{N}$ and $k \in \{1, \dots, n\}$ and assume for the moment that Λ has no mass at 0. Then, by (4),

$$\begin{aligned} \varphi(k, n) &= \binom{n}{k} \int_{(0,1)} (1 - (1-t))^k (1-t)^{n-k} \frac{\Lambda(dt)}{t} \\ &= \binom{n}{k} \int_{(0,1)} \sum_{j=0}^k \binom{k}{j} (-1-t)^{k-j} (1-t)^{n-k} \frac{\Lambda(dt)}{t} \\ &= \binom{n}{k} \int_{(0,1)} \sum_{j=0}^k \binom{k}{j} (-1)^{k-j} (1-t)^{n-j} \frac{\Lambda(dt)}{t}. \end{aligned}$$

Exploiting the relation $\sum_{j=0}^k \binom{k}{j} (-1)^{k-j} = (1-1)^k = 0$ yields

$$\begin{aligned} \varphi(k, n) &= \binom{n}{k} \int_{(0,1)} \sum_{j=0}^k \binom{k}{j} (-1)^{k-j} ((1-t)^{n-j} - 1) \frac{\Lambda(dt)}{t} \\ &= \binom{n}{k} \sum_{j=0}^k \binom{k}{j} (-1)^{k-j+1} \varphi(n-j). \end{aligned} \quad (11)$$

It is easily seen that (11) remains valid if $\Lambda(\{0\}) > 0$. The rates $\varphi(k, n)$ can hence be computed from the total rates (10) via (11). Similar relations between the rates and the total rates are known for coalescent processes with multiple collisions; see, for example, [21, Eq. (16)].

Remark (Relation to subordinators). The right-hand side in (10) is even defined with n replaced by an arbitrary real number $\eta \geq 0$. Let Q denote the measure on $(0, 1)$ defined via $Q(dt) := \Lambda(dt)/t$ and let $\varrho := Q_T$ denote the image of Q under the transformation $T(t) := -\log(1-t)$, $t \in (0, 1)$. Then, for all $\eta \geq 0$,

$$\varphi(\eta) := d\eta + \kappa + \int_{(0,1)} (1 - (1-t)^\eta) Q(dt) = d\eta + \kappa + \int_{(0,\infty)} (1 - e^{-\eta u}) \varrho(du).$$

Thus, the map $\eta \mapsto \varphi(\eta)$, $\eta \geq 0$, is the Laplace exponent of a Lévy process with non-decreasing paths (subordinator) $S = (S_t)_{t \geq 0}$ with drift d , killing rate κ and Lévy measure ϱ . Note that the generator L^S of the subordinator S acts on functions $f \in C^2([0, \infty))$ with $f, f', f'' \in C_0([0, \infty))$ via

$$L^S f(x) = df'(x) - \kappa f(x) + \int_{(0,\infty)} (f(x+u) - f(x)) \varrho(du), \quad x \in [0, \infty). \quad (12)$$

Obviously, from (7) it follows that the generator L^X of the process X obeys the decomposition $L^X = L_1 + L_2$, where

$$L_1 f(x) := -dx f'(x) + \kappa(f(0) - f(x)) + \int_{(0,1)} (f(x(1-t)) - f(x)) Q(dt)$$

and

$$L_2 f(x) := d(1-x)f'(x) + \kappa(f(1) - f(x)) + \int_{(0,1)} (f(x(1-t)+t) - f(x)) Q(dt).$$

Applying the transformation theorem to (12) it is readily checked that L_1 and L_2 are the generators of the processes $(e^{-S_t})_{t \geq 0}$ and $(1 - e^{-S_t})_{t \geq 0}$ respectively. Thus, the processes having generators L_1 and L_2 respectively both have a probabilistic interpretation in terms of the subordinator S . There does not seem to be a similar straightforward interpretation of the process X (having generator $L_1 + L_2$) in terms of the subordinator S .

Since the state space $[0, 1]$ is compact, the limiting process X in Theorem 2.2 has a unique stationary distribution μ (see, for example, Ethier and Kurtz [6, p. 240, Theorem 9.3 and Remark 9.4]). A recursion for the moments

$$\mu_n := \int f_n d\mu = \int x^n \mu(dx), \quad n \in \mathbb{N}_0, \quad (13)$$

of μ is obtained as follows. By [6, Proposition 9.2] and (9), for every $n \in \mathbb{N}_0$,

$$0 = \int L^X f_n d\mu = \int \left(\sum_{k=1}^n \varphi(k, n) f_{n-k} - 2\varphi(n) f_n \right) d\mu = \sum_{k=1}^n \varphi(k, n) \mu_{n-k} - 2\varphi(n) \mu_n.$$

If Λ is not the zero measure, then $\varphi(n) > 0$ for all $n \in \mathbb{N}$ and, hence, the moments satisfy the recursion $\mu_0 = 1$ and

$$\mu_n = \frac{1}{2\varphi(n)} \sum_{k=1}^n \varphi(k, n) \mu_{n-k}, \quad n \in \mathbb{N}, \quad (14)$$

and are hence uniquely determined. Note that the moments μ_0, μ_1, \dots uniquely determine the stationary distribution μ , since the state space $[0, 1]$ is compact. The following proposition provides a system of equations, which characterizes the moment generating function (mgf) of the stationary distribution μ .

Proposition 2.4 (Characterizing equation for the mgf of the stationary distribution). *The moment generating function (mgf) $g : [0, 1] \rightarrow [0, \infty)$ of the stationary distribution μ of X , defined via $g(x) := \sum_{n=0}^{\infty} \mu_n x^n / n!$ for all $x \in [0, 1]$, satisfies for every $x \in [0, 1]$ the equation*

$$2dxg'(x) + 2 \int_{(0,1]} \frac{g(x) - g(x(1-t))}{t} \Lambda(dt) = dxg(x) + \int_{(0,1]} \frac{e^{xt} - 1}{t} g(x(1-t)) \Lambda(dt) \quad (15)$$

with $d := \Lambda(\{0\})$. Conversely, if g is the mgf of some distribution on $[0, 1]$ and if g satisfies the equation (15) for all $x \in [0, 1]$, then g is the mgf of the stationary distribution μ of X .

Remark (Solving Eq. (15)). If $\Lambda = \delta_0$ is the Dirac measure at 0, then (15) reduces to the differential equation $2g' = g$ with solution $g(x) = e^{x/2}$ corresponding to $\mu = \delta_{1/2}$ as expected. For $\Lambda = \delta_1$, (15) immediately yields $g(x) = (e^x + 1)/2$ corresponding $\mu = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$. For general Λ it seems to be not straightforward to determine the solution g to (15). Equations of the form (15) are related to Carleman type integral equations, which could be solved using particular analytic techniques. We refer the reader exemplarily to Estrada and Kanwal [5] for such methods. The solutions typically involve residua and Cauchy principle values and are hence rarely known explicitly. We will see later that, for some particular but important classes of measures Λ , the stationary distribution μ is known explicitly.

We now turn to duality results for the limiting process X . We first consider the particular case when Λ is the uniform distribution, since in this case the duality result is explicit and straightforward to verify. Consider the duality kernel $D : [0, 1] \times \mathbb{N}_0 \rightarrow \mathbb{R}$ defined via

$$D(x, n) := (n+1)x^n, \quad x \in [0, 1], n \in \mathbb{N}_0. \quad (16)$$

With this D , the following duality result holds.

Proposition 2.5 (Duality, $\Lambda = U[0, 1]$). *Assume that Λ is the uniform distribution on $[0, 1]$. Then the process X is dual to Y with respect to the duality kernel D defined via (16), where Y is a death process with state space \mathbb{N}_0 and generator*

$$L^Y f(n) := \sum_{k=1}^n \left(\frac{1}{k} + \frac{1}{n-k+1} \right) (f(n-k) - f(n)), \quad n \in \mathbb{N}_0. \quad (17)$$

Proposition 2.5 generalizes for arbitrary non-zero measure Λ as follows.

Theorem 2.6 (Duality). *Assume that Λ is not the zero measure. Then X is dual to Y with respect to the duality kernel $D : [0, 1] \times \mathbb{N}_0 \rightarrow \mathbb{R}$ defined via*

$$D(x, n) := \frac{x^n}{\mu_n}, \quad x \in [0, 1], n \in \mathbb{N}_0, \quad (18)$$

where Y is a death process with state space \mathbb{N}_0 , whose generator L^Y acts on bounded functions $f : \mathbb{N}_0 \rightarrow \mathbb{R}$ via

$$L^Y f(n) = \frac{1}{\mu_n} \sum_{k=1}^n \varphi(k, n) \mu_{n-k} (f(n-k) - f(n)), \quad n \in \mathbb{N}_0. \quad (19)$$

Remark (Labelled particles). If the death process Y is in state $n \in \mathbb{N}$, then $k \in \{1, \dots, n\}$ specific particles are removed at the rate $\frac{\mu_{n-k}}{\mu_n} \int_{[0,1]} t^{k-1} (1-t)^{n-k} \Lambda(dt)$. Note that 0 is an absorbing state of Y .

Remark (Explicit duality function). Theorem 2.6 becomes explicit if the moments $\mu_n, n \in \mathbb{N}_0$, of μ are known. The most prominent example occurs when Λ is the uniform distribution. In this case (see Section 2.4) the stationary distribution μ is as well the uniform distribution having moments $\mu_n = 1/(n+1), n \in \mathbb{N}_0$. Thus, in this case the duality kernel is given by $D(x, n) = (n+1)x^n$, bringing us back to Proposition 2.5. Section 2.4 focusses more generally on the case when $\Lambda = \beta(a, b)$ is the beta distribution with parameters $a, b > 0$. In this case it turns out that the stationary distribution μ is a beta distribution if and only if $a \in \{1, 2\}$ or $a = b$. In this case it is the $\beta(b/a, b/a)$ -distribution. Section 2.4 also provides some information on the stationary distribution μ for the case $a + b = 1$. For all other cases, only little seems to be known on the stationary distribution μ .

We now turn to an analog duality result for the discrete process $X^{(N)}$. Instead of monomials, now descending factorials come into play. For $m \in \{0, \dots, N\}$ define $h_m^{(N)}(n) := (n)_m / (N)_m, n \in \{0, \dots, N\}$. Note that $1 = h_0^{(N)} \geq h_1^{(N)} \geq \dots \geq h_N^{(N)}$. The following key lemma is the discrete counterpart of Lemma 2.3. The proof of Lemma 2.7 is based on an elementary but useful sampling duality relation for binomial distributions, which is deferred to the proof section (see Lemma 3.1) for convenience.

Lemma 2.7 (Action of $L^{X^{(N)}}$ on $h_m^{(N)}$). *The generator $L^{X^{(N)}}$ acts on $h_m^{(N)}$ via*

$$L^{X^{(N)}} h_m^{(N)} = \sum_{k=1}^m \varphi(k, m) h_{m-k}^{(N)} - 2\varphi(m) h_m^{(N)}, \quad m \in \{0, \dots, N\}, \quad (20)$$

where $\varphi(\cdot, \cdot)$ and $\varphi(\cdot)$ are defined via (4) and (10).

Remark (Convergence of the stationary distributions). Clearly, the process $X^{(N)}$ has a stationary distribution, which is denoted by $\mu^{(N)}$ in the following. Define

$$\mu_m^{(N)} := \int h_m^{(N)} d\mu^{(N)} = \sum_{n=0}^N \frac{(n)_m}{(N)_m} \mu^{(N)}(\{n\}), \quad m \in \{0, \dots, N\}.$$

By [6, Proposition 9.2] and (20),

$$\begin{aligned} 0 &= \int L^{X^{(N)}} h_m^{(N)} d\mu^{(N)} = \int \left(\sum_{k=1}^m \varphi(k, m) h_{m-k}^{(N)} - 2\varphi(m) h_m^{(N)} \right) d\mu^{(N)} \\ &= \sum_{k=1}^m \varphi(k, m) \mu_{m-k}^{(N)} - 2\varphi(m) \mu_m^{(N)}. \end{aligned}$$

Thus, if $\Lambda \neq 0$, then the $\mu_m^{(N)}$ satisfy the recursion $\mu_0^{(N)} = 1$ and

$$\mu_m^{(N)} = \frac{1}{2\varphi(m)} \sum_{k=1}^m \varphi(k, m) \mu_{m-k}^{(N)}, \quad m \in \{1, \dots, N\}. \quad (21)$$

Comparing (21) with the recursion (14) it follows that $\mu_m^{(N)} = \mu_m$ if $m \in \{0, \dots, N\}$ and $\mu_m^{(N)} = 0$ otherwise. In particular, for all $m \in \{0, \dots, N\}$, $\mu_m^{(N)} = \mu_m$ does not depend on N , a crucial property simplifying the further analysis considerably. For example, for all $m \in \mathbb{N}_0$, $\mu_m^{(N)} \rightarrow \mu_m$ as $N \rightarrow \infty$. This convergence of moments implies the convergence $Z_N/N \rightarrow Z$ in distribution as $N \rightarrow \infty$, where Z_N and Z are random variables having distribution $\mu^{(N)}$ and μ respectively.

Theorem 2.8 (Duality for the process $X^{(N)}$). *Assume that Λ is not the zero measure. Then the process $X^{(N)}$ is dual to the process $Y^{(N)}$ with respect to the duality kernel $D^{(N)} : \{0, \dots, N\}^2 \rightarrow \mathbb{R}$ defined via*

$$D^{(N)}(n, m) := \frac{h_m^{(N)}(n)}{\mu_m^{(N)}} = \frac{(n)_m}{(N)_m \mu_m^{(N)}} = \frac{\binom{n}{m}}{\binom{N}{m} \mu_m^{(N)}}, \quad n, m \in \{0, \dots, N\}, \quad (22)$$

where $Y^{(N)}$ is a death process with state space $\{0, \dots, N\}$, whose generator $L^{Y^{(N)}}$ acts on functions $f : \{0, \dots, N\} \rightarrow \mathbb{R}$ via

$$L^{Y^{(N)}} f(m) = \frac{1}{\mu_m^{(N)}} \sum_{k=1}^m \varphi(k, m) \mu_{m-k}^{(N)} (f(m-k) - f(m)), \quad m \in \{0, \dots, N\}. \quad (23)$$

Remark ($Y^{(N)}$ as a restriction of Y). The right hand side in (23) does not depend on N , since $\mu_m^{(N)} = \mu_m$ for all $m \in \{0, \dots, N\}$ as observed after (21). The process $Y^{(N)}$ thus depends on N only via its state space $\{0, \dots, N\}$. One may interpret the process $Y^{(N)}$ simply as the process Y restricted to the state space $\{0, \dots, N\}$.

We finish our results for general Λ with a convergence result for the discrete dual process $Y^{(N)}$.

Theorem 2.9 (Convergence of the dual processes). *Assume that Λ is not the zero measure and that $Y_0^{(N)} \rightarrow Y_0$ in distribution as $N \rightarrow \infty$ for some random variable Y_0 . Then, as $N \rightarrow \infty$, the discrete dual process $Y^{(N)}$ with generator (23) converges in $D_{\mathbb{N}_0}([0, \infty))$ to the process Y with generator (19).*

The commutative diagram in Figure 1 illustrates these convergence and duality results obtained so far.

2.3 A more advanced duality result

We now turn to a more involved duality result. Having the particular duality statement in Proposition 2.5 for the uniform distribution $\Lambda = \beta(1, 1)$ in mind, it is natural to consider for

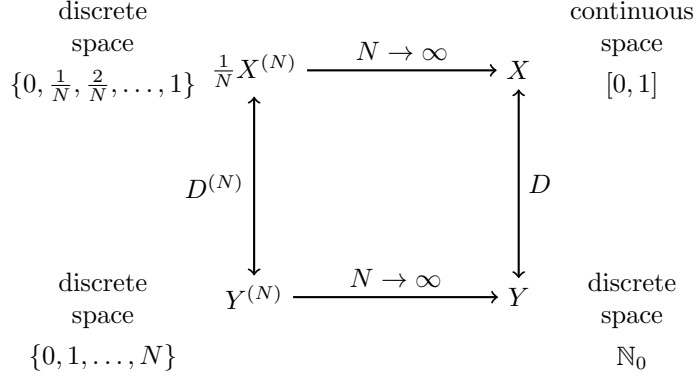


Figure 1: Commutative diagram for the process $X^{(N)}$ and the death process $Y^{(N)}$. The two right-arrows ‘ \rightarrow ’ stand for convergence in $D_{[0,1]}([0, \infty))$ and $D_{\mathbb{N}_0}([0, \infty))$ respectively as $N \rightarrow \infty$. The vertical updown-arrows ‘ \updownarrow ’ stand for duality, on the left hand side with respect to the sampling duality kernel $D^{(N)}(n, m) = \binom{n}{m} / \binom{N}{m} \mu_m^{(N)}$ and on the right hand side with respect to the moment duality kernel $D(x, n) = x^n / \mu_n$.

general finite measure Λ the Markov process Y with state space \mathbb{N}_0 and generator

$$L^Y f(n) := \sum_{k=1}^n (\varphi(k, n) + \varphi(n - k + 1, n)) (f(n - k) - f(n)), \quad n \in \mathbb{N}_0, \quad (24)$$

where $\varphi(\cdot, \cdot)$ is defined via (4). The following statement holds.

Theorem 2.10 (Polynomial duality with dual process (24)). *Let Λ be a finite measure on $[0, 1]$. Furthermore, let $r := (r_n)_{n \in \mathbb{N}_0}$ be a given sequence of real numbers. Then there exists a unique duality kernel $D = D_r : [0, 1] \times \mathbb{N}_0 \rightarrow \mathbb{R}$ (depending on the sequence r) satisfying the following three properties.*

(i) *The kernel D is of polynomial form*

$$D(x, n) = \sum_{j=0}^n d_{n,j} x^j, \quad x \in [0, 1], n \in \mathbb{N}_0, \quad (25)$$

for some real coefficients $d_{n,j}$, $n \in \mathbb{N}_0$, $j \in \{0, \dots, n\}$.

(ii) *The diagonal coefficients are given by $d_{n,n} = r_n$ for all $n \in \mathbb{N}_0$.*

(iii) *X is dual to Y with respect to D , where Y is the process with generator (24).*

Remark (Explicit polynomial duality). For general measure Λ the coefficients $d_{n,j}$, $n \in \mathbb{N}_0$, $j \in \{0, \dots, n\}$, do not have a simple form. One exception is uniform distribution $\Lambda = \beta(1, 1)$. In this case, choosing the diagonal entries as $d_{n,n} = n + 1$ ($=: r_n$), the recursion (68) yields $d_{n,j} = 0$ for all $0 \leq j < n$, and, hence, Theorem 2.10 boils down to Proposition 2.5. For $\Lambda \neq \beta(1, 1)$ it is impossible to choose the diagonal entries $d_{n,n}$, $n \in \mathbb{N}_0$, in such a way that all the off-diagonal entries $d_{n,j}$, $j < n$, vanish.

Remark (Space of duality functions). Theorem 2.10 shows that for each sequence $r = (r_n)_{n \in \mathbb{N}_0}$ there exists a duality kernel $D = D_r$ such that X is dual to Y with respect to D . It can be expected that the kernels D_{e_i} , where e_i denotes the i th unit vector in $\mathbb{R}^{\mathbb{N}_0}$, are linear independent elements of the Banach space $B([0, 1] \times \mathbb{N}_0)$ of all measurable, real-valued, bounded functions on $[0, 1] \times \mathbb{N}_0$. It is therefore conjectured that the space $\{D \in B([0, 1] \times \mathbb{N}_0) : X \text{ is dual to } Y \text{ with respect to } D\}$ of all duality kernels has infinite dimension.

2.4 The beta-mutation model

We focus in this section on the particular subclass of models when the measure Λ is the beta distribution $\beta(a, b)$ with parameters $a, b > 0$. Clearly, all results obtained so far are applicable. For example, the generator L^X of the limiting Markov process $X = (X_t)_{t \geq 0}$ in Theorem 2.2 acts on functions $f \in C^2([0, 1])$ via

$$L^X f(x) = \frac{1}{B(a, b)} \int_0^1 (f(x(1-t)) + f(x(1-t) + t) - 2f(x)) \frac{t^{a-1}(1-t)^{b-1}}{t} dt, \quad (26)$$

where $B(\cdot, \cdot)$ denotes the beta function. For obvious reasons we call this model the beta-mutation model. The rates (4) have the form

$$\varphi_{a,b}(k, n) := \binom{n}{k} \frac{B(k+a-1, n-k+b)}{B(a, b)}, \quad n \in \mathbb{N}, k \in \{1, \dots, n\} \quad (27)$$

and the total rates (10) reduce to

$$\begin{aligned} \varphi_{a,b}(n) &= \frac{1}{B(a, b)} \int_0^1 \frac{1 - (1-t)^n}{t} t^{a-1} (1-t)^{b-1} dt \\ &= \begin{cases} \frac{B(a-1, b) - B(a-1, n+b)}{B(a, b)} & \text{if } a \neq 1, \\ b(\Psi(n+b) - \Psi(b)) & \text{if } a = 1, \end{cases} \end{aligned} \quad (28)$$

where $\Psi := (\log \Gamma)' = \Gamma'/\Gamma$ denotes the digamma function (logarithmic derivative of the gamma function). Note that $B(a-1, b) - B(a-1, n+b) \rightarrow \Psi(n+b) - \Psi(b)$ as $a \rightarrow 1+$. Thus, the formula for $a = 1$ is the continuous extension of the formula for $a > 1$. For $0 < a < 1$, the negative argument $a-1 \in (-1, 0)$ occurring in the beta function of the above formula does not cause problems, since $B(p, q)$ is defined for all $p, q \in \mathbb{R} \setminus \{\dots, -2, -1, 0\}$ by analytic continuation.

The map $\eta \mapsto \varphi_{a,b}(\eta)$, $\eta \geq 0$, is the Laplace exponent of a drift-free subordinator, whose Lévy measure ϱ has density $u \mapsto e^{-bu}(1-e^{-u})^{a-2}/B(a, b)$, $u > 0$, with respect to the Lebesgue measure on $(0, \infty)$.

For $(a, b) = (1, 2s)$ with $s > 0$, the process X and the corresponding rates $\varphi_s(k, n) := \varphi_{1,2s}(k, n)$ play a certain role in the mathematical statistical physics literature on particle systems. We refer the reader to the work of Frassek, Giardinà and Kurchan [9, Eq. (2.30)] and to further studies on this model [8, 7, 3, 10, 4, 25]. In this case the model is called ‘harmonic’ because of the occurrence of the harmonic factor t^{-1} below the integral in (26) and since $\varphi_{\frac{1}{2}}(k, n) = \varphi_{1,1}(k, n) = 1/k$.

For $(a, b) = (1 - \alpha, \alpha)$ for some $\alpha \in (0, 1)$, the generator L^X occurs (up to a factor) as a mutation generator in the mathematical population genetics literature (see, for example, Handa [16, Eq. (1.3)] with $c_1 := c_2 := 1$).

The following proposition clarifies when the stationary distribution μ of the beta-mutation model is a beta distribution.

Proposition 2.11. *(Stationary distribution of the beta-mutation model). Let $\Lambda = \beta(a, b)$ with $a, b > 0$. Then the following two statements hold.*

- (i) *The stationary distribution μ of X is a beta distribution if and only if $a \in \{1, 2\}$ or $a = b$. In this case μ is the symmetric $\beta(b/a, b/a)$ -distribution.*
- (ii) *The stationary distribution μ of X is reversible if and only if $a = 1$.*

Example. (Duality for $\Lambda = \beta(1, b)$). Assume that $a = 1$. By Proposition 2.11 (i), the stationary distribution μ of X is the beta distribution $\beta(b, b)$ having moments

$$\mu_n := \frac{B(n+b, b)}{B(b, b)} = \frac{\Gamma(2b)\Gamma(n+b)}{\Gamma(b)\Gamma(n+2b)}, \quad n \in \mathbb{N}_0.$$

By Theorem 2.6, the process X is dual to Y with respect to the duality kernel D defined via

$$D(x, n) = \frac{x^n}{\mu_n} = \frac{\Gamma(b)\Gamma(n+2b)}{\Gamma(2b)\Gamma(n+b)} x^n, \quad x \in [0, 1], n \in \mathbb{N}_0,$$

where Y is a death process with state space \mathbb{N}_0 , whose generator L^Y acts on bounded functions $f : \mathbb{N}_0 \rightarrow \mathbb{R}$ via

$$L^Y f(n) := \frac{n!\Gamma(n+2b)}{(\Gamma(n+b))^2} \sum_{k=1}^n \frac{(\Gamma(n-k+2b))^2}{k(n-k)!\Gamma(n-k+2b)} (f(n-k) - f(n)), \quad n \in \mathbb{N}_0.$$

Remark (Spectral decomposition). In the reversible case, i.e. the $\beta(1, b)$ -model, it is possible to give a spectral decomposition for the generator of the process $X^{(N)}$, as well as for the process X . The eigenvectors are given by Hahn orthogonal polynomials in the first case, and Jacobi orthogonal polynomials in the second case. Because $a = 1$ is the only reversible case, this class is essentially the only model where the spectral decomposition can be obtained explicitly. The spectral decomposition immediately entails a closed formula for the semigroup of the processes. We shall report more about this in a future work.

The remaining part of this section deals with the $\beta(a, b)$ -mutation model for the case $a+b = 1$. Define the generalized Stieltjes transform S_b of the stationary distribution μ of the process X via $S_b(t) := \int (1+tx)^{-b} \mu(dx)$ for all $t \geq 0$.

Lemma 2.12 ($a+b = 1$, Stieltjes transform). *If $a+b = 1$, then the generalized Stieltjes transform S_b of the stationary distribution μ is given by*

$$S_b(t) = \frac{2}{1+(1+t)^b}, \quad t \geq 0. \tag{29}$$

Lemma 2.13 ($a+b = 1$, stationary distribution). *If $a+b = 1$, then the stationary distribution μ of the process X has a density f_μ with respect to the Lebesgue measure on $(0, 1)$, which is given by*

$$f_\mu(x) = \frac{b}{\pi i x^2} \int_C \frac{1}{(1+w)^{1-b} (1+\frac{w}{x})^{1-b} (1+(1+\frac{w}{x})^b)^2} dw, \quad x \in (0, 1), \quad (30)$$

where C is a counterclockwise contour that starts and ends at the point $w = -1$ and encloses the origin.

Remark (Contour integral). The contour C could be chosen to be the counterclockwise unit circle for instance. The function below the integral in (30) has three singularities, one at $w = -1$, one at $w = -x$ and another one at $w = x(e^{i\pi/b} - 1)$. We leave a further evaluation of the contour integral in (30) for future studies.

The following proposition provides formulas for the moments $\mu_n := \int x^n \mu(dx)$, $n \in \mathbb{N}_0$, of the stationary distribution of the process X .

Proposition 2.14 ($a + b = 1$, moments). *If $a + b = 1$ then the moments of the stationary distribution μ are given by*

$$\mu_n = \frac{1}{(-b)^n} \sum_{k=1}^n \frac{1}{2^k} \sum_{\ell=1}^k (-1)^\ell \binom{k}{\ell} \binom{b\ell}{n} = \frac{1}{(-b)^n} \sum_{k=1}^n \frac{(-1)^k}{2^k} k! \sum_{j=1}^n s(n, j) b^j S(j, k), \quad n \in \mathbb{N}, \quad (31)$$

where $s(., .)$ and $S(., .)$ denote the Stirling numbers of the first and second kind respectively.

Remark (Relation to generalized Stirling numbers) The sum over j arising in (31) is well-known to be of the form $\sum_{j=1}^n s(n, j) b^j S(j, k) = b^n S(n, k; 1/b, 1, 0) = b^k S(n, k; 1, b, 0)$, where $S(n, k; a, b, r)$ denote the generalized Stirling numbers studied by Hsu and Shiue [17]. Thus, the n th moment is also given by $\mu_n = \frac{1}{(-b)^n} \sum_{k=1}^n (-\frac{b}{2})^k k! S(n, k; 1, b, 0)$, $n \in \mathbb{N}$.

Remark (Additional comments). The first four moments are $\mu_1 = \frac{1}{2}$, $\mu_2 = \frac{1}{2(1+b)}$, $\mu_3 := \frac{2-b}{4(1+b)}$ and $\mu_4 := \frac{3(2-b^2)}{2(b+1)(b+2)(b+3)}$. Except for $b = \frac{1}{2}$, the fourth moment does not coincide with the fourth moment of the $\beta(b/a, b/a)$ -distribution, showing that for $b \neq \frac{1}{2}$ the stationary distribution cannot be a beta distribution. As expected, for $b \rightarrow 0$ one obtains $\mu_n = \frac{1}{2}$ for all $n \in \mathbb{N}$ corresponding to the stationary distribution $\mu = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$. Similarly, for $b \rightarrow 1$ one obtains $\mu_n = 2^{-n}$ for all $n \in \mathbb{N}$ corresponding to $\mu = \delta_{\frac{1}{2}}$.

The results of this section are illustrated in Figure 2.

3 Proofs

Proof of Proposition 2.1. Let us first verify the equivalence of (i) and (ii).

(i) \Rightarrow (ii): If the array $(\varphi(k, n))_{n \in \mathbb{N}, k \in \{1, \dots, n\}}$ satisfies the consistency relations (2), then the modified numbers $\lambda(k, n) := \varphi(k, n) / \binom{n}{k}$ obviously satisfy the modified consistency relations $\lambda(k, n) = \lambda(k, n+1) + \lambda(k+1, n+1)$, $n \in \mathbb{N}$, $k \in \{1, \dots, n\}$. Now proceed as in the proof

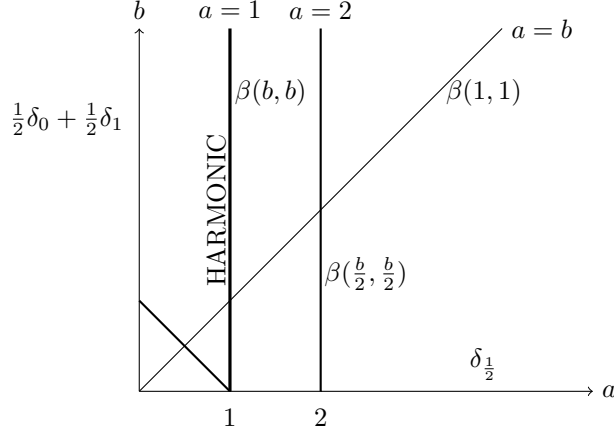


Figure 2: A graphical illustration of the results for the case when $\Lambda = \beta(a, b)$ is the beta distribution with parameters $a, b > 0$. If $a = 1$ then $\beta(b, b)$ is reversible. If $a + b = 1$ then $\mathbb{E}(1/(1 + tX_\infty)^b) = 2/(1 + (1 + t)^b)$.

of Pitman [23, Lemma 18]. If $\varphi(1, 1) = 0$, then (4) holds with $\Lambda := 0$. Assume now that $\varphi(1, 1) > 0$ and define $\mu(i, j) := \lambda(i + 1, i + j + 1)/\lambda(1, 1)$ for all $i, j \in \mathbb{N}_0$. Then, $\mu(0, 0) = 1$ and, for all $i, j \in \mathbb{N}_0$,

$$\begin{aligned} \mu(i + 1, j) + \mu(i, j + 1) &= \frac{\lambda(i + 2, i + j + 2) + \lambda(i + 1, i + j + 2)}{\lambda(1, 1)} \\ &= \frac{\lambda(i + 1, i + j + 1)}{\lambda(1, 1)} = \mu(i, j). \end{aligned}$$

De Finetti's representation for infinite exchangeable sequences of Bernoulli random variables implies that there exists a random variable S taking values in $[0, 1]$ such that $\mu(i, j) = \mathbb{E}(S^i(1 - S)^j)$ for all $i, j \in \mathbb{N}_0$. Thus, for all $n \in \mathbb{N}$ and $k \in \{1, \dots, n\}$, $\lambda(k, n) = \lambda(1, 1)\mu(k - 1, n - k) = \lambda(1, 1)\mathbb{E}(S^{k-1}(1 - S)^{n-k})$. Hence, (4) holds with $\Lambda := \lambda(1, 1)\mathbb{P}_S$, where \mathbb{P}_S denotes the distribution of S .

(ii) \Rightarrow (i): Conversely, assume that (4) holds for some finite measure Λ on $[0, 1]$. Then, a straightforward calculation shows that the numbers $\lambda(k, n) := \varphi(k, n)/\binom{n}{k} = \int_{[0, 1]} t^{k-1}(1 - t)^{n-k} \Lambda(dt)$ satisfy the relations $\lambda(k, n) = \lambda(k, n + 1) + \lambda(k + 1, n + 1)$, $n \in \mathbb{N}$, $k \in \{1, \dots, n\}$. Multiplication with $\binom{n}{k}$ shows that the numbers $\varphi(n, k)$ satisfy the consistency relations (2).

We now verify the equivalence of (i) and (iii).

(i) \Leftrightarrow (iii): We consider the generator L acting on bounded test functions $f : \mathbb{N}_0^2 \rightarrow \mathbb{R}$ via (5) and show that the consistency relations (2) for the rates are equivalent to

$$[L, A] = 0, \tag{32}$$

where A denotes the annihilation operator acting as

$$Af(n_1, n_2) = n_1 f(n_1 - 1, n_2) + n_2 f(n_1, n_2 - 1). \tag{33}$$

We will need the following:

$$\begin{aligned}
Lf(n_1 - 1, n_2) &= \sum_{k=1}^{n_1-1} \varphi(k, n_1 - 1)(f(n_1 - 1 - k, n_2 + k) - f(n_1 - 1, n_2)) \\
&+ \sum_{k=1}^{n_2} \varphi(k, n_2)(f(n_1 - 1 + k, n_2 - k) - f(n_1 - 1, n_2)), \quad (34)
\end{aligned}$$

$$\begin{aligned}
Lf(n_1, n_2 - 1) &= \sum_{k=1}^{n_1} \varphi(k, n_1)(f(n_1 - k, n_2 - 1 + k) - f(n_1, n_2 - 1)) \\
&+ \sum_{k=1}^{n_2-1} \varphi(k, n_2 - 1)(f(n_1 + k, n_2 - 1 - k) - f(n_1, n_2 - 1)), \quad (35)
\end{aligned}$$

as well as

$$Af(n_1 - k, n_2 + k) = (n_1 - k)f(n_1 - k - 1, n_2 + k) + (n_2 + k)f(n_1 - k, n_2 + k - 1) \quad (36)$$

and

$$Af(n_1 + k, n_2 - k) = (n_1 + k)f(n_1 + k - 1, n_2 - k) + (n_2 - k)f(n_1 + k, n_2 - k - 1). \quad (37)$$

We start by computing the action of LA . Using (5) we have

$$\begin{aligned}
LAf(n_1, n_2) &= \sum_{k=1}^{n_1} \varphi(k, n_1)(Af(n_1 - k, n_2 + k) - Af(n_1, n_2)) \\
&+ \sum_{k=1}^{n_2} \varphi(k, n_2)(Af(n_1 + k, n_2 - k) - Af(n_1, n_2)). \quad (38)
\end{aligned}$$

Inserting (33), (36) and (37) into the previous equation, we get

$$\begin{aligned}
LAf(n_1, n_2) &= \sum_{k=1}^{n_1} \varphi(k, n_1)(n_1 - k)f(n_1 - k - 1, n_2 + k) \\
&+ \sum_{k=1}^{n_1} \varphi(k, n_1)(n_2 + k)f(n_1 - k, n_2 + k - 1) \\
&- \sum_{k=1}^{n_1} \varphi(k, n_1)(n_1 f(n_1 - 1, n_2) + n_2 f(n_1, n_2 - 1)) \\
&+ \sum_{k=1}^{n_2} \varphi(k, n_2)(n_1 + k)f(n_1 + k - 1, n_2 - k) \\
&+ \sum_{k=1}^{n_2} \varphi(k, n_2)(n_2 - k)f(n_1 + k, n_2 - k - 1) \\
&- \sum_{k=1}^{n_2} \varphi(k, n_2)(n_1 f(n_1 - 1, n_2) + n_2 f(n_1, n_2 - 1)). \quad (39)
\end{aligned}$$

We now compute the action of AL . Using (33) we may write

$$ALf(n_1, n_2) = n_1Lf(n_1 - 1, n_2) + n_2Lf(n_1, n_2 - 1). \quad (40)$$

Inserting (34) and (35) into the previous equation, we get

$$\begin{aligned} ALf(n_1, n_2) &= n_1 \sum_{k=1}^{n_1-1} \varphi(k, n_1 - 1) (f(n_1 - 1 - k, n_2 + k) - f(n_1 - 1, n_2)) \\ &+ n_1 \sum_{k=1}^{n_2} \varphi(k, n_2) (f(n_1 - 1 + k, n_2 - k) - f(n_1 - 1, n_2)) \\ &+ n_2 \sum_{k=1}^{n_1} \varphi(k, n_1) (f(n_1 - k, n_2 - 1 + k) - f(n_1, n_2 - 1)) \\ &+ n_2 \sum_{k=1}^{n_2-1} \varphi(k, n_2 - 1) (f(n_1 + k, n_2 - 1 - k) - f(n_1, n_2 - 1)). \end{aligned} \quad (41)$$

Now we compare (39) with (41). We fix k with $2 \leq k \leq n_1$ and compare the coefficients of $f(n_1 - k, n_2 + k - 1)$. We get the equation

$$\varphi(k, n_1)(n_2 + k) + \varphi(k - 1, n_1)(n_1 - k + 1) = n_1\varphi(k - 1, n_1 - 1) + n_2\varphi(k, n_1). \quad (42)$$

The term $n_2\varphi(k, n_1)$ drops and we may rewrite the previous equation as

$$\varphi(k - 1, n_1 - 1) = \left(1 - \frac{k - 1}{n_1}\right)\varphi(k - 1, n_1) + \frac{k}{n_1}\varphi(k, n_1), \quad 2 \leq k \leq n_1. \quad (43)$$

If we now define $k' := k - 1$ and $n'_1 := n_1 - 1$, then we have

$$\varphi(k', n'_1) = \left(1 - \frac{k'}{n'_1 + 1}\right)\varphi(k', n'_1 + 1) + \frac{k' + 1}{n'_1 + 1}\varphi(k' + 1, n'_1 + 1), \quad 1 \leq k' \leq n'_1, \quad (44)$$

which is the consistency equation (2).

It remains to verify the equivalence of (i) and (iv). We essentially proceed as in the proof of the equivalence of (i) and (iii), but nevertheless provide the details because of the occurrence of the additional parameter N in the calculations.

(i) \Leftrightarrow (iv): For $N \in \mathbb{N}$ we consider the generator $L^{(N)} := L^{X^{(N)}}$ acting on test functions $f : E_N \rightarrow \mathbb{R}$ via (1) and show that the consistency relations (2) for the rates are equivalent to

$$[L^{(N)}, A^{(N)}] = 0 \quad \text{for all } N \in \mathbb{N}, \quad (45)$$

where $A^{(N)}$ denotes the annihilation operator acting as

$$A^{(N)}f(n) = nf(n - 1) + (N - n)f(n). \quad (46)$$

We will need the following:

$$\begin{aligned}
L^{(N)}f(n-1) &= \sum_{k=1}^{n-1} \varphi(k, n-1)(f(n-1-k) - f(n-1)) \\
&+ \sum_{k=1}^{N-n+1} \varphi(k, N-n+1)(f(n-1+k) - f(n-1)), \tag{47}
\end{aligned}$$

as well as

$$A^{(N)}f(n-k) = (n-k)f(n-k-1) + (N-n+k)f(n-k) \tag{48}$$

and

$$A^{(N)}f(n+k) = (n+k)f(n+k-1) + (N-n-k)f(n+k). \tag{49}$$

We start by computing the action of $L^{(N)}A^{(N)}$. Using (1) we have

$$\begin{aligned}
L^{(N)}A^{(N)}f(n) &= \sum_{k=1}^n \varphi(k, n)(A^{(N)}f(n-k) - A^{(N)}f(n)) \\
&+ \sum_{k=1}^{N-n} \varphi(k, N-n)(A^{(N)}f(n+k) - A^{(N)}f(n)). \tag{50}
\end{aligned}$$

Inserting (46), (48) and (49) into the previous equation, we get

$$\begin{aligned}
L^{(N)}A^{(N)}f(n) &= \sum_{k=1}^n \varphi(k, n)(n-k)f(n-k-1) \\
&+ \sum_{k=1}^n \varphi(k, n)(N-n+k)f(n-k) \\
&- \sum_{k=1}^n \varphi(k, n)(nf(n-1) + (N-n)f(n)) \\
&+ \sum_{k=1}^{N-n} \varphi(k, N-n)(n+k)f(n+k-1) \\
&+ \sum_{k=1}^{N-n} \varphi(k, N-n)(N-n-k)f(n+k) \\
&- \sum_{k=1}^{N-n} \varphi(k, N-n)(nf(n-1) + (N-n)f(n)). \tag{51}
\end{aligned}$$

We now compute the action of $A^{(N)}L^{(N)}$. Using (46) we may write

$$A^{(N)}L^{(N)}f(n) = nL^{(N)}f(n-1) + (N-n)L^{(N)}f(n). \tag{52}$$

Inserting (1) and (47) into the previous equation, we get

$$\begin{aligned}
A^{(N)}L^{(N)}f(n) &= n \sum_{k=1}^{n-1} \varphi(k, n-1)(f(n-1-k) - f(n-1)) \\
&+ n \sum_{k=1}^{N-n+1} \varphi(k, N-n+1)(f(n-1+k) - f(n-1)) \\
&+ (N-n) \sum_{k=1}^n \varphi(k, n)(f(n-k) - f(n)) \\
&+ (N-n) \sum_{k=1}^{N-n} \varphi(k, N-n)(f(n+k) - f(n)). \tag{53}
\end{aligned}$$

Now we compare (51) with (53). We fix k with $2 \leq k \leq n$ and compare the coefficients of $f(n-k)$. We get the equation

$$\varphi(k, n)(N-n+k) + \varphi(k-1, n)(n-k+1) = n\varphi(k-1, n-1) + (N-n)\varphi(k, n). \tag{54}$$

The term $(N-n)\varphi(k, n)$ drops and we may rewrite the previous equation as

$$\varphi(k-1, n-1) = \left(1 - \frac{k-1}{n}\right)\varphi(k-1, n) + \frac{k}{n}\varphi(k, n), \quad 2 \leq k \leq n. \tag{55}$$

If we now define $k' := k-1$ and $n' := n-1$, then we have

$$\varphi(k', n') = \left(1 - \frac{k'}{n'+1}\right)\varphi(k', n'+1) + \frac{k'+1}{n'+1}\varphi(k'+1, n'+1), \quad 1 \leq k' \leq n', \tag{56}$$

which is the consistency equation (2). \square

Proof of Theorem 2.2. Let us first verify that the space D of all polynomials $p : [0, 1] \rightarrow \mathbb{R}$ is a core for the generator $L := L^X$. Clearly, D is dense in the space H of all continuous functions $f : [0, 1] \rightarrow \mathbb{R}$. For $n \in \mathbb{N}_0$ let D_n denote the space of all polynomials of degree less than or equal to n . Then D_0, D_1, \dots is a sequence of finite-dimensional subspaces of D with $\bigcup_{n \in \mathbb{N}_0} D_n = D$. From (9) it follows that $LD_n \subseteq D_n$ for all $n \in \mathbb{N}_0$. We have $(\lambda - L)(D_n) = D_n$ for all λ not belonging to the set of eigenvalues of $L|_{D_n}$, i.e. for all but at most finitely many $\lambda > 0$. Thus, $(\lambda - L)(D) = (\lambda - L)(\bigcup_{n \in \mathbb{N}_0} D_n) = \bigcup_{n \in \mathbb{N}_0} D_n = D$ is dense in H for all but at most countably many $\lambda > 0$. By [6, p. 17, Proposition 3.1], D is a core for L .

Let us now turn to the generator $L^{X^{(N)}}$ of the process $X^{(N)}$. Recall that $d := \Lambda(\{0\})$. For $n \in \mathbb{N}_0$ and $p \in [0, 1]$ let $K_{n,p}$ be a random variable having a binomial distribution with parameters n and p . Plugging (4) into (1), a straightforward calculation shows that, for all $N \in \mathbb{N}$, $n \in \{0, \dots, N\}$ and all functions $f : \{0, \dots, N\} \rightarrow \mathbb{R}$,

$$\begin{aligned}
L^{X^{(N)}}f(n) &= dn(f(n-1) - f(n)) + \int_{(0,1]} \frac{\mathbb{E}(f(K_{n,1-t})) - f(n)}{t} \Lambda(dt) \\
&+ d(N-n)(f(n+1) - f(n)) + \int_{(0,1]} \frac{\mathbb{E}(f(N - K_{N-n,1-t})) - f(n)}{t} \Lambda(dt). \tag{57}
\end{aligned}$$

Now, let L_N denote the generator of the space-scaled process $X^{(N)}/N$ having state space $E_N := \{j/N : j \in \{0, \dots, N\}\}$. For any function $f : [0, 1] \rightarrow \mathbb{R}$ let the function $\pi_N f : E_N \rightarrow \mathbb{R}$ be defined by $\pi_N f(x) := f(x)$ for all $x \in E_N$. To prove convergence in $D_{[0,1]}([0, \infty))$, it suffices by standard theory on weak convergence of Markov processes (see, for example, Ethier and Kurtz [6, p. 28, Theorem 6.1, p. 28 and p. 232, Corollary 8.7] to verify that, for every $f \in D$,

$$\|L_N(\pi_N f) - \pi_N(Lf)\| := \sup_{x \in E_N} |L_N(\pi_N f)(x) - Lf(x)| \rightarrow 0, \quad N \rightarrow \infty. \quad (58)$$

The involved operators are linear. It hence suffices to verify (58) for every monomial $f = f_n$, $n \in \mathbb{N}_0$, defined via $f_n(x) := x^n$ for all $x \in [0, 1]$. In the following (58) is even verified for all monotone functions $f \in C^2([0, 1])$. Applying the space-scaling $x = n/N$ to (57) it follows for all $N \in \mathbb{N}$ and $x \in E_N$ that

$$\begin{aligned} L_N(\pi_N f)(x) &= dxN \left(f\left(x - \frac{1}{N}\right) - f(x) \right) + \int_{(0,1]} \frac{\mathbb{E}(f(\frac{K_{xN,1-t}}{N})) - f(x)}{t} \Lambda(dt) \\ &\quad + d(1-x)N \left(f\left(x + \frac{1}{N}\right) - f(x) \right) + \int_{(0,1]} \frac{\mathbb{E}(f(1 - \frac{K_{(1-x)N,1-t}}{N})) - f(x)}{t} \Lambda(dt). \end{aligned}$$

Therefore, for all $N \in \mathbb{N}$ and $x \in E_N$,

$$\begin{aligned} |L_N(\pi_N f)(x) - Lf(x)| &\leq dx \left| N \left(f\left(x - \frac{1}{N}\right) - f(x) \right) + f'(x) \right| \\ &\quad + \left| \int_{(0,1]} \frac{\mathbb{E}(f(\frac{K_{xN,1-t}}{N})) - f(x(1-t))}{t} \Lambda(dt) \right| \\ &\quad + d(1-x) \left| N \left(f\left(x + \frac{1}{N}\right) - f(x) \right) - f'(x) \right| \\ &\quad + \left| \int_{(0,1]} \frac{\mathbb{E}(f(1 - \frac{K_{(1-x)N,1-t}}{N})) - f(1 - (1-x)(1-t))}{t} \Lambda(dt) \right|. \end{aligned}$$

Uniformly over $x \in E_N$, all four terms on the right-hand side above converge to 0 as $N \rightarrow \infty$: For the first and the third term this uniform convergence to 0 is readily verified. For the second and the fourth term the pointwise convergence to 0 follows from $K_{n,p}/n \rightarrow p$ in distribution as $n \rightarrow \infty$ and by dominated convergence. Since f is assumed to be monotone, the maps $x \mapsto \mathbb{E}(f(K_{\lfloor xN \rfloor, 1-t}/N))$ and $x \mapsto \mathbb{E}(f(1 - K_{\lfloor (1-x)N \rfloor, 1-t}/N))$ are monotone in $x \in [0, 1]$. Thus, the convergence holds even uniformly for all $x \in [0, 1]$ by Pólya's [24] classical theorem on convergence of monotone sequences of functions. \square

Proof of Lemma 2.3. In this proof all functions inside integrals are understood to be contin-

uously extended at $t = 0$. For all $n \in \mathbb{N}$ and $x \in [0, 1]$,

$$\begin{aligned}
L^X f_n(x) &= \int_{[0,1]} (x^n(1-t)^n + (x(1-t) + t)^n - 2x^n) \frac{\Lambda(dt)}{t} \\
&= \int_{[0,1]} \left(x^n(1-t)^n + \sum_{k=0}^n \binom{n}{k} t^k (x(1-t))^{n-k} - 2x^n \right) \frac{\Lambda(dt)}{t} \\
&= \int_{[0,1]} \left(2x^n(1-t)^n + \sum_{k=1}^n \binom{n}{k} t^k (x(1-t))^{n-k} - 2x^n \right) \frac{\Lambda(dt)}{t} \\
&= 2x^n \int_{[0,1]} ((1-t)^n - 1) \frac{\Lambda(dt)}{t} + \sum_{k=1}^n \binom{n}{k} x^{n-k} \int_{[0,1]} t^k (1-t)^{n-k} \frac{\Lambda(dt)}{t} \\
&= -2\varphi(n)x^n + \sum_{k=1}^n \varphi(k, n)x^{n-k}
\end{aligned}$$

with $\varphi(k, n)$ and $\varphi(n)$ as defined in the statement of the lemma. \square

Proof of Proposition 2.4. We have $2\varphi_n \mu_n = \sum_{k=1}^n \mu_{n-k} \varphi(k, n)$, $n \in \mathbb{N}$. Multiplying this equation with $x^n/n!$ and summing over all $n \in \mathbb{N}$ yields

$$2 \sum_{n=1}^{\infty} \mu_n \frac{x^n}{n!} \varphi(n) = \sum_{n=1}^{\infty} \frac{x^n}{n!} \sum_{k=1}^n \mu_{n-k} \varphi(k, n).$$

The sum on the left-hand side is equal to

$$\begin{aligned}
&\sum_{n=1}^{\infty} \mu_n \frac{x^n}{n!} \left(nd + \int_{(0,1]} \frac{1 - (1-t)^n}{t} \Lambda(dt) \right) \\
&= d \sum_{n=1}^{\infty} \mu_n \frac{x^n}{(n-1)!} + \int_{(0,1]} \frac{1}{t} \sum_{n=1}^{\infty} \mu_n (1 - (1-t)^n) \frac{x^n}{n!} \Lambda(dt) \\
&= dx \frac{d}{dx} \sum_{n=1}^{\infty} \mu_n \frac{x^n}{n!} + \int_{(0,1]} \frac{1}{t} \sum_{n=1}^{\infty} \mu_n \frac{x^n - (x(1-t))^n}{n!} \Lambda(dt) \\
&= dxg'(x) + \int_{(0,1]} \frac{g(x) - g(x(1-t))}{t} \Lambda(dt).
\end{aligned}$$

The right-hand side is equal to

$$\begin{aligned}
&\sum_{n=1}^{\infty} \frac{x^n}{n!} \sum_{k=1}^n \mu_{n-k} \binom{n}{k} \int_{[0,1]} t^{k-1} (1-t)^{n-k} \Lambda(dt) \\
&= \int_{[0,1]} \sum_{k=1}^{\infty} \frac{x^k t^{k-1}}{k!} \sum_{n=k}^{\infty} \mu_{n-k} \frac{(x(1-t))^{n-k}}{(n-k)!} \Lambda(dt) \\
&= dxg(x) + \int_{(0,1]} \frac{e^{xt} - 1}{t} g(x(1-t)) \Lambda(dt),
\end{aligned}$$

which yields (15). Conversely, assume that (15) holds for all $x \in [0, 1]$, where g is the mgf of some probability measure Q on $[0, 1]$. Expanding all expressions below the integrals in terms of powers of x we can express both integrals in terms of the moments of Q . This yields that the moments of Q satisfy the recursion (14). Thus, $Q = \mu$ and, hence, g is the mgf of μ . \square

Proof of Proposition 2.5. Let $x \in [0, 1]$ and $n \in \mathbb{N}_0$. By the linearity of L^X and (9),

$$L^X D(\cdot, n)(x) = (n+1)L^X f_n(x) = -2(n+1)h_n x^n + (n+1) \sum_{k=1}^n \frac{x^{n-k}}{k},$$

where $h_n := \Psi(n+1) - \Psi(1) = \sum_{i=1}^n 1/i$ denotes the n th harmonic number. On the other hand, by the definition of the generator L^Y ,

$$\begin{aligned} L^Y D(x, \cdot)(n) &= \sum_{k=1}^n ((n-k+1)x^{n-k} - (n+1)x^n) \underbrace{\left(\frac{1}{k} + \frac{1}{n-k+1} \right)}_{= \frac{n+1}{k(n-k+1)}} \\ &= -(n+1)x^n \sum_{k=1}^n \left(\frac{1}{k} + \frac{1}{n-k+1} \right) + \sum_{k=1}^n (n-k+1)x^{n-k} \frac{n+1}{k(n-k+1)} \\ &= -2(n+1)h_n x^n + (n+1) \sum_{k=1}^n \frac{x^{n-k}}{k}. \end{aligned}$$

Thus, $L^X D(\cdot, n)(x) = L^Y D(x, \cdot)(n)$ for all $x \in [0, 1]$ and $n \in \mathbb{N}_0$, showing that X is dual to Y with respect to D . \square

Proof of Theorem 2.6. The linearity of the generator L^X and its action on the n th monomial f_n (see Lemma 2.3) yields

$$\begin{aligned} L^X D(\cdot, n)(x) &= \frac{L^X f_n(x)}{\mu_n} = \frac{1}{\mu_n} \sum_{k=1}^n \varphi(k, n) x^{n-k} - 2\varphi(n) \frac{x^n}{\mu_n} \\ &= \frac{1}{\mu_n} \sum_{k=1}^n \varphi(k, n) \mu_{n-k} D(x, n-k) - 2\varphi(n) D(x, n). \end{aligned}$$

Thanks to the recursion (14) for the moments μ_n , this expression can be rewritten as

$$\begin{aligned} L^X D(\cdot, n)(x) &= \frac{1}{\mu_n} \sum_{k=1}^n \varphi(k, n) \mu_{n-k} D(x, n-k) - \frac{1}{\mu_n} \sum_{k=1}^n \varphi(k, n) \mu_{n-k} D(x, n) \\ &= \frac{1}{\mu_n} \sum_{k=1}^n \mu_{n-k} \varphi(k, n) (D(x, n-k) - D(x, n)) = L^Y D(x, \cdot)(n), \end{aligned}$$

where Y is a continuous-time death process with state space \mathbb{N}_0 and generator (19). Thus, $L^X D(\cdot, n)(x) = L^Y D(x, \cdot)(n)$ for all $x \in [0, 1]$ and $n \in \mathbb{N}_0$, showing that X is dual to Y with respect to D . \square

In the following, as in the proof of Theorem 2.2, let $K_{n,p}$ denote a random variable having a binomial distribution with parameters $n \in \mathbb{N}$ and $p \in [0, 1]$. For the proof of Lemma 2.7, the following *sampling duality* result for binomial distributions is needed.

Lemma 3.1 (Sampling duality for the Binomial distribution). *For all $N \in \mathbb{N}$, $p \in [0, 1]$ and $n, m \in \{0, \dots, N\}$,*

$$\mathbb{E} \left(\frac{(n + K_{N-n,p})_m}{(N)_m} \right) = \mathbb{E} \left(\frac{\binom{n}{m-K_{m,p}}}{\binom{N}{m-K_{m,p}}} \right). \quad (59)$$

Proof. Consider a population of size N made of n individuals of type A and $N - n$ individuals of type B . Reproduction works as follows: each A always survives (equivalently, it produces exactly one descendant); each B survives independently with probability p (equivalently, it produces one offspring with probability p and has no offspring with probability $1 - p$). Thus, the number of B 's that survive is described by the binomial random variable $K_{N-n,p}$. After the reproduction step, the next generation contains $n + K_{N-n,p}$ individuals.

Consider sampling m individuals from the original population of size N . The result follows by computing the probability of the event $E := \{\text{all sampled individuals have descendants after the reproduction step}\}$ in the following two ways.

- *Thinning followed by sampling:* after thinning, A 's all survive, while exactly $K_{N-n,p}$ B 's survive. So there are $n + K_{N-n,p}$ individuals whose line survives. Given $K_{N-n,p}$, the probability that a sample of m individuals from the N original individuals all belong to this surviving set is $(n + K_{N-n,p})_m / (N)_m$. Taking expectation yields

$$\mathbb{P}(E) = \mathbb{E} \left(\frac{(n + K_{N-n,p})_m}{(N)_m} \right). \quad (60)$$

- *Sampling followed by thinning:* we now reverse the order and first sample m individuals. Let Z denote the number of B 's in the sample. Then Z has a hypergeometric distribution with parameters N , $N - n$ and m . Among these Z sampled B 's, each survives independently with probability p . Let Y be the number of sampled B 's that survive. Conditional on Z , we have $Y|Z \stackrel{d}{=} B(Z, p)$. A standard property of binomial thinning gives $Y \stackrel{d}{=} K_{m,p}$, i.e. the hypergeometric–binomial mixture collapses to a binomial distribution. For all sampled m individuals to survive, the remaining $m - Y$ individuals must correspond to A individuals. The probability that the $m - K_{m,p}$ surviving A 's come from the n A 's among the N individuals is $\binom{n}{m-K_{m,p}} / \binom{N}{m-K_{m,p}}$. Taking expectation yields

$$\mathbb{P}(E) = \mathbb{E} \left(\frac{\binom{n}{m-K_{m,p}}}{\binom{N}{m-K_{m,p}}} \right). \quad (61)$$

The proof is complete. □

Remark (Sampling duality). Lemma 3.1 can be stated in terms of Markov chains as follows. Let $B_j^{(n)}$, $j \in \mathbb{N}, n \in \mathbb{N}_0$, be iid Bernoulli random variables with parameter $p \in [0, 1]$. Fix

$N \in \mathbb{N}$. Given two initial variables X_0 and Y_0 taking values in $\{0, \dots, N\}$, define two Markov chains $X := (X_n)_{n \in \mathbb{N}_0}$ and $Y := (Y_n)_{n \in \mathbb{N}_0}$ recursively via

$$X_{n+1} := X_n + \sum_{j=1}^{N-X_n} B_j^{(n)} \quad \text{and} \quad Y_{n+1} := Y_n - \sum_{j=1}^{Y_n} B_j^{(n)}, \quad n \in \mathbb{N}_0.$$

Lemma 3.1 implies that X is dual to Y with respect to the sampling duality kernel $H_N : \{0, \dots, N\}^2 \rightarrow [0, 1]$ defined via $H_N(n, m) := (n)_m / (N)_m$, $n, m \in \{0, \dots, N\}$. Similar duality results are known for forward and backward processes in exchangeable Cannings population models (see, for example, [20, Theorem 3.1] and can be traced back to works of Cannings [2] and Gladstien [11, 12, 13], where sampling duality results are stated in elementary matrix notation without knowing the notion of duality. For a more recent related work using sampling duality techniques we refer the reader to González Casanova and Spanò [15].

Proof of Lemma 2.7. We have

$$L^{X^{(N)}} h_m^{(N)}(n) = \sum_{k=1}^n (h_m^{(N)}(n-k) - h_m^{(N)}(n)) \varphi(k, n) + \sum_{k=1}^{N-n} (h_m^{(N)}(n+k) - h_m^{(N)}(n)) \varphi(k, N-n).$$

The proof of Lemma 2.7 can hence be split into the proof of

$$\sum_{k=1}^n (h_m^{(N)}(n-k) - h_m^{(N)}(n)) \varphi(k, n) = -\varphi(m) h_m^{(N)}(n) \quad (62)$$

and of

$$\sum_{k=1}^{N-n} (h_m^{(N)}(n+k) - h_m^{(N)}(n)) \varphi(k, N-n) = \sum_{k=1}^m \varphi(k, m) h_{m-k}^{(N)}(n) - \varphi(m) h_m^{(N)}(n). \quad (63)$$

As in the proof of Theorem 2.2, let $K_{n,p}$ denote a random variable having a binomial distribution with parameters $n \in \mathbb{N}$ and $p \in [0, 1]$. Note that $K_{n,p}$ has descending factorial moments $\mathbb{E}((K_{n,p})_m) = p^m (n)_m$, $m \in \mathbb{N}_0$. Division by $(N)_m$ shows that $\mathbb{E}(h_m^{(N)}(K_{n,p})) = p^m h_m^{(N)}(n)$, $m \in \mathbb{N}_0$. Note furthermore that $n - K_{n,p}$ has the same distribution as $K_{n,1-p}$. Thus,

$$\mathbb{E}(h_m^{(N)}(n - K_{n,p})) = \mathbb{E}(h_m^{(N)}(K_{n,1-p})) = (1-p)^m h_m^{(N)}(n), \quad m \in \mathbb{N}_0. \quad (64)$$

In the following all functions inside integrals are understood to be continuously extended at

$t = 0$. Using the definition of $\varphi(k, n)$ and (64), the left hand side of (62) simplifies to

$$\begin{aligned}
& \sum_{k=1}^n (h_m^{(N)}(n-k) - h_m^{(N)}(n)) \varphi(k, n) \\
&= \sum_{k=1}^n (h_m^{(N)}(n-k) - h_m^{(N)}(n)) \int_{[0,1]} \mathbb{P}(K_{n,t} = k) \frac{\Lambda(dt)}{t} \\
&= \int_{[0,1]} \sum_{k=0}^n (h_m^{(N)}(n-k) - h_m^{(N)}(n)) \mathbb{P}(K_{n,t} = k) \frac{\Lambda(dt)}{t} \\
&= \int_{[0,1]} \left(\mathbb{E}(h_m^{(N)}(n - K_{n,t})) - h_m^{(N)}(n) \right) \frac{\Lambda(dt)}{t} \\
&= \int_{[0,1]} ((1-t)^m h_m^{(N)}(n) - h_m^{(N)}(n)) \frac{\Lambda(dt)}{t} \\
&= h_m^{(N)}(n) \int_{[0,1]} \frac{(1-t)^m - 1}{t} \Lambda(dt) = -\varphi(m) h_m^{(N)}(n),
\end{aligned}$$

where the last equality holds by the integral formula for the total rate $\varphi(m)$. Thus, (62) is established. It remains to verify (63). Using the definition of $\varphi(k, n)$, the left hand side of (63) turns into

$$\begin{aligned}
& \sum_{k=1}^{N-n} (h_m^{(N)}(n+k) - h_m^{(N)}(n)) \varphi(k, N-n) \\
&= \sum_{k=1}^{N-n} (h_m^{(N)}(n+k) - h_m^{(N)}(n)) \int_{[0,1]} \mathbb{P}(K_{N-n,t} = k) \frac{\Lambda(dt)}{t} \\
&= \int_{[0,1]} \sum_{k=0}^{N-n} (h_m^{(N)}(n+k) - h_m^{(N)}(n)) \mathbb{P}(K_{N-n,t} = k) \frac{\Lambda(dt)}{t} \\
&= \int_{[0,1]} \left(\mathbb{E}(h_m^{(N)}(n + K_{N-n,t})) - h_m^{(N)}(n) \right) \frac{\Lambda(dt)}{t}
\end{aligned}$$

Similarly, for the right hand side of (63) we obtain

$$\sum_{k=1}^m \varphi(k, m) h_{m-k}^{(N)}(n) - \varphi(m) h_m^{(N)}(n) = \int_{[0,1]} \left(\mathbb{E}(h_{m-K_{m,t}}^{(N)}(n)) - h_m^{(N)}(n) \right) \frac{\Lambda(dt)}{t}.$$

Thus, (63) holds, since $\mathbb{E}(h_m^{(N)}(n + K_{N-n,t})) = \mathbb{E}(h_{m-K_{m,t}}^{(N)}(n))$ by Lemma 3.1. \square

Proof of Theorem 2.8. We essentially follow the proof of Theorem 2.6. The calculations are slightly more involved due to the discrete nature of the involved processes, requiring a replacement of monomials by factorials. The linearity of the generator $L^{X^{(N)}}$ and its action on

the function $h_m^{(N)}$ (see Lemma 2.7) yields

$$\begin{aligned} L^{X^{(N)}} D^{(N)}(., m)(n) &= \frac{L^{X^{(N)}} h_m^{(N)}(n)}{\mu_m^{(N)}} = \frac{1}{\mu_m^{(N)}} \sum_{k=1}^m \varphi(k, m) h_{m-k}^{(N)}(n) - 2\varphi(m) \frac{h_m^{(N)}(n)}{\mu_m^{(N)}} \\ &= \frac{1}{\mu_m^{(N)}} \sum_{k=1}^m \varphi(k, m) \mu_{m-k}^{(N)} D^{(N)}(n, m-k) - 2\varphi(m) D^{(N)}(n, m). \end{aligned}$$

Thanks to the recursion (21) for the $\mu_m^{(N)}$, this expression can be rewritten as

$$\begin{aligned} L^{X^{(N)}} D^{(N)}(., m)(n) &= \frac{1}{\mu_m^{(N)}} \sum_{k=1}^m \varphi(k, m) \mu_{m-k}^{(N)} D^{(N)}(n, m-k) - \frac{1}{\mu_m^{(N)}} \sum_{k=1}^m \varphi(k, m) \mu_{m-k}^{(N)} D^{(N)}(n, m) \\ &= \frac{1}{\mu_m^{(N)}} \sum_{k=1}^m \mu_{m-k}^{(N)} \varphi(k, m) (D^{(N)}(n, m-k) - D^{(N)}(n, m)) = L^{Y^{(N)}} D^{(N)}(n, .)(m), \end{aligned}$$

where $Y^{(N)}$ is a continuous-time death process with state space $\{0, \dots, N\}$ and generator (23). Thus, $L^{X^{(N)}} D^{(N)}(., m)(n) = L^{Y^{(N)}} D^{(N)}(n, .)(m)$ for all $n, m \in \{0, \dots, N\}$ showing that $X^{(N)}$ is dual to $Y^{(N)}$ with respect to $D^{(N)}$. \square

Proof of Theorem 2.9. For any function $f : \mathbb{N}_0 \rightarrow \mathbb{R}$ let the function $\pi_N f : \{0, \dots, N\} \rightarrow \mathbb{R}$ be defined by $\pi_N f(m) := f(m)$ for all $m \in \{0, \dots, N\}$. To prove the convergence $Y^{(N)} \rightarrow Y$ in $D_{\mathbb{N}_0}([0, \infty))$, it suffices by standard theory on weak convergence of Markov processes (see, for example, Ethier and Kurtz [6, p. 28, Theorem 6.1 and p. 232, Corollary 8.7]) to verify that, for every bounded function $f : \mathbb{N}_0 \rightarrow \mathbb{R}$,

$$\|L^{Y^{(N)}}(\pi_N f) - \pi_N(L^Y f)\| := \sup_{m \in \{0, \dots, N\}} |L^{Y^{(N)}}(\pi_N f)(m) - L^Y f(m)| \rightarrow 0, \quad N \rightarrow \infty. \quad (65)$$

In the following it is shown that the left hand side in (65) is even equal to 0 for all $N \in \mathbb{N}$. As observed after Eq. (21), the scaled factorial moment $\mu_m^{(N)}$ coincides with the moment μ_m as long as $m \leq N$. Hence, for all $N \in \mathbb{N}$ and all $m \in \{0, \dots, N\}$, by (23) and (19),

$$L^{Y^{(N)}}(\pi_N f)(m) = \frac{1}{\mu_m} \sum_{k=1}^m \varphi(k, m) \mu_{m-k} (f(m-k) - f(m)) = L^Y f(m)$$

does not depend on $N \in \mathbb{N}$. Thus, the equality $L^{Y^{(N)}}(\pi_N f) = \pi_N(L^Y f)$ holds for all $N \in \mathbb{N}$, showing that the left hand side in (65) is equal to 0 for all $N \in \mathbb{N}$. \square

Proof of Theorem 2.10. Let D be a kernel of the polynomial form (25). Then, from the

linearity of the generator L^X and from (9) it follows that

$$\begin{aligned}
L^X D(\cdot, n)(x) &= \sum_{j=0}^n d_{n,j} L^X f_j(x) = \sum_{j=0}^n d_{n,j} \left(-2\varphi(j)x^j + \sum_{k=1}^j \varphi(k, j)x^{j-k} \right) \\
&= -2 \sum_{j=0}^n d_{n,j} \varphi(j)x^j + \sum_{j=0}^n d_{n,j} \sum_{i=0}^{j-1} \varphi(j-i, j)x^i \\
&= -2 \sum_{i=0}^n d_{n,i} \varphi(i)x^i + \sum_{i=0}^{n-1} \sum_{j=i+1}^n d_{n,j} \varphi(j-i, j)x^i.
\end{aligned}$$

This expression is a polynomial of degree less than or equal to n and the coefficient in front of x^i , $i \in \{0, \dots, n\}$, is

$$-2d_{n,i}\varphi(i) + \sum_{j=i+1}^n d_{n,j}\varphi(j-i, j), \quad i \in \{0, \dots, n\}. \quad (66)$$

On the other hand, by the definition (24) of the generator L^Y ,

$$\begin{aligned}
L^Y D(x, \cdot)(n) &= \sum_{k=1}^n (\varphi(k, n) + \varphi(n-k+1, n))(D(x, n-k) - D(x, n)) \\
&= \sum_{k=1}^n (\varphi(k, n) + \varphi(n-k+1, n)) \left(\sum_{j=0}^{n-k} d_{n-k,j}x^j - \sum_{j=0}^n d_{n,j}x^j \right) \\
&= \sum_{j=0}^{n-1} x^j \sum_{k=1}^{n-j} d_{n-k,j} (\varphi(k, n) + \varphi(n-k+1, n)) \\
&\quad - \sum_{j=0}^n d_{n,j} x^j \underbrace{\sum_{k=1}^n (\varphi(k, n) + \varphi(n-k+1, n))}_{=2\varphi(n)}.
\end{aligned}$$

Again, this expression is a polynomial of degree less than or equal to n and the coefficient in front of x^i , $i \in \{0, \dots, n\}$, is

$$\sum_{k=1}^{n-i} d_{n-k,i} (\varphi(k, n) + \varphi(n-k+1, n)) - 2d_{n,i}\varphi(n). \quad (67)$$

We would like to choose the coefficients $d_{n,i}$ in such a way that (66) and (67) coincide, since then we will have the equality $L^X D(\cdot, n)(x) = L^Y D(x, \cdot)(n)$, which is the desired duality result. Putting (66) equal to (67) yields the equation

$$2d_{n,i}\varphi(i) = \sum_{k=1}^{n-i} d_{n-k,i} (\varphi(k, n) + \varphi(n-k+1, n)) - \sum_{j=i+1}^n d_{n,j}\varphi(j-i, j). \quad (68)$$

Now note that (68) always holds for $i = n$ no matter how $d_{n,n}$ is chosen. Thus, one can choose the diagonal coefficients $d_{n,n}$, $n \in \mathbb{N}_0$, arbitrarily. Once these diagonal coefficients are chosen, say as $d_{n,n} := r_n$ for all $n \in \mathbb{N}_0$, Eq. (68) is a recursion forwards over n and for each n backwards over $i = n - 1, n - 2, \dots, 1, 0$, showing that all the off-diagonal coefficients $d_{n,i}$, $n \in \mathbb{N}$, $i \in \{0, \dots, n - 1\}$, are uniquely determined by the diagonal sequence $r = (r_n)_{n \in \mathbb{N}_0}$. \square

Proof of Proposition 2.11. (i) “ \Rightarrow ”: Assume first that $\mu = \beta(p, q)$, i.e. the stationary distribution of the $\beta(a, b)$ -mutation process X is a beta distribution for some parameters $p, q > 0$. We have to verify that $a \in \{1, 2\}$ or $a = b$. From the recursion (14) it follows that the first four moments of μ are

$$\begin{aligned} \mu_1 &= \frac{1}{2}, \quad \mu_2 = \frac{a+b}{2(a+2b)}, \quad \mu_3 = \frac{2a+b}{4(a+2b)} \quad \text{and} \\ \mu_4 &= \frac{(a+b+2)(a^3+a^2+3a^2b+4ab+4ab^2+b^2+b^3)}{2(a+2b)(a+2b+2)(a^2+2ab+a+4b+2b^2)}. \end{aligned} \tag{69}$$

Note that the formula for μ_4 is a rather involved expression in a and b . Comparing $\mu_1 = \frac{1}{2}$ with the mean $\frac{p}{p+q}$ of the $\beta(p, q)$ -distribution it follows that $p = q$. The symmetric beta distribution $\beta(p, p)$ has second moment $\frac{p(p+1)}{2p(2p+1)} = \frac{p+1}{2(2p+1)}$. Since this expression is equal to μ_2 it follows that $p = b/a$. For this p , even the third moment $\frac{p(p+1)(p+2)}{2p(2p+1)(2p+2)} = \frac{p+2}{4(2p+1)} = \frac{2a+b}{4(a+2b)}$ of the $\beta(p, p)$ -distribution coincides with μ_3 . However, the fourth moment $\frac{p(p+1)(p+2)(p+3)}{2p(2p+1)(2p+2)(2p+3)} = \frac{(p+2)(p+3)}{4(2p+1)(2p+3)} = \frac{(2a+b)(3a+b)}{4(a+2b)(3a+2b)}$ only coincides with μ_4 in particular cases. A technical but straightforward calculation shows that these cases are $a \in \{1, 2\}$ or $a = b$.

“ \Leftarrow ”: We have to show that, for $a \in \{1, 2\}$ or $a = b$, the stationary distribution is the $\beta(b/a, b/a)$ -distribution. To see this it suffices to verify that the moments of the beta distribution $\mu := \beta(b/a, b/a)$ satisfies the recursion (14). We do this separately for the three cases $a = 1$, $a = 2$ or $a = b$.

Case 1: Assume that $a = 1$. In the following, instead of verify the recursion (14), it is even shown that $\mu := \beta(b, b)$ is reversible, i.e. for all functions $f, g \in L_2(\mu)$ it is shown that

$$\langle L^X f, g \rangle = \langle f, L^X g \rangle, \tag{70}$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product in $L_2(\mu)$. Recalling the definition of L^X we have $\langle L^X f, g \rangle = I_1 + I_2$, where

$$I_1 := \int_0^1 \left(\int_0^1 (f(x(1-t)) - f(x)) \frac{1}{t} (1-t)^{b-1} dt \right) g(x) \frac{x^{b-1} (1-x)^{b-1}}{B(b, b)} dx$$

and

$$I_2 := \int_0^1 \left(\int_0^1 (f(x(1-t)+t) - f(x)) \frac{1}{t} (1-t)^{b-1} dt \right) g(x) \frac{x^{b-1} (1-x)^{b-1}}{B(b, b)} dx.$$

The change of variables $(x, t) \mapsto (x, y)$ with $y := x(1-t)$ yields

$$I_1 = \frac{1}{B(b, b)} \iint_{0 < y < x < 1} (f(y) - f(x)) g(x) \frac{y^{b-1} (1-x)^{b-1}}{x-y} dx dy$$

and changing variables $(x, t) \mapsto (x, y)$ with $y := x(1 - t) + t$ gives

$$I_2 = \frac{1}{\mathbf{B}(b, b)} \iint_{0 < x < y < 1} (f(y) - f(x))g(x) \frac{(1-y)^{b-1}x^{b-1}}{y-x} dx dy.$$

In a similar manner, one shows that $\langle f, L^X g \rangle = I_2 + I_1$. As a consequence (70) holds.

Case 2: Assume that $a = 2$. The beta distribution $\beta(\frac{b}{2}, \frac{b}{2})$ has moments $m_n := \frac{\Gamma(b)\Gamma(n+\frac{b}{2})}{\Gamma(\frac{b}{2})\Gamma(n+b)}$. Therefore, using (27),

$$\begin{aligned} \sum_{k=1}^n m_{n-k} \varphi_{2,b}(k, n) &= \sum_{k=1}^n \frac{\Gamma(b)\Gamma(n-k+\frac{b}{2})}{\Gamma(\frac{b}{2})\Gamma(n-k+b)} \binom{n}{k} \frac{1}{\mathbf{B}(2, b)} \frac{\Gamma(k+1)\Gamma(n-k+b)}{\Gamma(n+b+1)} \\ &= \frac{1}{\mathbf{B}(2, b)} \frac{\Gamma(b)}{\Gamma(\frac{b}{2})} \frac{n!}{\Gamma(n+b+1)} \sum_{k=1}^n \frac{\Gamma(n-k+\frac{b}{2})}{(n-k)!}. \end{aligned}$$

The last sum simplifies to $\sum_{i=0}^{n-1} \frac{\Gamma(i+\frac{b}{2})}{i!} = \frac{\Gamma(n+\frac{b}{2})}{\frac{b}{2}\Gamma(n)}$, which is easily verified by induction on $n \in \mathbb{N}$. Thus,

$$\begin{aligned} \sum_{k=1}^n m_{n-k} \varphi_{2,b}(k, n) &= \frac{1}{\mathbf{B}(2, b)} \frac{\Gamma(b)}{\Gamma(\frac{b}{2})} \frac{n!}{\Gamma(n+b+1)} \frac{\Gamma(n+\frac{b}{2})}{\frac{b}{2}\Gamma(n)} \\ &= \frac{1}{\mathbf{B}(2, b)} 2 \frac{\Gamma(b)\Gamma(n+\frac{b}{2})}{\Gamma(\frac{b}{2})\Gamma(n+b)} \frac{n}{b(n+b)} = 2m_n \varphi_{2,b}(n), \end{aligned}$$

since $\varphi_{2,b}(n) = \frac{1}{\mathbf{B}(2,b)} \frac{n}{b(n+b)}$ by (28). Thus, the moments m_n satisfy the recursion (14).

Case 3: For the case $a = b > 0$, the moments $m_n := 1/(n+1)$ of the uniform distribution satisfy

$$\begin{aligned} \sum_{k=1}^n m_{n-k} \varphi_{a,a}(k, n) &= \sum_{k=1}^n \frac{1}{n-k+1} \binom{n}{k} \frac{1}{\mathbf{B}(a, a)} \int_0^1 t^{k+a-2} (1-t)^{n-k+a-1} dt \\ &= \frac{1}{n+1} \frac{1}{\mathbf{B}(a, a)} \int_0^1 t^{a-2} (1-t)^{a-2} \sum_{k=1}^n \binom{n+1}{k} t^k (1-t)^{n+1-k} dt \\ &= m_n \frac{1}{\mathbf{B}(a, a)} \int_0^1 t^{a-2} (1-t)^{a-2} (1 - (1-t)^{n+1} - t^{n+1}) dt. \end{aligned}$$

For $a \neq 1$, the last integral is equal to $\mathbf{B}(a-1, a-1) - \mathbf{B}(a-1, n+a) - \mathbf{B}(n+a, a-1) = 2\mathbf{B}(a-1, a) - 2\mathbf{B}(a-1, n+a) = 2\mathbf{B}(a, a)\varphi_{a,a}(n)$, since $\mathbf{B}(a-1, a-1) = 2\mathbf{B}(a-1, a)$, and for $a = 1$ this integral is equal to

$$\begin{aligned} \int_0^1 \frac{1 - (1-t)^{n+1} - t^{n+1}}{t(1-t)} dt &= \int_0^1 \left(\frac{1-t^n}{1-t} + \frac{1-(1-t)^n}{t} \right) dt \\ &= 2(\Psi(n+1) - \Psi(1)) = 2\varphi_{1,1}(n). \end{aligned}$$

Thus, again, the recursion (14) holds, which shows that the uniform distribution is the stationary distribution for the case $a = b$.

(ii) If $a = 1$ then it was already verified in (i) that the detailed balance condition holds. Thus, for $a = 1$ the stationary distribution is reversible.

Conversely, assume that the stationary distribution μ is reversible. We have to verify that $a = 1$. Since μ is reversible, we have $\langle L^X f_n, f_m \rangle = \langle f_n, L^X f_m \rangle$ for all $n, m \in \mathbb{N}_0$. Using the action (9) of L^X on monomials it is readily seen that these equations are equivalent to

$$\sum_{k=1}^n \varphi(k, n) \mu_{n+m-k} - 2\varphi(n) \mu_{n+m} = \sum_{k=1}^m \varphi(k, m) \mu_{n+m-k} - 2\varphi(m) \mu_{n+m}, \quad n, m \in \mathbb{N}_0, \quad (71)$$

where $\mu_n := \int f_n d\mu$ denotes the n th moment of μ , $n \in \mathbb{N}_0$. For $m = 0$, the right hand side in (71) vanishes, leading to the recursion (14) for the moments of μ . In particular, the first four moments of μ are given by (69). The obvious idea of obtaining the desired result $a = 1$ by exploiting (71) for $(m, n) = (1, 2)$ fails, since it turns out that (71) is satisfied for all $a, b > 0$ in this case. The next attempt is hence to consider $(m, n) = (1, 3)$, in which case (71) turns into

$$\begin{aligned} 0 &= (\varphi(1, 3) - \varphi(1, 1))\mu_3 + \varphi(2, 3)\mu_2 + \varphi(3, 3)\mu_1 + 2(\varphi(1) - \varphi(3))\mu_4 \\ &= \left(\frac{3b(b+1)}{(a+b)(a+b+1)} - 1 \right) \frac{2a+b}{4(a+2b)} \\ &\quad + \frac{3ab}{(a+b)(a+b+1)} \frac{a+b}{2(a+2b)} + \frac{a(a+1)}{(a+b)(a+b+1)} \frac{1}{2} \\ &\quad + 2 \left(1 - \frac{a^2 + 3ab + a + 3b^2 + 3b}{(a+b)(a+b+1)} \right) \frac{(a+b+2)(a^3 + a^2 + 3a^2b + 4ab + 4ab^2 + b^2 + b^3)}{2(a+2b)(a+2b+2)(a^2 + 2ab + a + 4b + 2b^2)} \\ &= \frac{ab(1-a^2)}{4(a^2 + 2ab + a + 4b + 2b^2)(a+b)(a+b+1)}, \end{aligned}$$

where the last equality follows via tedious but straightforward calculus. The latter fraction is equal to 0 only for $a = 1$. The proof is complete. \square

Proof of Lemma 2.12. As in Handa [16], define $G_t(x) := 1/(1+tx)$. A straightforward calculation shows that

$$G_t(x(1-u)) + G_t(x(1-u)+u) - 2G_t(x) = -\frac{tu}{1+tx} \left(\frac{1-x}{1+tx(1-u)+tu} - \frac{x}{1+tx(1-u)} \right).$$

Therefore,

$$\begin{aligned} L^X G_t(x) &= \int_0^1 \frac{u^{a-1}(1-u)^{b-1}}{u} (G_t(x(1-u)) + G_t(x(1-u)+u) - 2G_t(x)) du \\ &= -\int_0^1 u^{a-1}(1-u)^{b-1} \frac{t}{1+tx} \left(\frac{1-x}{1+tx(1-u)+tu} - \frac{x}{1+tx(1-u)} \right) du \\ &= -\frac{t(1-x)}{1+tx} \int_0^1 \frac{u^{a-1}(1-u)^{b-1}}{1+tx(1-u)+tu} du + \frac{tx}{1+tx} \int_0^1 \frac{u^{a-1}(1-u)^{b-1}}{1+tx(1-u)} du. \end{aligned}$$

Since $a + b = 1$, both integrals are known explicitly (Handa [16, Eq. (2.4)]), which yields

$$\begin{aligned} L^X G_t(x) &= -\frac{t(1-x)}{1+tx} \mathbf{B}(a, b)(1+t)^{-a}(1+tx)^{-b} + \frac{tx}{1+tx} \mathbf{B}(a, b)(1+tx)^{-b} \\ &= -\mathbf{B}(a, b)t \left(\frac{1-x}{(1+t)^{1-b}(1+tx)^{b+1}} - \frac{x}{(1+tx)^{b+1}} \right). \end{aligned}$$

For all $t \geq 0$, $\int L^X G_t(x) \mu(dx) = 0$, which gives

$$\int \frac{x}{(1+tx)^{b+1}} \mu(dx) = \frac{1}{(1+t)^{1-b}} \int \frac{1-x}{(1+tx)^{b+1}} \mu(dx).$$

The left and right integrals are equal to $-S'_b(t)/b$ and $\frac{1+t}{b} S'_b(t) + S_b(t)$ respectively, which yields the differential equation $((1+t)^{1-b} + 1+t) S'_b(t) = -b S_b(t)$ and, hence, via integration

$$-\log S_b(t) = \int_0^t \frac{b}{(1+u)^{1-b} + 1+u} du, \quad t \geq 0.$$

An antiderivative function is $F(u) := \log(1+(1+u)^b)$, which yields $-\log S_b(t) = F(t) - F(0) = \log(1+(1+t)^b) - \log 2 = \log((1+(1+t)^b)/2)$ and, hence, the solution (29). \square

Proof of Lemma 2.13. The substitution $y := 1/x$ yields

$$S_b(z) = \int_0^1 (1+zx)^{-b} f_\mu(x) dx = \int_1^\infty (1+z/y)^{-b} f_\mu(1/y) y^{-2} dy = \int_1^\infty (y+z)^{-b} F(y) dy,$$

where $F(y) := y^{b-2} f_\mu(1/y)$ for all $y > 1$. By inversion (see, for example, Schwarz [26, Eq. (9)]),

$$F(y) := -\frac{1}{2\pi i} y^b \int_C (1+w)^{b-1} S'_b(yw) dw, \quad y \in (1, \infty),$$

where C is a contour that starts and ends at the point $w = -1$ and encloses the origin. Multiplication with y^{2-b} yields

$$f_\mu(1/y) = y^{2-b} F(y) = -\frac{1}{2\pi i} y^2 \int_C (1+w)^{b-1} S'_b(yw) dw, \quad y \in (1, \infty),$$

and back substitution $y := 1/x$ leads to

$$f_\mu(x) = -\frac{1}{2\pi i} x^{-2} \int_C (1+w)^{b-1} S'_b(w/x) dw, \quad x \in (0, 1).$$

The result now follows from $S'_b(z) = -\frac{2b(1+z)^{b-1}}{(1+(1+z)^b)^2}$. \square

Proof of Proposition 2.14. The expansion of the generalized Stieltjes transform $S_b(t)$ in pow-

ers of t is obtained via

$$\begin{aligned}
S_b(t) &= \frac{2}{1 + (1+t)^b} = \sum_{k=0}^{\infty} \left(\frac{1 - (1+t)^b}{2} \right)^k \\
&= 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{2^k} ((1+t)^b - 1)^k = 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{2^k} \left(\sum_{i=1}^{\infty} \binom{b}{i} t^i \right)^k \\
&= 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{2^k} \sum_{i_1, \dots, i_k \in \mathbb{N}} \binom{b}{i_1} \cdots \binom{b}{i_k} t^{i_1 + \dots + i_k} \\
&= 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{2^k} \sum_{n=k}^{\infty} t^n \sum_{\substack{i_1, \dots, i_k \in \mathbb{N} \\ i_1 + \dots + i_k = n}} \binom{b}{i_1} \cdots \binom{b}{i_k} \\
&= 1 + \sum_{n=1}^{\infty} t^n \sum_{k=1}^n \frac{(-1)^k}{2^k} \sum_{\substack{i_1, \dots, i_k \in \mathbb{N} \\ i_1 + \dots + i_k = n}} \binom{b}{i_1} \cdots \binom{b}{i_k}.
\end{aligned}$$

On the other hand, $S_b(t) = \int (1+tx)^{-b} \mu(dx) = \int \sum_{n=0}^{\infty} \binom{-b}{n} (tx)^n \mu(dx) = \sum_{n=0}^{\infty} t^n \binom{-b}{n} \mu_n$. Comparing the coefficient in front of t^n yields the solution

$$\mu_n = \frac{1}{\binom{-b}{n}} \sum_{k=1}^n \frac{(-1)^k}{2^k} \sum_{\substack{i_1, \dots, i_k \in \mathbb{N} \\ i_1 + \dots + i_k = n}} \binom{b}{i_1} \cdots \binom{b}{i_k}, \quad n \in \mathbb{N}.$$

The last multi sum is expressed by a single sum (see, for example, [22, Eq. (13)] as

$$\sum_{\substack{i_1, \dots, i_k \in \mathbb{N} \\ i_1 + \dots + i_k = n}} \binom{b}{i_1} \cdots \binom{b}{i_k} = \sum_{\ell=1}^k (-1)^{k-\ell} \binom{k}{\ell} \binom{b\ell}{n}, \quad n, k \in \mathbb{N}, \quad (72)$$

which yields the first formula in (31). Using well-known formulas for the standard Stirling numbers $s(\cdot, \cdot)$ and $S(\cdot, \cdot)$ of the first and second kind respectively we obtain

$$\begin{aligned}
\sum_{\ell=1}^k (-1)^{k-\ell} \binom{k}{\ell} \binom{b\ell}{n} &= \frac{1}{n!} \sum_{\ell=1}^k (-1)^{k-\ell} \binom{k}{\ell} (b\ell)_n = \frac{1}{n!} \sum_{\ell=1}^k (-1)^{k-\ell} \binom{k}{\ell} \sum_{j=1}^n s(n, j) (b\ell)^j \\
&= \frac{1}{n!} \sum_{j=1}^n s(n, j) b^j \sum_{\ell=1}^k (-1)^{k-\ell} \binom{k}{\ell} \ell^j = \frac{k!}{n!} \sum_{j=1}^n s(n, j) b^j S(j, k),
\end{aligned}$$

which yields the second formula in (31). \square

Acknowledgements. The authors would like to thank Frank Redig for many enlightening discussions, in particular for his suggestion of the dual process when the measure Λ is the uniform distribution. The authors also thank the University of Tübingen and the Modena and Reggio Emilia University for supporting two one-month research visits where parts of this work were conducted. The whole probability group at the University of Tübingen and the mathematical statistical physics group at the Modena and Reggio Emilia University are warmly acknowledged.

References

- [1] ALDOUS, D. AND LANOUE, D. (2012) A lecture on the averaging process. *Probab. Surv.* **9**, 90–102 MR2908618
- [2] CANNINGS, C. (1974) The latent roots of certain Markov chains arising in genetics: a new approach, I. Haploid models. *Advances in Appl. Probability* **6**, 260–290. MR0343949
- [3] CARINCI, G., FRANCESCHINI, C., FRASSEK, R., GIARDINÀ, C. AND REDIG, F. (2025) Large deviations and additivity principle for the open harmonic process. *Comm. Math. Phys.* **406**(5), Paper No. 103, 40pp. MR4887622
- [4] CARINCI, G., FRANCESCHINI, C., GABRIELLI, D., GIARDINÀ, C. AND TSAGKARO-GIANNIS, D. (2025) Solvable stationary non equilibrium states. *J. Stat. Phys.* **191**(1), Paper No. 10, 14pp. MR4691917
- [5] ESTRADA, R. AND KANWAL, R. P. (1987) The Carleman type singular integral equations. *SIAM Review* **29**(2), 263–290. MR number not available
- [6] ETHIER, S. N. AND KURTZ, T. G. (2005) *Markov Processes, Characterization and Convergence*. Wiley. MR0838085
- [7] FRANCESCHINI, C., FRASSEK, R. AND GIARDINÀ, C. (2023) Integrable heat conduction model. *J. Math. Phys.* **64**(4), Paper No. 043304, 21 pp. MR4575353
- [8] FRASSEK, R. AND GIARDINÀ, C. (2022) Exact solution of an integrable non-equilibrium particle system. *J. Math. Phys.* **63**(10), Paper No. 103301, 35pp. MR4493574
- [9] FRASSEK, R., GIARDINÀ, C. AND KURCHAN, J. (2020) Non-compact quantum spin chains as integrable stochastic particle processes. *J. Stat. Phys.* **180**(1–6), 135–171. MR4130985
- [10] GIARDINÀ, C., REDIG, F. AND VAN TOL, B. (2025) Intertwining and propagation of mixtures for generalized KMP models and harmonic models. *J. Stat. Phys.* **192**(2), Paper No. 21, 33pp. MR4856237
- [11] GLADSTIEN, K. (1976) Loss of alleles in a haploid population with varying environment. *Theoret. Population Biol.* **10**, 383–394. MR0496833
- [12] GLADSTIEN, K. (1977) Haploid populations subject to varying environment: the characteristic values and the rate of loss of alleles. *SIAM J. Appl. Math.* **32**, 778–783. MR0434496
- [13] GLADSTIEN, K. (1978) The characteristic values and vectors for a class of stochastic matrices arising in genetics. *SIAM J. Appl. Math.* **34**(4), 630–642. MR0475977
- [14] GRIFFITHS, R. C. (2014) The Λ -Fleming–Viot process and a connection with Wright–Fisher diffusion. *Adv. in Appl. Probab.* **46**(4), 1009–1035. MR3290427
- [15] GONZÁLEZ CASANOVA, A. AND SPANÓ, D. (2018) Duality and fixation in Ξ -Wright–Fisher processes with frequency-dependent selection. *Ann. Appl. Probab.* **28**(1), 250–284. MR3770877

- [16] HANDA, K. (2014) Stationary distributions for a class of generalized Fleming–Viot processes. *Ann. Probab.* **42**(3), 1257–1284. MR3189071
- [17] HSU, L. C. AND SHIUE, P. J.-S. (1998) A unified approach to generalized Stirling numbers. *Adv. in Appl. Math.* **20**(3), 366–384. MR1618435
- [18] KIM, S., QUATTROPANI, M. AND SAU, F. (2025) Spectral gap of the KMP and other stochastic exchange models on arbitrary graphs. *arXiv:2505.02400*
- [19] KIPNIS, C., MARCHIORO, C. AND PRESUTTI, E. (1982) Heat flow in an exactly solvable model. *J. Stat. Phys.* **27**(1), 65–74. MR0656869
- [20] MÖHLE, M. (1999) The concept of duality and applications to Markov processes arising in neutral population genetics models. *Bernoulli* **5**(5), 761–777. MR1724816
- [21] MÖHLE, M. (2006) On sampling distributions for coalescent processes with simultaneous multiple collisions. *Bernoulli* **12**(1), 35–53. MR2202319
- [22] MÖHLE, M. (2010) Looking forwards and backwards in the multi-allelic neutral Cannings population model. *J. Appl. Probab.* **47**(3), 713–731. MR2731344
- [23] PITMAN, J. (1999) Coalescents with multiple collisions. *Ann. Probab.* **27**(4), 1870–1902. MR1742892
- [24] PÓLYA, G. (1920) Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentenproblem. *Math. Z.* **8**, 171–181. MR15444437
- [25] REDIG, F. AND VAN TOL, B. (2026) Ergodic properties of the harmonic process. *arXiv:2604.17469*
- [26] SCHWARZ, J. H. (2005) The generalized Stieltjes transform and its inverse. *J. Math. Phys.* **46**(1), 013501, 8pp. MR2113770