Non-local Solvable Birth–Death Processes

Giacomo Ascione[1] · Nikolai Leonenko[2] · Enrica Pirozzi[1]

Abstract

In this paper, we study strong solutions of some non-local difference–differential equations linked to a class of birth–death processes arising as discrete approximations of Pearson diffusions by means of a spectral decomposition in terms of orthogonal polynomials and eigenfunctions of some non-local derivatives. Moreover, we give a stochastic representation of such solutions in terms of time-changed birth–death processes and study their invariant and their limit distribution. Finally, we describe the correlation structure of the aforementioned time-changed birth–death processes.

Keywords Subordinator · Bernstein function · Classical orthogonal polynomial of discrete variable

Mathematics Subject Classification (2020) 60K15 · 33C45 · 60G22

1 Introduction

Birth–death processes constitute a fundamental class of continuous-time Markov chain that are widely used in applications such as, for instance, evolutionary dynamics [39]

✉ Enrica Pirozzi
  enrica.pirozzi@unina.it

  Giacomo Ascione
  giacomo.ascione@unina.it

  Nikolai Leonenko
  leonenkon@cardiff.ac.uk

1   Dipartimento di Matematica e Applicazioni "Renato Caccioppoli", Università degli Studi di Napoli Federico II, 80126 Napoli, Italy

2   School of Mathematics, Cardiff University, Cardiff CF24 4AG, UK

and queueing theory [26]. In particular, one can achieve a complete characterization of birth–death processes by families of classical orthogonal polynomials of discrete variable. This theory, linked to the solution of the Stieltjes moment problem, has been widely studied by Karlin and McGregor in their seminal papers [22,25].

As birth–death processes are linked to difference–differential equations, *fractionalization* of such processes can be used to study the solutions of fractional difference–differential equations. Indeed, with this idea in mind, a fractional version of the Poisson process has been introduced in [10,11] and fractional versions of some birth–death processes, for instance, in [40–42].

In the case of Pearson diffusions, one can use a spectral approach to study strong solutions of fractional backward and forward Kolmogorov equations and, at the same time, define the fractional Pearson diffusions by means of a time-change via an inverse stable subordinator (see, for instance, [31–33]). The same approach has been used to study the case of fractional immigration-death processes in [8]. Let us also stress out that this approach can be used to study a fractional $M/M/\infty$ queue or a fractional $M/M/1$ queue with acceleration of service (for some models of fractional queues, we refer to [6,7,17]).

However, one could consider a time-change with a different inverse subordinator. In such case, in place of the fractional derivative in time, one obtains a more general non-local operator. Such kind of operators have been introduced in [27] for the class of complete Bernstein functions and extended in [49] for any Bernstein function. A first step towards the theory of general time-changed Pearson diffusions has been achieved in [20].

In this work, we describe a general theory for non-local solvable birth–death processes in terms of orthogonal polynomials, where such processes are defined by means of a time-change with a general inverse subordinator. In particular, we focus on the strong solutions of general non-local backward and forward Kolmogorov equations associated to such processes and on the stochastic representation of such solutions. In particular, the paper is structured as follows:

- In Sect. 2 we introduce the theory of solvable birth–death processes as discrete approximations of Pearson diffusions and we state the main hypotheses we need on the starting birth–death process;
- In Sect. 3 we give some preliminaries on inverse subordinators and non-local time derivatives. In particular we focus on the eigenvalue equation for such derivatives and on some upper bounds for the eigenfunctions. Let us stress out that some properties of such eigenfunctions are expressed in [27,28] in the complete Bernstein case, in [36] in the special case and in [4] in the general case. Moreover, a series expansion in terms of convolutions of potential densities in the special Bernstein case is obtained in [4];
- In Sect. 4 we focus on the spectral decomposition of the strong solutions of non-local forward and backward Kolmogorov equations in terms of orthogonal polynomials of discrete variable and eigenfunctions of the non-local time derivatives;
- In Sect. 5 we introduce the time-changed birth–death processes and we study the stochastic representation of the aforementioned strong solutions in terms of such

processes. In particular we obtain that the time-changed process still admits the same invariant measure that is also the limit measure for any starting distribution;

– Finally, in Sect. 6 we study the correlation structure of the time-changed birth–death processes in terms of the potential measure of the involved subordinator and the eigenfunctions of the non-local time derivatives. In particular, the non-stationarity of the process is evident in the expression of the covariance, thus, to give some information on the memory of the process, we have to refer to a non-stationary extension of the definition of long-range and short-range dependence suggested by the necessary conditions given in [12, Lemmas 2.1 and 2.2].

2 Solvable Birth–Death Processes

Let us fix a filtered space $(\Omega, \{\mathcal{F}_t\}_{t \in \mathbb{R}^+}, \mathbb{P})$ and consider a Birth–Death process $\{N(t), t \geq 0\}$ on it. Let us denote by $E \subseteq \mathbb{Z}$ its state space, that will be finite or at most countable. In particular we can always suppose that $E \subseteq \mathbb{N}_0$ and is a segment, i.e. for any $n_1, n_2 \in E$ and $n \in \mathbb{N}_0$ such that $n_1 \leq n \leq n_2$ it holds $n \in E$, with $\min E = 0$. For a Birth–Death process $N(t)$ we can consider the semigroup $T_t : f \in \mathfrak{b} \mapsto T_t f(y) = \mathbb{E}_y[f(N(t))]$ where $\mathfrak{b}$ is a suitable Banach sequence space containing c_0 (i.e. the space of infinitesimal sequences) whenever $E = \mathbb{N}_0$, while $\mathfrak{b} = \mathbb{R}^{|E|}$ if E is finite. Since we will consider Birth–Death processes with a stationary measure $\mathbf{m}$, we will always set $\mathfrak{b} = \ell^2(\mathbf{m})$ (which is equivalent to $\mathbb{R}^{|E|}$ if E is finite). Let us recall that the generator $\mathcal{G}$ of a Birth–Death process can be always expressed as

$$\mathcal{G}f(x) = (b(x) - d(x))\nabla^+ f(x) + d(x)\Delta f(x), \qquad x \in E,$$

where $d(x)$ are the death rates, $b(x)$ are the birth rates (recalling that d, b must be non-negative in E), $\nabla^\pm$ are the first-order forward and backward finite differences defined as

$$\nabla^+ f(x) = f(x+1) - f(x) \qquad \nabla^- f(x) = f(x) - f(x-1)$$

and Δ is the second-order central finite difference

$$\Delta f(x) = f(x+1) - 2f(x) + f(x-1) = \nabla^- \nabla^+ f(x).$$

Let us stress out that $\mathcal{G}$ is actually a tridiagonal matrix, possibly infinite if $E = \mathbb{N}_0$. Moreover, as T_t is defined on $\ell^2(\mathbf{m})$, also the generator $\mathcal{G}$ admits $\ell^2(\mathbf{m})$ as domain.

Here we want to introduce some birth–death version of Pearson diffusions (see, for instance, [32]). To do this, we refer to the theory of birth–death polynomials, whose main papers are [22,25].

Definition 1 We say the process $N(t)$ is solvable if

– $N(t)$ is irreducible and recurrent;
– the spectrum of $\mathcal{G}$ is purely discrete with non-positive eigenvalues $(\lambda_n)_{n \in E}$ such that $\lambda_0 = 0$ and $\lambda_n < 0$ for any $n \geq 1$;

– its eigenfunctions $(P_n)_{n \in E}$ are classical orthogonal polynomials of discrete variable with orthogonality measure $\mathbf{m}$ which is the invariant and stationary measure of $N(t)$;

– the function $m(x) = \mathbf{m}(\{x\})$ solves the following discrete Pearson equation:

$$\nabla^+(d(\cdot)m(\cdot))(x) = (b(x) - d(x))m(x) \qquad (1)$$

– $d(\cdot)$ is a polynomial of degree at most 2 and $b(\cdot) - d(\cdot)$ is a polynomial of degree at most 1.

Concerning solvable birth–death processes, they arise as lattice approximations of Pearson diffusions. In particular, one has in such case

$$\lambda_n = n\nabla^+(b(\cdot) - d(\cdot))(x) + \frac{1}{2}n(n-1)\Delta d(x).$$

Concerning classical orthogonal polynomials of discrete variable, we mainly refer to [38,46]. Their orthogonality relation is expressed as

$$\sum_{x \in E} P_n(x)P_m(x)m(x) = \eth_n^2 \delta_{n,m}, \qquad n, m \in E$$

where $\delta_{n,m}$ is Kronecker delta symbol. In particular one obtains that $\|P_n\|_{\ell^2(\mathbf{m})} = \eth_n$ and then we can introduce the normalized polynomials as $Q_n(x) = \frac{P_n(x)}{\eth_n}$, such that

$$\sum_{x \in E} Q_n(x)Q_m(x)m(x) = \delta_{n,m}, \qquad n, m \in E.$$

On the other hand, the function $\widetilde{m}(n) = \frac{1}{\eth_n^2}$ defines a measure on E. Thus, by proceeding with a Gram-Schmidt orthogonalization procedure on the monomials $(1, x, x^2, \cdots)$, we can define a family of orthogonal polynomials $\widetilde{P}_n(x)$ that satisfies the following orthogonality condition

$$\sum_{x \in E} \widetilde{P}_n(x)\widetilde{P}_m(x)\widetilde{m}(x) = \frac{1}{m(n)}\delta_{n,m}, \qquad n, m \in E.$$

The family of polynomials $(\widetilde{P}_n)_{n \in E}$ is called the dual family of $(P_n)_{n \in E}$, see [38]. Let us give some examples:

– Immigration-death processes (see [2]) are defined by a constant birth rate b and a linear death rate $d(x) = dx$. In such case the invariant measure is given by the Poisson distribution $m(x) = e^{-\rho}\frac{\rho^x}{x!}$ for $x \in \mathbb{N}$, where $\rho = \frac{b}{d}$. The orthogonal polynomials are Charlier polynomials of parameter ρ, that we denote by $P_n(x) = C_n(x; \rho)$ and they satisfy the self-duality relation

$$P_n(x) = P_x(n) \qquad (2)$$

and the eigenvalues are given by $\lambda_n = -bn$. In particular the polynomials P_n coincide with the family of dual orthogonal polynomials $\widetilde{P}_n$. This kind of process arises as a lattice approximation of the Ornstein-Uhlenbeck process.

– Let us consider a linear death rate $d(x) = dx$ and a linear birth rate $b(x) = (x + \beta)b$, with $b, d, \beta > 0$ and $b < d$. In such case the birth death process admits state space $E = \mathbb{N}$ and the generator is given by

$$\mathcal{G} = (\beta b + (b - d)x)\nabla^+ + dx\Delta.$$

Defining $\rho = \frac{b}{d}$, we have that the orthogonal polynomials are Meixner polynomials of parameters ρ and β, that we denote by $P_n(x) = M_n(x; \rho, \beta)$ and they are orthogonal with respect to the invariant measure

$$m(x) = \frac{(\beta)_x \rho^x}{x!(1 - \rho)^\beta},$$

where $(\beta)_x = \frac{\Gamma(\beta+x)}{\Gamma(\beta)}$, which is a Pascal (or negative binomial) distribution of parameters β and ρ. Also Meixner polynomials satisfy the self-duality relation (2) and coincide with their dual polynomials. Finally, let us observe that the eigenvalues are given by $\lambda_n = -(d - b)n$. This process is called the Meixner process and arises as lattice approximation of the Cox-Ingersoll-Ross process. Meixner processes are discussed for instance in [23].

– Another case is given by a linear death rate $d(x) = dx$ and a linear decreasing birth rate $b(x) = (N - x)b$ with $b, d > 0$ and $N \in \mathbb{N}$. In such case the birth–death process admits finite state space $E = \{0, \ldots, N\}$ and the generator is given by

$$\mathcal{G} = (Nb - (b + d)x)\nabla^+ + dx\Delta.$$

Defining $p = \frac{b}{b+d}$ and $q = 1 - p$, we achieve the invariant distribution given by

$$m(x) = \binom{N}{x} p^x q^{N-x}$$

which is a Binomial distribution over E. The orthogonal polynomials are Krawtchouk polynomials of parameters N and p, that we denote by $P_n(x) = K_n(x; N, p)$ and satisfy the self-duality relation (2). The eigenvalues of the generator are given by $\lambda_n = -n(b+d)$. Let us recall that this is actually a time-continuous version of the Ehrenfest urn model (see, for instance, [24]).

– Another interesting case is given by a quadratic one. Indeed let us consider for some $d > 0$ and $\alpha, \beta, N \in \mathbb{N}$

$$d(x) = dx(N + \beta + 1 - x) \quad b(x) = d[N(\alpha + 1) + x(N - 1 - \alpha) - x^2].$$

Let us observe that $b(N) = 0$, thus the state space of the process is given by $E = \{0, \ldots, N\}$. Moreover, its generator is given by

$$\mathcal{G} = d(N(\alpha + 1) - (\beta + \alpha + 2)x)\nabla^+ + dx(N + \beta + 1 - x)\Delta$$

with eigenvalues
$$\lambda_n = -dn[n + 1 + \alpha + \beta].$$

The invariant measure of this birth–death process is an hypergeometric distribution on E given by

$$m(x) = \binom{\alpha + x}{x}\binom{\beta + N - x}{N - x}$$

and the orthogonal polynomials are the Hahn polynomials, that we denote by $P_n(x) = H_n(x; \alpha, \beta, N)$. In this case we do not have self-duality relation, but the family of dual Hahn polynomials, that we denote by $\widetilde{P}_n(x) = R_n(x; \alpha, \beta, N)$, is linked to the Hahn polynomials by the relation

$$H_n(x; \alpha, \beta, N) = R_x(n(n + \alpha + \beta + 1); \alpha, \beta, N).$$

This particular birth–death process is a lattice approximation of the Jacobi process. For such process, we refer directly to [46].

In the examples we have not only the lattice approximations of the *light-tailed* Pearson diffusions, but also another process, which is the continuous-time Ehrenfest urn process. Hence, we can observe that with the definition of solvable birth–death process we do not only cover these lattice approximations, but we also gain some other birth–death processes that are not covered in the theory of light-tailed Pearson diffusions (see [32,33]).

Let us also recall that if we consider the matrix $p(t) = (p(t, x; y))_{x,y \in E}$ where $p(t, x; y) = \mathbb{P}(N(t + s) = x | N(s) = y)$ is the transition probability function, then $p(t)$ solves both the so-called backward and forward equations (see [46])

$$p'(t) = \mathcal{G} \cdot p(t) \qquad\qquad p'(t) = p(t) \cdot \mathcal{G}.$$

In particular we can consider the adjoint operator $\mathcal{L}$ of $\mathcal{G}$ (still defined on $\ell^2(\mathbf{m})$) such that $p'(t) = \mathcal{L}p(t)$. For this reason, we refer to $\mathcal{L}$ as the forward operator. Let us observe that for a birth–death process $N(t)$ the forward operator $\mathcal{L}$ is defined as

$$\mathcal{L}f(x) = -\nabla^-((b(\cdot) - d(\cdot))f(\cdot))(x) + \Delta(d(\cdot)f(\cdot))(x).$$

Moreover, let us observe that $\mathcal{L}$ can be considered as a (possibly infinite) tridiagonal matrix such that $p'(t) = \mathcal{L} \cdot p(t)^{\mathbf{t}}$ where with $\mathbf{t}$ we denote the transpose. Now let us show what the orthogonal polynomials $P_n(x)$ represent for the forward operator.

Lemma 1 *Let $N(t)$ be a solvable birth–death process with forward operator $\mathcal{L}$, invariant measure $\mathbf{m}$ and associated family of orthonormal polynomials $Q_n(x)$. Then we*

have for any $n \in E$ and any $x \in E$,

$$\mathcal{L}(mQ_n)(x) = m(x)\lambda_n Q_n(x).$$

Proof We have, recalling that $\Delta = \nabla^-\nabla^+$ and that ∇^- is a linear operator,

$$\mathcal{L}(mQ_n)(x) = -\nabla^-((b-d)mQ_n)(x) + \Delta(dmQ_n)(x)$$
$$= \nabla^-[-(b-d)mQ_n + \nabla^+(dmQ_n)](x).$$

Now, by discrete Leibniz rule on ∇^+, denoting $\widetilde{d}(x) = d(x+1)$ and $\widetilde{m} = m(x+1)$, we have

$$\mathcal{L}(mQ_n)(x) = \nabla^-[-(b-d)mQ_n + Q_n\nabla^+(dm) + \widetilde{d}\widetilde{m}\nabla^+Q_n](x)$$
$$= \nabla^-[Q_n(-(b-d)mQ_n + \nabla^+(dm)) + \widetilde{d}\widetilde{m}\nabla^+Q_n](x)$$
$$= \nabla^-(\widetilde{d}\widetilde{m}\nabla^+Q_n)(x).$$

Now let us use again Leibniz rule on ∇^- to achieve

$$\mathcal{L}(mQ_n)(x) = d(x)m(x)\Delta Q_n(x) + \nabla^+Q_n(x)\nabla^-(\widetilde{d}\widetilde{m})(x).$$

Now let us work with $\nabla^-(\widetilde{d}\widetilde{m})(x)$. We have

$$\nabla^-(\widetilde{d}\widetilde{m})(x) = d(x+1)m(x+1) - d(x)m(x).$$

However, by the discrete Pearson equation (1) we obtain

$$d(x+1)m(x+1) - d(x)m(x) = (b(x) - d(x))m(x)$$

and then we have
$$\nabla^-(\widetilde{d}\widetilde{m})(x) = (b(x) - d(x))m(x).$$

Hence, we achieve

$$\mathcal{L}(mQ_n)(x) = m(x)[(b(x) - d(x))\nabla^+Q_n(x) + d(x)\Delta Q_n(x)]$$
$$= m(x)\mathcal{G}Q_n(x) = m(x)\lambda_n Q_n(x),$$

concluding the proof.   □

Thus we have, as a consequence of the discrete Pearson equation, the discrete version of the spectral decomposition for parabolic problems with the generator and the forward operator of a light-tailed Pearson diffusion (which is a direct consequence of the results in [22] once we notice that in our case the spectral measure coincides with the stationary one):

Theorem 1 *Let $N(t)$ be a solvable birth–death process with state space E, generator $\mathcal{G}$, forward operator $\mathcal{L}$, invariant measure $\mathbf{m}$ and family of associated orthonormal polynomials $(Q_n)_{n \in E}$. Then the following assertions hold true:*

– *The transition probability function $p(t, x; y) = \mathbb{P}(N(t+s) = x | N(s) = y)$ for $x, y \in E$ and $t, s \geq 0$ admit the following spectral representation:*

$$p(t, x; y) = m(x) \sum_{n \in E} e^{\lambda_n t} Q_n(x) Q_n(y)$$

for any $x, y \in E$ and $t \geq 0$.
– *If $g \in \ell^2(\mathbf{m})$ with $g(x) = \sum_{n \in E} g_n Q_n(x)$ then the strong solution of the Cauchy problem*

$$\begin{cases} \frac{du}{dt}(t, y) = \mathcal{G}u(t, y) & t \geq 0, y \in E \\ u(0, y) = g(y) & y \in E \end{cases}$$

is given by

$$u(t, y) = \sum_{n \in E} g_n e^{\lambda_n t} Q_n(y) = \sum_{x \in E} p(t, x; y) g(x).$$

In particular $p(t, x; y)$ is the fundamental solution of $\frac{du}{dt}(t, y) = \mathcal{G}u(t, y)$ and u admits the stochastic interpretation:

$$u(t, y) = \mathbb{E}_y[g(N(t))]$$

where $\mathbb{E}_y[\cdot] = \mathbb{E}[\cdot | N(0) = y]$.
– *If $f/m \in \ell^2(\mathbf{m})$ with $\frac{f(x)}{m(x)} = \sum_{n \in E} f_n Q_n(x)$ the strong solution of the Cauchy problem*

$$\begin{cases} \frac{dv}{dt}(t, x) = \mathcal{L}u(t, x) & t \geq 0, x \in E \\ v(0, x) = f(x) & x \in E \end{cases}$$

is given by

$$v(t, x) = m(x) \sum_{n \in E} f_n e^{\lambda_n t} Q_n(x) = \sum_{y \in E} p(t, x; y) f(y).$$

In particular $p(t, x; y)$ is the fundamental solution of $\frac{dv}{dt}(t, x) = \mathcal{L}v(t, x)$ and, if $f \geq 0$ with $\|f\|_{\ell^1} = 1$, v admits the stochastic interpretation:

$$v(t, x) = \mathbb{P}_f(N(t) = x)$$

where $\mathbb{P}_f$ is the probability measure obtained by conditioning with respect to the fact that $N(0)$ admits distribution f.

From this theorem, it is easy to determine the covariance of any solvable birth–death process in its stationary form. First of all, let us observe that the stationary version of $N(t)$ admits moments of any order. This is obvious if E is finite. To show this when $E = \mathbb{N}_0$, let us first show the following Proposition.

Proposition 1 *Let $N(t)$ be a solvable birth–death process with invariant measure* **m** *and state space $E = \mathbb{N}_0$. Then there exists a constant $\rho < 1$ and a state $x_0 \in E$ such that for any $x \geq x_0$ it holds*

$$m(x) \leq \rho^{x-x_0} m(x_0) \tag{3}$$

Proof First of all, let us observe that since $b(x)$ and $d(x)$ are polynomials, then $\lim_{x \to +\infty} \frac{b(x)}{d(x+1)}$ always exists. Moreover, we can rewrite the discrete Pearson equation as

$$m(x+1) = \frac{b(x)}{d(x+1)} m(x).$$

Thus, we have that m is well defined if and only if

$$\sum_{x=1}^{+\infty} \prod_{k=0}^{x} \frac{b(k)}{d(k+1)} < +\infty.$$

It is easy to see that such condition implies $\lim_{x \to +\infty} \frac{b(x)}{d(x+1)} \leq 1$. Let us suppose $\lim_{x \to +\infty} \frac{b(x)}{d(x+1)} = 1$. This could happen only if $b(x)$ and $d(x)$ are polynomials of the same degree and with the same director coefficient. However, if $b(x)$ and $d(x)$ are polynomials of degree at most 1, then $\lambda_n = 0$ for any $n \geq 1$, that is absurd. Thus, we have that $b(x)$ and $d(x)$ are polynomials of degree 2. However, since $\lambda_n < 0$ for any $n \geq 1$, it follows that the coefficient director of $d(x)$ must be negative. However, this means that for x big enough it holds $d(x) < 0$, which is absurd. Thus, we conclude that

$$\lim_{x \to +\infty} \frac{b(x)}{d(x+1)} = l < 1.$$

Now let us consider $\rho \in (l, 1)$. Then there exists a state $x_0 \in E$ such that $\frac{b(x)}{d(x+1)} < \rho$ as $x \geq x_0$. Thus, we have

$$m(x+1) < \rho m(x)$$

for any $x \geq x_0$. Finally, the assertion follows from the previous inequality by induction. $\square$

As a direct consequence of the previous proposition we obtain

Corollary 1 *Let $N(t)$ be a solvable birth–death process with invariant distribution* **m** *such that $N(0)$ admits* **m** *as distribution. Then $N(t)$ admits moments of any order.*

Now we can focus on the autocovariance function of the process $N(t)$.

Corollary 2 *Let $N(t)$ be a solvable birth–death process with invariant measure* **m**. *Then there exists a constant $a_1 \in \mathbb{R}$ such that, for any $t, s \geq 0$*

$$Cov_m(N(t), N(s)) = a_1^2 e^{\lambda_1 |t-s|}.$$

Proof First of all let us recall that the stationary version of $N(t)$ admits moments of any order, thus in particular also second-order moments and then the autocovariance is well-defined. Since we are supposing that $N(t)$ is stationary, we have, for any $t \geq s$,

$$Cov_m(N(t), N(s)) = Cov_m(N(t-s)N(0))$$

Thus, let us consider $t \geq 0$ and let us evaluate $Cov_m(N(t), N(0))$. To do this, let us rewrite

$$
\begin{aligned}
Cov_m(N(t), N(0)) &= \mathbb{E}_m[N(t)N(0)] - \mathbb{E}_m[N(t)]\mathbb{E}_m[N(0)] \\
&= \mathbb{E}_m[N(t)N(0)] - \mathbb{E}_m[N(0)]^2.
\end{aligned}
$$

Now let us first evaluate $\mathbb{E}_x[N(t)]$. Since **m** admits second moment then $\iota(x) = x$ is in $\ell^2(\mathbf{m})$. Moreover, since $\deg(\iota(x)) = 1$, it can be written as a linear combination of $Q_0 = 1$ and Q_1. Let then

$$\iota(x) = a_0 + a_1 Q_1(x).$$

By the previous theorem we have that

$$\mathbb{E}_x[N(t)] = a_0 + a_1 e^{\lambda_1 t} Q_1(x)$$

(recalling that $\lambda_0 = 0$ for any solvable birth–death process). Starting from this observation, we have that

$$\mathbb{E}_m[N(t)N(0)] = \sum_{x \in E} x m(x) \mathbb{E}_x[N(t)] = a_0 \sum_{x \in E} x m(x) + a_1 e^{\lambda_1 t} \sum_{x \in E} x m(x) Q_1(x).$$

As we stated before, we can write $x = a_0 + a_1 Q_1(x)$, thus we have

$$
\begin{aligned}
\mathbb{E}_m[N(t)N(0)] = {}& a_0^2 \sum_{x \in E} m(x) + a_0 a_1 \sum_{x \in E} Q_0 Q_1(x) m(x) \\
&+ a_0 a_1 e^{\lambda_1 t} \sum_{x \in E} Q_0 m(x) Q_1(x) + a_1^2 e^{\lambda_1 t} \sum_{x \in E} m(x) Q_1^2(x).
\end{aligned}
$$

By using the orthonormality relation we have

$$\mathbb{E}_m[N(t)N(0)] = a_0^2 + a_1^2 e^{\lambda_1 t}.$$

Now let us evaluate $\mathbb{E}_m[N(0)]$. We have

$$
\begin{aligned}
\mathbb{E}_m[N(0)] &= \sum_{x \in E} x m(x) = a_0 + a_1 \sum_{x \in E} Q_1(x) m(x) \\
&= a_0 + a_1 \sum_{x \in E} Q_0 Q_1(x) m(x) = a_0.
\end{aligned}
$$

We finally achieve

$$Cov_m(N(t), N(0)) = a_1^2 e^{\lambda_1 t}$$

concluding the proof.   □

2.1 Classification of Solvable Birth–Death Processes

We can actually improve the result in Proposition 1 by obtaining a complete classification of solvable birth–death processes. Indeed we have the following Proposition.

Proposition 2 *Let $N(t)$ be a solvable birth–death process with state space E. Then one of the following statements holds true:*

– *E is finite;*
– *$N(t)$ is an immigration-death process;*
– *$N(t)$ is a Meixner process.*

In particular, if $(P_n)_{n \in E}$ is the family of orthogonal polynomials associated to $N(t)$, then either E is finite or $(P_n)_{n \in E}$ coincide with its dual family and $P_n(x) = P_x(n)$ for any $x, n \in E$.

Proof First of all, suppose $E = \mathbb{N}_0$. Then we have $\lim_{x \to +\infty} \frac{b(x)}{d(x+1)} < 1$. In particular this implies that $\deg b \le \deg d$. If $\deg d = 0$, then also $\deg b = 0$ and in such case $\lambda_n = 0$ for any $n \in \mathbb{N}$, which is absurd. Thus, $\deg d = 1, 2$. However, we have already seen that if $\deg d = 2$, then, since $\lambda_n < 0$ for any $n \ge 1$, the director coefficient of d must be negative and this is a contradiction with the fact that d is non negative on E. Thus, we conclude that if $E = \mathbb{N}_0$, then $\deg d = 1$. Moreover, arguing as before, we know that the director coefficient of d must be positive (since if it is negative then d is negative for big values of $x \in E$) and, being also $d(0) = 0$, it must hold $d(x) = dx$ for some $d > 0$. Now let us consider b. Since we want $E = \mathbb{N}_0$, arguing as we did with d, we need the director coefficient of b to be positive. Hence, we have two possibilities:

– $\deg b = 0$, thus $b > 0$ is constant and we get an immigration-death process;
– $\deg b = 1$, thus $b(x) = b(x + \beta)$ for some $b, \beta > 0$ and then we get a Meixner process (since also $b < d$ by the condition $\lim_{x \to +\infty} \frac{b(x+\beta)}{d(x+1)} = \frac{b}{d} < 1$).

This concludes the proof.   □

As we can see from the previous Proposition, we already discussed as examples the unique two cases in which the state space is countably infinite, that are actually the ones for which the proof of the main results (that will follow) are more articulated. Let us also observe some other particular properties concerning the classification of solvable birth–death processes:

– In the class of solvable birth–death processes we have lattice approximations of Pearson diffusions of first spectral category. Pearson diffusions are statistically tractable diffusions (see [19]), however, their spectral behaviour can be distinguished in three different classes (see [33]). The first spectral class (the one that contains Pearson diffusions whose generators admit purely discrete spectrum) is composed of the Ornstein-Uhlenbeck, Cox-Ingersoll-Ross and Jacobi diffusions.

These diffusions are approximated, respectively, by immigration-death, Meixner and Hahn birth–death processes;

– However, we do not obtain the analogous of Pearson diffusions of the second spectral category: this is due to the fact that to obtain these, we need $d(x)$ to be a polynomial of degree 2 and with positive director coefficient, which goes in contradiction with the request that the spectrum of the generator of a solvable birth–death process is purely discrete and non-positive together with the fact that the state space is infinite. However, if we reduce the state space to a finite segment of $\mathbb{N}_0$, we can still cover these cases, in which $d(x)$ admits discriminant $\Delta \geq 0$ and $d(0) = 0$, by asking that $b(x)$ admits the first root $x_0 \in \mathbb{N}$ and $E = \{0, \ldots, x_0\}$. If $\Delta > 0$ we also have to ask that the other root of $d(x)$ is negative. In any case, these lattice schemes do not approximate reciprocal Gamma and Fisher-Snedecor diffusions in the support of their invariant measure, but they still provide, in some cases, an approximation scheme for the backward and forward Kolmogorov equations in a subset of the full support;

– Even considering a finite segment in $\mathbb{N}_0$, we cannot approximate by a solvable birth–death process the Student distribution, which belongs to the third spectral category: this is due to the fact that to achieve this approximation, $d(x)$ must be a polynomial of degree 2, with positive director coefficient and negative discriminant, which is in contradiction with the condition $d(0) = 0$;

– We also get some birth–death processes that are not actual lattice approximation of any Pearson diffusion, such as in the Ehrenfest process case, whose state space is finite but $d(x)$ is a polynomial of degree 1.

– In particular, let us notice that for any solvable birth–death processes $N(t)$ such that the polynomial $b(x-1) - d(x)$ is of degree 1, the invariant measure $\mathbf{m}$ is in the Ord family (see [21]). Indeed if $\deg(b) = \deg(d) = 2$, we can suppose that (since $d(0) = 0$)

$$d(x) = \widetilde{a}x^2 + \widetilde{d}_1 x, \qquad b(x) = \widetilde{a}x^2 + \widetilde{b}_1 x + \widetilde{b}_2.$$

By using the relation $m(x-1) = \frac{d(x)}{b(x-1)} m(x)$ and setting

$$k = 2\widetilde{a} + \widetilde{d}_1 - \widetilde{b}_1, \qquad a = \frac{\widetilde{a} - \widetilde{b}_1 + \widetilde{b}_2}{k},$$

$$b_0 = 0, \qquad b_1 = \frac{\widetilde{d}_1 + \widetilde{a}}{k}, \qquad b_2 = \frac{\widetilde{a}}{k},$$

where $k \neq 0$ since it is the director coefficient of the first degree polynomial $b(x-1) - d(x)$, we get the equation

$$\frac{\nabla^+ m(x-1)}{m(x)} = \frac{a - x}{(a + b_0) + (b_1 - 1)x + b_2 x(x-1)} \tag{4}$$

which is the characterizing equation of the Ord family.

If $\deg(d) = \deg(b) = 1$, then we can suppose that

$$d(x) = \tilde{d}_1 x, \qquad b(x) = \tilde{b}_1 x + \tilde{b}_0.$$

To have $\deg(b(x-1) - d(x)) = 1$, we need $\tilde{d}_1 \neq \tilde{b}_1$. Thus, setting

$$k = \tilde{d}_1 - \tilde{b}_1 \neq 0, \qquad\qquad a = \frac{\tilde{b}_0 - \tilde{b}_1}{k},$$

$$b_0 = 0, \qquad\qquad b_1 = \frac{\tilde{b}_0}{k}, \qquad\qquad b_2 = 0,$$

we still get Eq. (4). In particular we get the form

$$\frac{\nabla^+ m(x-1)}{m(x)} = \frac{a - x}{a + (b_1 - 1)x} \tag{5}$$

which is the characterizing equation of the Katz family. Finally, last case is $\deg(d) = 1$ and $\deg(b) = 0$, that is to say

$$d(x) = \tilde{d}_1 x \qquad b(x) \equiv \tilde{b}_0$$

in which the substitution that leads to Eq. (4) is given by

$$k = \tilde{d}_1, \qquad\qquad a = \frac{\tilde{b}_0}{k},$$
$$b_0 = 0, \qquad\qquad b_1 = 1, \qquad\qquad b_2 = 0.$$

Even in this case we are actually in the Katz family. Finally, let us observe that in such case the state space has to be infinite (since $b(x) \equiv \tilde{b}_0 > 0$) and then we are considering an immigration-death process (and the invariant measure is a Poisson measure). In particular we cover all the distributions of the Katz family (see [21]).

– It is also interesting to see that we cover the Poisson, Binomial, Negative Binomial and Hypergeometric invariant distribution cases, which are all in the cumulative Ord family (see [1]).

3 Inverse Subordinators and Non-local Convolution Derivatives

Now let us introduce our main object of study. Let us denote by $\mathcal{BF}$ the convex cone of Bernstein functions, that is to say $\Phi \in \mathcal{BF}$ if and only if $\Phi \in C^\infty(\mathbb{R}^+)$, $\Phi(\lambda) \geq 0$ and for any $n \in \mathbb{N}$

$$(-1)^n \frac{\mathrm{d}^n \Phi}{\mathrm{d}\lambda^n}(\lambda) \leq 0.$$

In particular it is known that for $\Phi \in \mathcal{BF}$ the following Lévy-Khintchine representation ([45]) is given

$$\Phi(\lambda) = a + b\lambda + \int_0^{+\infty} (1 - e^{-\lambda t}) \nu(\mathrm{d}t) \tag{6}$$

where $a, b \geq 0$ and ν is a Lévy measure on $\mathbb{R}^+$ such that

$$\int_0^{+\infty} (1 \wedge t) \nu(\mathrm{d}t) < +\infty. \tag{7}$$

The triple (a, b, ν) is called the Lévy triple of Φ. Also the vice versa can be shown, i.e. for any Lévy triple (a, b, ν) such that ν is a Lévy measure satisfying the integral condition (7) there exists a unique Bernstein function Φ such that Eq. (6) holds. In the rest of the paper, we suppose that $a = b = 0$ and $\nu(0, +\infty) = +\infty$. It is also known (see [45]) that for each Bernstein function $\Phi \in \mathcal{BF}$ there exists a unique subordinator $\sigma_\Phi = \{\sigma_\Phi(y), y \geq 0\}$ (i. e. an increasing Lévy process) such that

$$\mathbb{E}[e^{-\lambda \sigma_\Phi(y)}] = e^{-y\Phi(\lambda)}.$$

For general notion on subordinators we refer to [13, Chapter 3] and [14]. In particular the hypothesis $b = 0$ ensure that σ_Φ is a pure jump process, $a = 0$ implies that it is not killed and $\nu(0, +\infty) = +\infty$ implies that σ_Φ is strictly increasing (a. s.).

Let us now fix our Bernstein function Φ and its associated subordinator σ_Φ. Now we can define the inverse subordinator E_Φ as, for any $t > 0$

$$E_\Phi(t) := \inf\{y \geq 0 : \sigma_\Phi(y) > t\}.$$

Under our hypotheses, we have that $E_\Phi(t)$ is absolutely continuous for any $t > 0$ and its sample paths are almost surely continuous. Let us denote by $f_\Phi(s; t)$ its density (see [34]). Let us recall (see [34]) that, denoting by $\overline{f}_\Phi(s; \lambda)$ the Laplace transform of $f_\Phi(s; t)$ with respect to t,

$$\overline{f}_\Phi(s; \lambda) = \frac{\Phi(\lambda)}{\lambda} e^{-s\Phi(\lambda)}.$$

Now let us introduce the non-local convolution derivatives (of Caputo type) associated with Φ. Indeed, for Φ identified by the Lévy triple $(0, 0, \nu)$, let us define the Lévy tail $\overline{\nu}(t) = \nu(t, +\infty)$. Now let us recall the definition of non-local convolution derivative, defined in [27,49].

Definition 2 Let $f : \mathbb{R}^+ \to \mathbb{R}$ be an absolutely continuous function. Then we define the non-local convolution derivative induced by Φ of f as

$$\frac{\mathrm{d}^\Phi}{\mathrm{d}t^\Phi} f(t) = \int_0^t f'(\tau) \overline{\nu}(t - \tau) \mathrm{d}\tau. \tag{8}$$

Let us observe that one can define also the regularized version of the non-local convolution derivative as

$$\frac{d^{\Phi}}{dt^{\Phi}} f(t) = \frac{d}{dt} \int_0^t (f(\tau) - f(0+))\bar{\nu}(t - \tau)d\tau \qquad (9)$$

observing that it coincides with the previous definition on absolutely continuous functions. Here we will always consider the regularized form (9).

It can be shown, by Laplace transform arguments (see, for instance [4,28]) or by Green functions arguments (see [29]), that the (eigenvalue) Cauchy problem

$$\begin{cases} \frac{d^{\Phi}}{dt^{\Phi}} \mathfrak{e}_{\Phi}(t; \lambda) = \lambda \mathfrak{e}_{\Phi}(t; \lambda) & t > 0 \\ \mathfrak{e}_{\Phi}(0; \lambda) = 1 \end{cases}$$

admits a unique solution for any $\lambda > 0$ and it is given by $\mathfrak{e}_{\Phi}(t; \lambda) := \mathbb{E}[e^{\lambda E_{\Phi}(t)}]$ (hence, in particular, it is a completely monotone function in λ for fixed t). Let us recall that if $\Phi(\lambda) = \lambda^{\alpha}$ for $\alpha \in (0, 1)$, then $\bar{\nu}(t) = \frac{t^{-\alpha}}{\Gamma(1-\alpha)}$ and $\frac{d^{\Phi}}{dt^{\Phi}}$ coincides with the fractional Caputo derivative of order α. In particular this means that in this case $\mathfrak{e}_{\Phi}(t; \lambda) = E_{\alpha}(\lambda t^{\alpha})$ where E_{α} is the one-parameter Mittag-Leffler function defined, for $t \in \mathbb{R}$ as

$$E_{\alpha}(t) = \sum_{k=0}^{+\infty} \frac{t^k}{\Gamma(\alpha k + 1)}.$$

Let us recall (see [48]) that

$$E_{\alpha}(-\lambda t^{\alpha}) \leq \frac{1}{1 + \frac{t^{\alpha}}{\Gamma(1+\alpha)}\lambda}$$

hence it is not difficult to show the following Proposition.

Proposition 3 *For any $\lambda > 0$ it holds*

$$\lambda E_{\alpha}(-\lambda t^{\alpha}) \leq \frac{\Gamma(1 + \alpha)}{t^{\alpha}}.$$

The proof is identical to the one of [8, Lemma 4.2]. We want to achieve a similar bound for any inverse subordinator. This is done by means of the following proposition.

Proposition 4 *Fix $t > 0$. Then there exists a constant $K(t)$ such that*

$$\lambda \mathfrak{e}_{\Phi}(t; -\lambda) \leq K(t), \ \forall \lambda \in [0, +\infty). \qquad (10)$$

Proof Let us first recall that $\mathfrak{e}_{\Phi}(t; -\lambda) = \mathbb{E}[e^{-\lambda E_{\Phi}(t)}]$, thus it is the Laplace transform of $f_{\Phi}(s; t)$ with respect to s. In particular it is completely monotone in λ and $\mathfrak{e}_{\Phi}(t; 0) = 1$. Now let us recall that $f_{\Phi}(0+; t) = \bar{\nu}(t)$ (see, for instance, [49, Theorem 4.1]). On

the other hand, by the initial-value theorem (see, for instance, [18, Section 17.8]), we have

$$\lim_{\lambda \to +\infty} \lambda \mathfrak{e}_\Phi(t; -\lambda) = f_\Phi(0+; t) = \overline{v}(t) < +\infty.$$

Hence, we can consider the continuous function $\lambda \in [0, +\infty] \mapsto \lambda \mathfrak{e}_\Phi(t; -\lambda) \in \mathbb{R}^+$ and obtain (10) by Weierstrass theorem.   □

Let us give some examples of Bernstein functions and associated subordinators.

– We have already referred to the α-stable subordinator, i.e. the one we get when we choose $\Phi(\lambda) = \lambda^\alpha$ for $\alpha \in (0, 1)$. In such case, extensive informations on inverse α-stable subordinators are given in [35]. As we stated before, we have in particular $\mathfrak{e}_\Phi(t; \lambda) = E_\alpha(\lambda t^\alpha)$, where E_α is the one-parameter Mittag-Leffler function (see [15]). As a particular property, let us recall that, if we denote by σ_α the α-stable subordinator and g_α the density of the random variable $\sigma_\alpha(1)$, then the inverse α-stable subordinator $E_\Phi(t)$ admits density

$$f_\alpha(s; t) = \frac{t}{\beta} s^{-1-\frac{1}{\beta}} g_\alpha(ts^{-\frac{1}{\beta}});$$

– If we fix a constant $\theta > 0$ and define $\Phi(\lambda) = (\lambda+\theta)^\alpha - \theta^\alpha$ we obtain the tempered α-stable subordinator with tempering parameter $\theta > 0$. Denoting by $\sigma_{\alpha,\theta}(t)$ this subordinator, one can show that the density of the subordinator is given by

$$g_{\alpha,\theta}(s; t) = e^{-\theta s + t\theta^\alpha} g_\alpha(s; t)$$

where g_α is the density of the α-stable subordinator $\sigma_\alpha(t)$. An important property to recall is that the introduction of the tempering parameter implies the existence of all the moments of $\sigma_{\alpha,\theta}(t)$ (while this is not true for σ_α). Inverse tempered stable subordinators are studied for instance in [30]. Moreover, it can be shown that the Lévy tail $\overline{v}$ is given by

$$\overline{v}(t) = \frac{\alpha\theta^\alpha \Gamma(-\alpha; t)}{\Gamma(1-\alpha)},$$

where $\Gamma(\alpha; x) = \int_x^{+\infty} t^{\alpha-1} e^{-t} dt$ is the upper incomplete Gamma function;

– For $\Phi(\lambda) = \log(1+\lambda^\alpha)$ as $\alpha \in (0, 1)$ we obtain the geometric α-stable subordinator. From the form of the Bernstein function associated to the geometric α-stable subordinator, one obtains (see [47, Theorem 2.6]) that the density $g_{G,\alpha}$ of the random variable $\sigma_{G,\alpha}(1)$ (where $\sigma_{G,\alpha}(t)$ is the geometric α-stable subordinator) satisfies the following asymptotics:

$$g_{G,\alpha}(s) \sim \frac{s^{\alpha-1}}{\Gamma(\alpha)} \qquad \text{as } s \to 0^+;$$

$$g_{G,\alpha}(s) \sim 2\pi \sin\left(\frac{\alpha\pi}{2}\right) \Gamma(1+\alpha) s^{-\alpha-1} \qquad \text{as } s \to +\infty.$$

Concerning the Lévy tail $\bar{\nu}$, it cannot be explicitly expressed, but it has been shown in [47, Theorem 2.5] that it satisfies the following asymptotic relation:

$$\bar{\nu}(t) \sim \frac{t^{-\alpha}}{\Gamma(1-\alpha)} \qquad \text{as } t \to +\infty.$$

– If in the previous example we consider $\alpha = 1$ we obtain the Gamma subordinator. In this specific case, one can obtain explicitly the Lévy tail $\bar{\nu}$ as (see, for instance, [47] and references therein)

$$\bar{\nu}(t) = \Gamma(0; t).$$

4 Non-local Forward and Backward Equations

Let us consider $N(t)$ to be a solvable birth–death process with state space E and invariant measure $\mathbf{m}$ and let us denote by $\mathcal{G}$ and $\mathcal{L}$ its backward and forward operators, respectively. Let us first focus on the backward equation.

4.1 Heuristic Derivation of the Strong Solution

Let us consider a Bernstein function Φ and the Cauchy problem

$$\begin{cases} \frac{\partial^{\Phi}}{\partial t^{\Phi}} u(t,x) = \mathcal{G}u(t,x) & t > 0, \ x \in E \\ u(0,x) = g(x) & x \in E. \end{cases} \tag{11}$$

Suppose that $g \in \ell^2(\mathbf{m})$. Let us consider $(Q_n)_{n \in E}$ the family of orthonormal polynomials associated to $N(t)$. Then we can decompose g as

$$g(x) = \sum_{n \in E} g_n Q_n(x)$$

for some coefficients $(g_n)_{n \in E}$. Let us suppose we want to find a solution $u(t,x)$ by separation of variables. Thus, let us suppose $u(t,x) = T(t)\varphi(x)$. If we substitute this relation in the first equation of (11) we obtain the following coupled equations

$$\begin{cases} \mathcal{G}\varphi(x) = \lambda\varphi(x) & x \in E \\ \frac{d^{\Phi}}{dt^{\Phi}} T(t) = \lambda T(t) & t > 0. \end{cases}$$

Concerning the first equation, let us observe that we need to set $\varphi(x) = Q_n(x)$ (up to a multiplicative constant) and $\lambda = \lambda_n$ for some $n \in E$. Let us also recall that $\lambda_n < 0$. Concerning the second equation we get

$$T(t) = \mathfrak{e}_{\Phi}(t; \lambda_n).$$

and then we have for some $n \in E$

$$u(t, x) = Q_n(x)\mathfrak{e}_\Phi(t; \lambda_n).$$

We can also have solutions that are linear combinations of $Q_n \mathfrak{e}_\Phi$. Let us suppose that we can consider eventually infinite linear combinations. Then we expect a solution of the form

$$u(t, x) = \sum_{n \geq 0} u_n Q_n(x)\mathfrak{e}_\Phi(t; \lambda_n)$$

for some coefficients $(u_n)_{n \in E}$. Finally, let us observe that

$$\sum_{n \in E} g_n Q_n(x) = g(x) = u(0, x) = \sum_{n \in E} u_n Q_n(x)$$

then, since the components $(g_n)_{n \in E}$ are uniquely determined, then $u_n = g_n$ for any $n \in E$.

Finally, we expect the solution to be of the form

$$u(t, x) = \sum_{n \in E} g_n Q_n(x)\mathfrak{e}_\Phi(t; \lambda_n).$$

Now we want to formalize this reasoning.

4.2 The Backward Equation

Before working in the general case, we need to exploit what will be our fundamental solution. To do this, let us show the following Lemma.

Lemma 2 *Let $N(t)$ be a solvable birth–death process with state space E, generator $\mathcal{G}$, invariant measure $\mathbf{m}$ and family of associated classical orthogonal polynomials $(P_n)_{n \in E}$. Then the series*

$$p_\Phi(t, x; y) = m(x) \sum_{n \in E} \mathfrak{e}_\Phi(t; \lambda_n) Q_n(x) Q_n(y), \tag{12}$$

where Q_n are the normalized orthogonal polynomials, absolutely converges for fixed $t \geq 0$ and $x, y \in E$.

Proof Let us first observe that if E is finite, then the summation (12) is actually finite. Thus, let us consider the case in which $E = \mathbb{N}_0$. Let us denote by $\mathfrak{d}_n = \|P_n\|_{\ell^2}$ and let us recall that the dual polynomials $\widetilde{P}_n$ exhibit orthogonality with respect to the measure $\widetilde{m}(x) = \frac{1}{\mathfrak{d}_x^2}$. However, since, if $E = \mathbb{N}_0$, $N(t)$ is either an immigration-death process or a Meixner process, then the dual polynomials $\widetilde{P}_n$ coincide with the polynomials P_n

themselves. We have, by using the self-duality relation $P_n(x) = P_x(n)$,

$$
\begin{aligned}
p_\Phi(t, x; y) &= m(x) \sum_{n=0}^{+\infty} \mathfrak{e}_\Phi(t; \lambda_n) Q_n(x) Q_n(y) \\
&= m(x) \sum_{n=0}^{+\infty} \widetilde{m}(n) \mathfrak{e}_\Phi(t; \lambda_n) P_n(x) P_n(y) \\
&= m(x) \sum_{n=0}^{+\infty} \widetilde{m}(n) \mathfrak{e}_\Phi(t; \lambda_n) P_x(n) P_y(n).
\end{aligned}
$$

Now let us denote by $\mathrm{root}(x)$ the set of all the roots of the polynomial $P_x(n)$. These sets are finite with cardinality at most x. Thus, we can define

$$
n_0 = \lceil \max(\mathrm{root}(x) \cup \mathrm{root}(y)) \rceil + 1.
$$

To show the absolute convergence of the series in $p_\Phi(t, x; y)$, we only need to show the absolute convergence of

$$
\sum_{n=n_0}^{+\infty} \widetilde{m}(n) \mathfrak{e}_\Phi(t; \lambda_n) P_x(n) P_y(n).
$$

Let us observe (see [38, Table 2.3]) that the sign of the director coefficient of P_x (both in the Charlier than in the Meixner case) depends on the parity of x. In particular we have that for any $n \geq n_0$ it holds $\mathrm{sign}(P_x(n) P_y(n)) = (-1)^{x+y}$. So we get

$$
\begin{aligned}
\sum_{n=n_0}^{+\infty} &\left| \widetilde{m}(n) \mathfrak{e}_\Phi(t; \lambda_n) P_x(n) P_y(n) \right| \\
&\leq (-1)^{x+y} \sum_{n=n_0}^{+\infty} \widetilde{m}(n) P_x(n) P_y(n),
\end{aligned}
$$

where we have to observe that $\mathfrak{e}_\Phi(t; \lambda_n) \leq 1$ since $\lambda_n \leq 0$. As before, the series at the right-hand side converges if and only if

$$
\sum_{n=0}^{+\infty} \widetilde{m}(n) P_x(n) P_y(n)
$$

converges. However, by the dual orthogonal relation and the self-duality relation of Charlier and Meixner polynomials, we achieve

$$
\sum_{n=0}^{+\infty} \widetilde{m}(n) P_x(n) P_y(n) = \frac{1}{m(x)} \delta_{x,y}.
$$

for any $x, y \in \mathbb{N}$, concluding the proof.   □

Before exploiting the strong solution, let us show the normal convergence of some auxiliary series of functions.

Lemma 3 *Let $N(t)$ be a solvable birth–death process with state space $E = \mathbb{N}_0$, generator $\mathcal{G}$, invariant measure $\mathbf{m}$ and family of associated classical orthogonal polynomials $(P_n)_{n \geq 0}$. Let $g \in \ell^2(\mathbf{m})$ such that $g(x) = \sum_{n \geq 0} g_n Q_n(x)$ for $x \in E$ and some constants $(g_n)_{n \geq 0}$, where $(Q_n)_{n \geq 0}$ are the normalized orthogonal polynomials. Then*

1. *For any $x \in E$ it holds $\sum_{n \geq 0} |g_n Q_n(x)| \leq \frac{\|g\|_{\ell^2(\mathbf{m})}}{\sqrt{m(x)}}$;*
2. *For any fixed $x \in E$ the sum $\sum_{n \geq 0} \mathfrak{e}_\Phi(t; \lambda_n) g_n Q_n(x)$ normally converges for $t \in [0, +\infty)$;*
3. *For any fixed $x \in E$ and $T_1 > 0$ the series $\sum_{n \geq 0} \lambda_n \mathfrak{e}_\Phi(t; \lambda_n) g_n Q_n(x)$ normally converges for $t \in [T_1, +\infty)$.*

Proof Let us show property (1). Since $g(x) = \sum_{n \geq 0} g_n Q_n(x)$ and Q_n is an orthonormal basis of $\ell^2(\mathbf{m})$, it holds $\sum_{n \geq 0} g_n^2 = \|g\|_{\ell^2(\mathbf{m})}^2$. In particular it holds, by Cauchy-Schwartz inequality and self-duality relation for Charlier and Meixner polynomials

$$
\sum_{n \geq 0} |g_n Q_n(x)| = \sum_{n \geq 0} \sqrt{\widetilde{m}(n)} |g_n P_n(x)|
$$

$$
\leq \left(\sum_{n \geq 0} \widetilde{m}(n) P_n^2(x) \right)^{\frac{1}{2}} \|g\|_{\ell^2(\mathbf{m})}
$$

$$
= \left(\sum_{n \geq 0} \widetilde{m}(n) P_x^2(n) \right)^{\frac{1}{2}} \|g\|_{\ell^2(\mathbf{m})} = \frac{\|g\|_{\ell^2(\mathbf{m})}}{\sqrt{m(x)}}.
$$

To show property (2), let us just observe that $\mathfrak{e}_\Phi(t; \lambda_n) \leq 1$ since $\lambda_n \leq 0$ and then

$$
\sum_{n \geq 0} |\mathfrak{e}_\Phi(t; \lambda_n) g_n Q_n(x)| \leq \sum_{n \geq 0} |g_n Q_n(x)|
$$

where the series on the right-hand side is convergent and independent of $t \geq 0$.

Concerning property (3), by Eq. (10) we obtain for some constant $K(T_1) > 0$, since $\mathfrak{e}_\Phi(t; \lambda_n)$ is decreasing,

$$
\sum_{n \geq 0} |\lambda_n \mathfrak{e}_\Phi(t; \lambda_n) g_n Q_n(x)| \leq \sum_{n \geq 0} |\lambda_n \mathfrak{e}_\Phi(T_1; \lambda_n) g_n Q_n(x)| \leq K(T_1) \sum_{n \geq 0} |g_n Q_n(x)|,
$$

concluding the proof.   □

Now we are ready to show that for the initial datum $g \in \ell^2(\mathbf{m})$ our backward problem admits a solution.

Theorem 2 *Let $N(t)$ be a solvable birth–death process with state space E, generator $\mathcal{G}$, invariant measure* **m** *and family of associated classical orthogonal polynomials $(P_n)_{n \in E}$. Let $g \in \ell^2(\mathbf{m})$ such that $g(x) = \sum_{n \in E} g_n Q_n(x)$ for $x \in E$ and some constants $(g_n)_{n \in E}$, where $(Q_n)_{n \in E}$ are the normalized orthogonal polynomials. Then the Cauchy problem*

$$\begin{cases} \frac{\partial^\Phi u}{\partial t^\Phi}(t, y) = \mathcal{G}u(t, y) & t > 0, \ y \in E \\ u(0, y) = g(y) & y \in E \end{cases} \tag{13}$$

admits a unique strong solution of the form

$$u(t, y) = \sum_{n \in E} \mathfrak{e}_\Phi(t; \lambda_n) g_n Q_n(y). \tag{14}$$

In particular $\sup_{t \geq 0} \|u(t, \cdot)\|_{\ell^2(\mathbf{m})} \leq \|g\|_{\ell^2(\mathbf{m})}$.
Finally $p_\Phi(t, x; y)$ is the fundamental solution of (13)*, in the sense that it is the strong solution of* (13) *for $g(y) = \delta_x(y)$ and for any $g \in \ell^2(\mathbf{m})$ it holds*

$$u(t, y) = \sum_{x \in E} p_\Phi(t, x; y) g(x).$$

Before giving the proof of the Theorem, let us state formally what we mean as strong solution.

Definition 3 A function $u(t, y)$ is a strong solution of the Cauchy problem (13) if:

- $\frac{\partial^\Phi u}{\partial t^\Phi}(t, y)$ exists for any $t > 0$ and $y \in E$;
- The convolution integral in $\frac{\partial^\Phi u}{\partial t^\Phi}(t, y)$ is well-defined as a Bochner integral and the differentiation operator is a strong differentiation in $\ell^2(\mathbf{m})$;
- The equations in (13) hold pointwise;
- $u(t; \cdot) \in C([0, +\infty); \ell^2(\mathbf{m}))$;
- $\frac{\partial^\Phi u}{\partial t^\Phi}(t; \cdot) \in C((0, +\infty); \ell^2(\mathbf{m}))$.

Proof of Theorem 2 Let us first show that $u(t, y)$ in the form of (14) is a strong solution for (13).

First of all, let us recall that, by definition of $\mathfrak{e}_\Phi$ and Q_n, it holds

$$\mathcal{G}[\mathfrak{e}_\Phi(t; \lambda_n) Q_n g_n](x) = \mathfrak{e}_\Phi(t; \lambda_n) g_n \mathcal{G} Q_n(x)$$
$$= \lambda_n \mathfrak{e}_\Phi(t; \lambda_n) g_n Q_n(x) = \frac{\partial^\Phi}{\partial t^\Phi} \mathfrak{e}_\Phi(t; \lambda_n) g_n Q_n(x).$$

Now let us observe that if E is finite then $n_E \in \mathbb{N}$ and we have that $u(t, y)$ is a strong solution of (13) just by linearity of the involved operators.

Let us now suppose that E is countably infinite. First of all, let us show that the series (14) converges in $\ell^2(\mathbf{m})$. To do this, define for $N \in \mathbb{N}$

$$u_N(t, y) = \sum_{n=0}^{N} \mathfrak{e}_\Phi(t; \lambda_n) g_n Q_n(y).$$

Now consider $N < M$ in $\mathbb{N}$ and observe that, being $\lambda_n \leq 0$ and $\mathfrak{e}_\Phi(t; \lambda_n) \leq 1$, it holds

$$\|u_N(t, \cdot) - u_M(t, \cdot)\|^2_{\ell^2(\mathbf{m})} \leq \sum_{n=N}^{M} g_n^2$$

thus, by Cauchy's criterion, we know that the series converges in $\ell^2(\mathbf{m})$.

Now let us denote

$$I_\nu(t) = \int_0^t \overline{\nu}(\tau) d\tau$$

that is increasing and non-negative.

Let us then observe that

$$\int_0^t (u(\tau, y) - u(0+, y))\overline{\nu}(t - \tau) d\tau = \int_0^t (u(\tau, y) - u(0+, y)) dI_\nu(t - \tau).$$

Since we have shown in Lemma 3 that the series defining $u(t, y)$ normally converges for fixed $y \in E$, then we can use [43, Theorem 7.16] to write:

$$\int_0^t (u(\tau, y) - u(0+, y))\overline{\nu}(t-\tau)d\tau = \sum_{n=0}^{+\infty} \left(\int_0^t (\mathfrak{e}_\Phi(\tau; \lambda_n) - 1)\overline{\nu}(t - \tau)d\tau \right) Q_n(y)g_n. \tag{15}$$

As next step, we want to differentiate under the series sign. However, we have to show uniform convergence for t in any compact set $[T_1, T_2]$ of the series of the derivatives to use [43, Theorem 7.17]. Recalling (9), one has

$$\sum_{n=0}^{+\infty} \frac{\partial}{\partial t} \left(\int_0^t (\mathfrak{e}_\Phi(\tau; \lambda_n) - 1)\overline{\nu}(t - \tau)d\tau \right) Q_n(y)g_n = \sum_{n=0}^{+\infty} \frac{\partial^\Phi}{\partial t^\Phi} \mathfrak{e}_\Phi(t; \lambda_n) Q_n(y)g_n$$

$$= \sum_{n=0}^{+\infty} \lambda_n \mathfrak{e}_\Phi(t; \lambda_n) Q_n(y)g_n$$

that normally converges by statement (3) of Lemma 3.

Hence we obtain, differentiating on both sides in (15),

$$\frac{\partial^\Phi}{\partial t^\Phi} u(t, y) = \sum_{n=0}^{+\infty} \frac{\partial^\Phi}{\partial t^\Phi} \mathfrak{e}_\Phi(t; \lambda_n) Q_n(y)g_n = \sum_{n=0}^{+\infty} \mathfrak{e}_\Phi(t; \lambda_n) \mathcal{G} Q_n(y)g_n.$$

Now let us recall that $\mathcal{G} = (b(x) - d(x))\nabla^+ + d(x)\Delta$, hence we have to show that we can exchange ∇^+ with the sign of series. This is due to the commutative property of normally convergent series. Indeed we have

$$
\begin{aligned}
\nabla^+ \sum_{n=0}^{+\infty} e_\Phi(t; \lambda_n) Q_n(y) g_n &= \sum_{n=0}^{+\infty} e_\Phi(t; \lambda_n) Q_n(y+1) g_n - \sum_{n=0}^{+\infty} e_\Phi(t; \lambda_n) Q_n(y) g_n \\
&= \lim_{N \to +\infty} \left(\sum_{n=0}^{N} e_\Phi(t; \lambda_n) Q_n(y+1) g_n - \sum_{n=0}^{N} e_\Phi(t; \lambda_n) Q_n(y) g_n \right) \\
&= \lim_{N \to +\infty} \sum_{n=0}^{N} e_\Phi(t; \lambda_n) \nabla^+ Q_n(y) g_n = \sum_{n=0}^{+\infty} e_\Phi(t; \lambda_n) \nabla^+ Q_n(y) g_n
\end{aligned}
$$

where all the passages are justified by the fact that the two series $\sum_{n=0}^{+\infty} e_\Phi(t; \lambda_n) Q_n(y+1) g_n$, $\sum_{n=0}^{+\infty} e_\Phi(t; \lambda_n) Q_n(y) g_n$ both normally converge by Lemma 3. The same holds for Δ. Thus, we finally have

$$
\frac{\partial^\Phi}{\partial t^\Phi} u(t, y) = \sum_{n=0}^{+\infty} e_\Phi(t; \lambda_n) \mathcal{G} Q_n(y) g_n = \mathcal{G} u(t, y)
$$

for any $t > 0$. Concerning the initial datum, we have

$$
u(0, y) = \sum_{n \in E} g_n Q_n(y) = g(y).
$$

Now let us show strong continuity of u in $[0, +\infty)$ and of $\frac{\partial^\Phi}{\partial t^\Phi} u$ in $(0, +\infty)$. These properties are obvious as E is finite, thus let us suppose $E = \mathbb{N}_0$. Concerning the continuity of u, let us show it in 0^+, since for any other $t \in (0, +\infty)$ the proof is analogous. We have

$$
\|u(t, \cdot) - g(\cdot)\|^2_{\ell^2(\mathbf{m})} = \sum_{n=1}^{+\infty} (1 - e_\Phi(t; \lambda_n))^2 g_n^2.
$$

Now fix $\varepsilon > 0$. Since $(g_n)_{n \geq 0}$ belongs to ℓ^2, there exists $n(\varepsilon) \geq 0$ such that $\sum_{n=n(\varepsilon)}^{+\infty} g_n^2 \leq \varepsilon$. By using the fact that $(1 - e_\Phi(t; \lambda))^2 \leq 1$ for any $\lambda < 0$, we get

$$
\|u(t, \cdot) - g(\cdot)\|^2_{\ell^2(\mathbf{m})} = \sum_{n=1}^{n(\varepsilon)} (1 - e_\Phi(t; \lambda_n))^2 g_n^2 + \varepsilon.
$$

Sending $t \to 0^+$ (using the fact that $e_\Phi(t; \lambda_n)$ is continuous in t and that the first summation is finite) and then $\varepsilon \to 0^+$ we obtain strong continuity of u. Let us discuss the continuity of $\frac{\partial^\Phi}{\partial t^\Phi} u$ in $(0, +\infty)$. To do this, let us consider $t_0 \in [t_1, t_2]$ with $t_1 > 0$.

We have, for any $t \in [t_1, t_2]$, arguing as before and using (10),

$$
\left\| \frac{\partial^\Phi}{\partial t^\Phi} u(t, \cdot) - \frac{\partial^\Phi}{\partial t^\Phi} u(t_0, \cdot) \right\|^2_{\ell^2(\mathbf{m})} = \left\| \sum_{n \in E} \lambda_n (\mathfrak{e}_\Phi(t; \lambda_n) - \mathfrak{e}_\Phi(t_0; \lambda_n)) Q_n(\cdot) g_n \right\|^2_{\ell^2(\mathbf{m})}
$$

$$
= \sum_{n=0}^{+\infty} \lambda_n^2 (\mathfrak{e}_\Phi(t; \lambda_n) - \mathfrak{e}_\Phi(t_0; \lambda_n))^2 g_n^2
$$

$$
\leq \sum_{n=0}^{n(\varepsilon)} \lambda_n^2 (\mathfrak{e}_\Phi(t; \lambda_n) - \mathfrak{e}_\Phi(t_0; \lambda_n))^2 g_n^2 + K(t_1)\varepsilon.
$$

Thus, sending $t \to t_0$ (observing that the first sum is finite) and then $\varepsilon \to 0^+$ we obtain the desired continuity. Uniqueness follows easily from the fact that $(Q_n)_{n \geq 0}$ is an orthonormal system in $\ell^2(\mathbf{m})$ hence the coefficients are unique.
Now let us show the bound of the $\ell^2(\mathbf{m})$ norm. We have

$$
\|u(t, \cdot)\|^2_{\ell^2(\mathbf{m})} = \sum_{n=0}^{+\infty} \mathfrak{e}_\Phi^2(t; \lambda_n) g_n^2 \leq \sum_{n=0}^{+\infty} g_n^2 = \|g\|_{\ell^2(\mathbf{m})}.
$$

We only have to show the second property of Definition 3. Let us first show that the convolution integral involved in the Φ-derivative $\frac{\partial^\Phi}{\partial t^\Phi} u(t, \cdot)$ is a Bochner integral. To do this, we can use Bochner's theorem (see [3, Theorem 1.1.4]). We just have to show that

$$
\int_0^t \|u(\tau, \cdot) - u(0+, \cdot)\|_{\ell^2(\mathbf{m})} \, \overline{v}(t - \tau) \mathrm{d}\tau < +\infty;
$$

to do this, we will work with the square of this quantity. By Jensen's inequality (see [44, Theorem 3.3]) we have

$$
\left(\int_0^t \|u(\tau, \cdot) - u(0+, \cdot)\|_{\ell^2(\mathbf{m})} \, \overline{v}(t - \tau) \mathrm{d}\tau \right)^2
$$

$$
\leq I_v(t) \int_0^t \|u(\tau, \cdot) - u(0+, \cdot)\|^2_{\ell^2(\mathbf{m})} \, \overline{v}(t - \tau) \mathrm{d}\tau
$$

$$
= I_v(t) \int_0^t \sum_{n=1}^{+\infty} (1 - \mathfrak{e}_\Phi(\tau; \lambda_n))^2 g_n^2 \overline{v}(t - \tau) \mathrm{d}\tau,
$$

and then, by monotone convergence theorem and the fact that $0 \le 1 - \mathfrak{e}_\Phi(t; \lambda_n) \le 1$,

$$
\left(\int_0^t \|u(\tau, \cdot) - u(0+, \cdot)\|_{\ell^2(\mathbf{m})} \, \bar{v}(t - \tau) \mathrm{d}\tau \right)^2
$$

$$
\le I_v(t) \sum_{n=1}^{+\infty} \int_0^t (1 - \mathfrak{e}_\Phi(\tau; \lambda_n))^2 \bar{v}(t - \tau) \mathrm{d}\tau \, g_n^2
$$

$$
\le I_v^2(t) \|g\|_{\ell^2(\mathbf{m})} < +\infty,
$$

concluding the proof of the statement.

Next, let us show that the differential involved in $\frac{\partial^\Phi}{\partial t^\Phi} u(t, y)$ is a strong differential in $\ell^2(\mathbf{m})$, that is to say

$$
\lim_{h \to 0} \left\| \frac{1}{h} \left(\int_0^{t+h} (u(t + h - \tau, \cdot) - u(0+, \cdot)) \bar{v}(\tau) \mathrm{d}\tau \right. \right.
$$

$$
\left. \left. - \int_0^t (u(t - \tau, \cdot) - u(0+, \cdot)) \bar{v}(\tau) \mathrm{d}\tau \right) - \sum_{n=0}^{+\infty} \lambda_n \mathfrak{e}_\Phi(t; \lambda_n) Q_n g_n \right\|_{\ell^2(\mathbf{m})} = 0.
$$

$$(16)$$

Let us first argue as $h > 0$. First of all, let us rewrite the difference of the convolution integrals in the following way

$$
\int_0^{t+h} (u(t + h - \tau, y) - u(0+, y)) \bar{v}(\tau) \mathrm{d}\tau - \int_0^t (u(t - \tau, y) - u(0+, y)) \bar{v}(\tau) \mathrm{d}\tau
$$

$$
= \int_0^t (u(t + h - \tau, y) - u(t - \tau, y)) \bar{v}(\tau) \mathrm{d}\tau
$$

$$
+ \int_t^{t+h} (u(t + h - \tau, y) - u(0+, y)) \bar{v}(\tau) \mathrm{d}\tau,
$$

thus we can use the triangular inequality to achieve

$$
\left\| \frac{1}{h} \left(\int_0^{t+h} (u(t + h - \tau, \cdot) - u(0+, \cdot)) \bar{v}(\tau) \mathrm{d}\tau \right. \right.
$$

$$
\left. \left. - \int_0^t (u(t - \tau, \cdot) - u(0+, \cdot)) \bar{v}(\tau) \mathrm{d}\tau \right) - \sum_{n=0}^{+\infty} \lambda_n \mathfrak{e}_\Phi(t; \lambda_n) Q_n g_n \right\|_{\ell^2(\mathbf{m})}
$$

$$
\le \left\| \frac{1}{h} \int_0^t (u(t + h - \tau, \cdot) - u(t - \tau, \cdot)) \bar{v}(\tau) \mathrm{d}\tau - \sum_{n=0}^{+\infty} \lambda_n \mathfrak{e}_\Phi(t; \lambda_n) Q_n g_n \right\|_{\ell^2(\mathbf{m})}
$$

$$
+ \left\| \frac{1}{h} \int_t^{t+h} (u(t + h - \tau, \cdot) - u(0+, \cdot)) \bar{v}(\tau) \mathrm{d}\tau \right\|_{\ell^2(\mathbf{m})}.
$$

Let us first consider the second summand. By Bochner's theorem and the fact that $\bar{v}(t)$ is decreasing we have

$$\left\| \frac{1}{h} \int_t^{t+h} (u(t+h-\tau,\cdot) - u(0+,\cdot))\bar{v}(\tau)\mathrm{d}\tau \right\|_{\ell^2(\mathbf{m})}$$

$$\le \bar{v}(t)\frac{1}{h} \int_0^h \|u(\tau,\cdot) - u(0+,\cdot)\|_{\ell^2(\mathbf{m})} \, \mathrm{d}\tau.$$

Since we have already shown strong continuity of $t \in [0, +\infty) \mapsto u(t,\cdot) \in \ell^2(\mathbf{m})$, we know that $\tau \mapsto \|u(\tau,\cdot) - u(0+,\cdot)\|_{\ell^2(\mathbf{m})}$ is a continuous function, hence, by the integral mean theorem, there exists $\theta(h) \in [0, h]$ such that

$$\frac{1}{h} \int_0^h \|u(\tau,\cdot) - u(0+,\cdot)\|_{\ell^2(\mathbf{m})} \, \mathrm{d}\tau = \|u(\theta(h),\cdot) - u(0+,\cdot)\|_{\ell^2(\mathbf{m})}.$$

Taking the limit as $h \to 0^+$ we conclude that

$$\lim_{h\to 0^+} \frac{1}{h} \int_0^h \|u(\tau,\cdot) - u(0+,\cdot)\|_{\ell^2(\mathbf{m})} \, \mathrm{d}\tau = 0,$$

and then

$$\lim_{h\to 0^+} \left\| \frac{1}{h} \int_t^{t+h} (u(t+h-\tau,\cdot) - u(0+,\cdot))\bar{v}(\tau)\mathrm{d}\tau \right\|_{\ell^2(\mathbf{m})} = 0. \qquad (17)$$

Now let us consider the first summand. By definition of $u(t, y)$ we have

$$\int_0^t (u(t+h-\tau,y) - u(t-\tau,y))\bar{v}(\tau)\mathrm{d}\tau$$

$$= \int_0^t \sum_{n=1}^{+\infty} (\mathfrak{e}_\Phi(t+h-\tau;\lambda_n) - \mathfrak{e}_\Phi(t-\tau;\lambda_n))g_n Q_n(y)\bar{v}(\tau)\mathrm{d}\tau.$$

Let us recall that for any $t, s > 0$ it holds $|\mathfrak{e}_\Phi(t;\lambda_n) - \mathfrak{e}_\Phi(s;\lambda_n)| \le 1$. By Lemma 3, property 1, we have

$$\sum_{n=1}^{+\infty} |\mathfrak{e}_\Phi(t+h-\tau;\lambda_n) - \mathfrak{e}_\Phi(t-\tau;\lambda_n)||g_n Q_n(y)| \le \frac{\|g\|_{\ell^2(\mathbf{m})}}{\sqrt{m(y)}},$$

thus, being $\overline{\nu}$ integrable in $(0, t)$, we can use dominated convergence theorem to achieve

$$\int_0^t (u(t + h - \tau, y) - u(t - \tau, y))\overline{\nu}(\tau)\mathrm{d}\tau$$
$$= \sum_{n=1}^{+\infty} \left(\int_0^t (\mathfrak{e}_\Phi(t + h - \tau; \lambda_n) - \mathfrak{e}_\Phi(t - \tau; \lambda_n))\overline{\nu}(\tau)\mathrm{d}\tau \right) g_n Q_n(y).$$

Observing that

$$\sum_{n=1}^{+\infty} \left| \frac{1}{h} \int_0^t (\mathfrak{e}_\Phi(t + h - \tau; \lambda_n) - \mathfrak{e}_\Phi(t - \tau; \lambda_n))\overline{\nu}(\tau)\mathrm{d}\tau \right| |g_n Q_n(y)| \le \frac{I_\nu(t) \|g\|_{\ell^2(\mathbf{m})}}{h\sqrt{m(y)}}$$

and

$$\sum_{n=1}^{+\infty} |\lambda_n \mathfrak{e}_\Phi(t; \lambda_n)||g_n Q_n(y)| \le \frac{K(t) \|g\|_{\ell^2(\mathbf{m})}}{\sqrt{m(y)}}$$

for a suitable $K(t)$, we have, being all the involved series absolutely convergent,

$$\sum_{n=1}^{+\infty} \left(\frac{1}{h} \int_0^t (\mathfrak{e}_\Phi(t + h - \tau; \lambda_n) - \mathfrak{e}_\Phi(t - \tau; \lambda_n))\overline{\nu}(\tau)\mathrm{d}\tau \right) g_n Q_n(y)$$
$$- \sum_{n=1}^{+\infty} \lambda_n \mathfrak{e}_\Phi(t; \lambda_n) g_n Q_n(y)$$
$$= \sum_{n=1}^{+\infty} \left(\frac{1}{h} \int_0^t (\mathfrak{e}_\Phi(t + h - \tau; \lambda_n) - \mathfrak{e}_\Phi(t - \tau; \lambda_n))\overline{\nu}(\tau)\mathrm{d}\tau \right.$$
$$\left. - \lambda_n \mathfrak{e}_\Phi(t; \lambda_n) \right) g_n Q_n(y). \tag{18}$$

To show that the previous series converges in $\ell^2(\mathbf{m})$, let us define

$$S_N = \sum_{n=1}^{N} \left(\frac{1}{h} \int_0^t (\mathfrak{e}_\Phi(t + h - \tau; \lambda_n) - \mathfrak{e}_\Phi(t - \tau; \lambda_n))\overline{\nu}(\tau)\mathrm{d}\tau \right.$$
$$\left. - \lambda_n \mathfrak{e}_\Phi(t; \lambda_n) \right) g_n Q_n(y)$$

and observe that, for any $M > N$,

$$
\begin{aligned}
\|S_M - S_N\|^2_{\ell^2(\mathbf{m})} &= \sum_{n=N+1}^{M} \left(\frac{1}{h} \int_0^t (\mathfrak{e}_\Phi(t+h-\tau;\lambda_n) \right. \\
&\quad \left. - \mathfrak{e}_\Phi(t-\tau;\lambda_n))\bar{v}(\tau)\mathrm{d}\tau - \lambda_n \mathfrak{e}_\Phi(t;\lambda_n) \right)^2 g_n^2 \\
&\le \sum_{n=N+1}^{M} \left(\left(\frac{1}{h} \int_0^t (\mathfrak{e}_\Phi(t+h-\tau;\lambda_n) \right. \right. \\
&\quad \left. \left. - \mathfrak{e}_\Phi(t-\tau;\lambda_n))\bar{v}(\tau)\mathrm{d}\tau \right)^2 + (\lambda_n \mathfrak{e}_\Phi(t;\lambda_n))^2 \right) g_n^2 \\
&\le \left(\frac{I_v^2(t)}{h^2} + K^2(t) \right) \sum_{n=N+1}^{M} g_n^2
\end{aligned}
$$

for a suitable $K(t)$. Being $\sum_{n=0}^{+\infty} g_n^2$ convergent, by Cauchy's criterion we know that for any $\varepsilon > 0$ there exists $N_0 > 0$ such that, for any $N_0 \le N < M$, it holds $\sum_{n=N+1}^{M} g_n^2 < \frac{\varepsilon}{\left(\frac{I_v^2(t)}{h^2} + K^2(t) \right)}$. In particular, choosing $N_0 \le N < M$ we have that $\|S_M - S_N\|_{\ell^2(\mathbf{m})} < \varepsilon$, thus S_N converges in $\ell^2(\mathbf{m})$. In particular, this means that we can use Parseval's identity on the series in the right-hand side of Eq. (18), hence

$$
\begin{aligned}
&\left\| \frac{1}{h} \int_0^t (u(t+h-\tau,\cdot) - u(t-\tau,\cdot))\bar{v}(\tau)\mathrm{d}\tau - \sum_{n=1}^{+\infty} \lambda_n \mathfrak{e}_\Phi(t;\lambda_n) Q_n g_n \right\|^2_{\ell^2(\mathbf{m})} \\
&= \sum_{n=1}^{+\infty} \left(\frac{1}{h} \int_0^t (\mathfrak{e}_\Phi(t+h-\tau;\lambda_n) - \mathfrak{e}_\Phi(t-\tau;\lambda_n))\bar{v}(\tau)\mathrm{d}\tau - \lambda_n \mathfrak{e}_\Phi(t;\lambda_n) \right)^2 g_n^2 \\
&= \sum_{n=1}^{+\infty} \left(\frac{1}{h} \left(\int_0^{t+h} (\mathfrak{e}_\Phi(t+h-\tau;\lambda_n) - 1)\bar{v}(\tau)\mathrm{d}\tau \right. \right. \\
&\quad \left. - \int_0^t (\mathfrak{e}_\Phi(t-\tau;\lambda_n) - 1)\bar{v}(\tau)\mathrm{d}\tau \right) \\
&\quad \left. - \frac{1}{h} \int_t^{t+h} (\mathfrak{e}_\Phi(t+h-\tau;\lambda_n) - 1)\bar{v}(\tau)\mathrm{d}\tau - \lambda_n \mathfrak{e}_\Phi(t;\lambda_n) \right)^2 g_n^2 \\
&\le 2 \sum_{n=1}^{+\infty} \left(\left(\frac{I_n(t+h) - I_n(t)}{h} - \lambda_n \mathfrak{e}_\Phi(t;\lambda_n) \right)^2 \right. \\
&\quad \left. + \bar{v}^2(t) \left(\frac{1}{h} \int_0^h (\mathfrak{e}_\Phi(s;\lambda_n) - 1)^2 \mathrm{d}s \right) \right) g_n^2,
\end{aligned}
$$

where we set

$$I_n(t) = \int_0^t (\mathfrak{e}_\Phi(t - \tau; \lambda_n) - 1)\overline{v}(\tau)\mathrm{d}\tau$$

and we used Jensen's inequality.

Observe that $I_n(t)$ is a C^1 function with $I_n'(t) = \lambda_n \mathfrak{e}_\Phi(t; \lambda_n)$. By Lagrange's theorem there exists a function $\theta_n(h) \in [t, t + h]$, for each $n \geq 1$, such that

$$\frac{I_n(t + h) - I_n(t)}{h} = I_n'(\theta_n(h)) = \lambda_n \mathfrak{e}_\Phi(\theta_n(h); \lambda_n),$$

and then we obtain

$$\begin{aligned}
&\left\| \frac{1}{h} \int_0^{t+h} (u(t + h - \tau, \cdot) - u(t - \tau, \cdot))\overline{v}(\tau)\mathrm{d}\tau \right. \\
&\left. - \sum_{n=0}^{+\infty} \lambda_n \mathfrak{e}_\Phi(t; \lambda_n) Q_n g_n \right\|_{\ell^2(\mathbf{m})}^2 \\
&2 \leq \sum_{n=1}^{+\infty} \left((\lambda_n \mathfrak{e}_\Phi(\theta_n(h); \lambda_n) - \lambda_n \mathfrak{e}_\Phi(t; \lambda_n))^2 \right. \\
&\left. + \overline{v}^2(t) \left(\frac{1}{h} \int_0^h (\mathfrak{e}_\Phi(s; \lambda_n) - 1)^2 \mathrm{d}s \right) \right) g_n^2.
\end{aligned}$$

Let us show that the series in the previous equation normally converges for fixed $t > 0$. To do this, just observe that

$$\begin{aligned}
&\sum_{n=1}^{+\infty} \left((\lambda_n \mathfrak{e}_\Phi(\theta_n(h); \lambda_n) - \lambda_n \mathfrak{e}_\Phi(t; \lambda_n))^2 \right. \\
&\left. + \overline{v}^2(t) \left(\frac{1}{h} \int_0^h (\mathfrak{e}_\Phi(s; \lambda_n) - 1)^2 \mathrm{d}s \right) \right) g_n^2 \\
&\leq (2K^2(t) + \overline{v}^2(t)) \|g\|_{\ell^2(\mathbf{m})}.
\end{aligned}$$

Thus, we can take the limit inside the summation sign to obtain

$$\begin{aligned}
&\lim_{h \to 0^+} \sum_{n=1}^{+\infty} \left((\lambda_n \mathfrak{e}_\Phi(\theta_n(h); \lambda_n) - \lambda_n \mathfrak{e}_\Phi(t; \lambda_n))^2 \right. \\
&\left. + \overline{v}^2(t) \left(\frac{1}{h} \int_0^h (\mathfrak{e}_\Phi(s; \lambda_n) - 1)^2 \mathrm{d}s \right) \right) g_n^2 \\
&= \sum_{n=1}^{+\infty} \left(\lambda_n^2 \lim_{h \to 0^+} (\mathfrak{e}_\Phi(\theta_n(h); \lambda_n) - \mathfrak{e}_\Phi(t; \lambda_n))^2 \right. \\
&\left. + \overline{v}^2(t) \lim_{h \to 0^+} \left(\frac{1}{h} \int_0^h (\mathfrak{e}_\Phi(s; \lambda_n) - 1)^2 \mathrm{d}s \right) \right) g_n^2 = 0,
\end{aligned}$$

where we used the fact that $\varepsilon_\Phi(t; \lambda_n)$ is continuous. Last relation implies

$$
\lim_{h \to 0^+} \left\| \frac{1}{h} \int_0^{t+h} (u(t + h - \tau, \cdot) - u(t - \tau, \cdot)) \bar{v}(\tau) d\tau \right.
$$
$$
\left. - \sum_{n=0}^{+\infty} \lambda_n \varepsilon_\Phi(t; \lambda_n) Q_n g_n \right\|_{\ell^2(\mathbf{m})}^2 = 0. \tag{19}
$$

Combining Eqs. (17) and (19) we finally obtain

$$
\lim_{h \to 0^+} \left\| \frac{1}{h} \left(\int_0^{t+h} (u(t + h - \tau, \cdot) - u(0+, \cdot)) \bar{v}(\tau) d\tau \right. \right.
$$
$$
\left. \left. - \int_0^t (u(t - \tau, \cdot) - u(0+, \cdot)) \bar{v}(\tau) d\tau \right) - \sum_{n=0}^{+\infty} \lambda_n \varepsilon_\Phi(t; \lambda_n) Q_n g_n \right\|_{\ell^2(\mathbf{m})} = 0. \tag{20}
$$

Concerning the case $h < 0$, after observing that

$$
\int_0^{t+h} (u(t + h - \tau, y) - u(0+, y)) \bar{v}(\tau) d\tau - \int_0^t (u(t - \tau, y) - u(0+, y)) \bar{v}(\tau) d\tau
$$
$$
= \int_0^{t+h} (u(t + h - \tau, y) - u(t - \tau, y)) \bar{v}(\tau) d\tau
$$
$$
+ \int_{t+h}^t (u(t - \tau, y) - u(0+, y)) \bar{v}(\tau) d\tau,
$$

one can conclude with the same argument as in the case $h > 0$ that

$$
\lim_{h \to 0^-} \left\| \frac{1}{h} \left(\int_0^{t+h} (u(t + h - \tau, \cdot) - u(0+, \cdot)) \bar{v}(\tau) d\tau \right. \right.
$$
$$
\left. \left. - \int_0^t (u(t - \tau, \cdot) - u(0+, \cdot)) \bar{v}(\tau) d\tau \right) - \sum_{n=0}^{+\infty} \lambda_n \varepsilon_\Phi(t; \lambda_n) Q_n g_n \right\|_{\ell^2(\mathbf{m})} = 0. \tag{21}
$$

Combining Eqs. (20) and (21) we finally obtain Eq. (16).

Now let us show that $p_\Phi(t, x; y)$ is the fundamental solution. Let us consider a function $g \in \ell^2(\mathbf{m})$. Then we have

$$\sum_{x \in E} p_\Phi(t, x; y) g(x) = \sum_{x \in E} m(x) \left(\sum_{n \in E} \mathfrak{e}_\Phi(t; \lambda_n) Q_n(x) Q_n(y) \right) g(x)$$

$$= \sum_{n \in E} Q_n(y) \mathfrak{e}_\Phi(t; \lambda_n) \sum_{x \in E} m(x) Q_n(x) g(x)$$

$$= \sum_{n \in E} Q_n(y) \mathfrak{e}_\Phi(t; \lambda_n) g_n = u(t, y)$$

where we could exchange the order of the series since or E is finite, and then the sums are finite, or E is countably infinite and all the series involved are normally convergent in compact sets containing t.

Finally, let us observe that if, for some $z \in E$, $g(x) = \delta_z(x)$, then

$$u(t, y) = \sum_{x \in E} p_\Phi(t, x; y) \delta_z(x) = p_\Phi(t, z; y),$$

concluding the proof.   □

4.3 The Forward Equation

Now let us apply the same strategy to study the Cauchy problem associated with $\mathcal{L}$.

Theorem 3 *Let $N(t)$ be a solvable birth–death process with state space E, forward operator $\mathcal{L}$, invariant measure $\mathbf{m}$ and family of associated classical orthogonal polynomials $(P_n)_{n \in E}$. Let $f/m \in \ell^2(\mathbf{m})$ such that $f(x) = m(x) \sum_{n \in E} f_n Q_n(x)$ for $x \in E$ and some constants $(f_n)_{n \in E}$, where $(Q_n)_{n \in E}$ are the normalized orthogonal polynomials. Then the Cauchy problem*

$$\begin{cases} \frac{\partial^\Phi v}{\partial t^\Phi}(t, x) = \mathcal{L}v(t, x) & t > 0, \ x \in E \\ v(0, x) = f(x) & x \in E \end{cases} \tag{22}$$

admits a unique strong solution of the form

$$v(t, x) = m(x) \sum_{n \in E} \mathfrak{e}_\Phi(t; \lambda_n) f_n Q_n(x), \tag{23}$$

such that $\sup_{t \geq 0} \|v(t, \cdot)\|_{\ell^2(\mathbf{m})} \leq \|f/m\|_{\ell^2(\mathbf{m})}$. Finally $p_\Phi(t, x; y)$ is the fundamental solution of (13), in the sense that it is the strong solution of (13) for $f(x) = \delta_y(x)$ and for any $f/m \in \ell^2(\mathbf{m})$ it holds

$$v(t, x) = \sum_{y \in E} p_\Phi(t, x; y) f(y).$$

Proof Let us first observe, by Lemma 1

$$\mathcal{L}m(x)\mathfrak{e}_\Phi(t;\lambda_n)f_nQ_n(x) = \mathfrak{e}_\Phi(t;\lambda_n)f_n\mathcal{L}_{z\to x}(m(z)Q_n(z))(x)$$

$$= \lambda_n\mathfrak{e}_\Phi(t;\lambda_n)f_nm(x)Q_n(x) = \frac{\partial^\Phi}{\partial t^\Phi}m(x)\mathfrak{e}_\Phi(t;\lambda_n)f_nQ_n(x).$$

Thus, the exact same strategy of the proof of Theorem 2 leads to formula (23) and uniqueness follows as before. Concerning the estimate on the norm, it holds, since **m** is a probability measure and then $m(x) \leq 1$,

$$\|v(t,x)\|^2_{\ell^2(\mathbf{m})} = \sum_{x\in E} m(x)\left(m(x)\sum_{n\in E}\mathfrak{e}_\Phi(t;\lambda_n)f_nQ_n(x)\right)^2$$

$$\leq \sum_{x\in E} m(x)\left(\sum_{n\in E}f_n\mathfrak{e}_\Phi(t;\lambda_n)Q_n(x)\right)^2$$

$$= \sum_{n\in E}f_n^2\mathfrak{e}_\Phi^2(t;\lambda_n) \leq \|f/m\|^2_{\ell^2(\mathbf{m})}.$$

Moreover, let us observe that

$$\sum_{y\in E} p_\Phi(t,x;y)f(y) = m(x)\sum_{y\in E}\left(\sum_{n\in E}\mathfrak{e}_\Phi(t;\lambda_n)Q_n(x)Q_n(y)\right)\frac{f(y)}{m(y)}m(y)$$

$$= m(x)\sum_{n\in E}\mathfrak{e}_\Phi(t;\lambda_n)Q_n(x)\sum_{y\in E}Q_n(y)\frac{f(y)}{m(y)}m(y)$$

$$= m(x)\sum_{n\in E}\mathfrak{e}_\Phi(t;\lambda_n)f_nQ_n(x) = v(t,x).$$

Finally, observe that if, for some fixed $z \in E$, $f(y) = \delta_z(y)$, then obviously $f/m \in \ell^2(\mathbf{m})$ and we have

$$v(t,x) = \sum_{y\in E} p_\Phi(t,x;y)f(y) = \sum_{y\in E} p_\Phi(t,x;y)\delta_z(y) = p_\Phi(t,x;z)$$

concluding the proof. □

Remark 1 One also obtains the following estimate on the norm:

$$\sup_{t\geq 0}\|v(t,\cdot)/m(\cdot)\|_{\ell^2(\mathbf{m})} \leq \|f/m\|_{\ell^2(\mathbf{m})}.$$

Next step is to identify some processes such that $p_\Phi(t,x;y)$ is its *transition probability function* (in some sense) and then give some stochastic representation of the strong solutions of the Cauchy problems (14) and (23).

5 Non-local Solvable Birth–Death Processes

Let us now consider a solvable birth–death process N and the subordinator σ_Φ associated to the Bernstein function Φ, with its inverse E_Φ. Let us suppose N and E_Φ are independent.

Definition 4 The non-local solvable birth–death process induced by N and Φ is defined as

$$N_\Phi(t) := N(E_\Phi(t)) \tag{24}$$

for any $t \geq 0$, where $E_\Phi(t)$ is independent of $N(t)$. Moreover, we define its transition probability function

$$p_\Phi(t, x; y) := \mathbb{P}(N_\Phi(t) = x | N_\Phi(0) = y).$$

We used the same notation for the transition probability function and the fundamental solution of the Cauchy problems (14) and (23). Indeed, we can show the following result.

Theorem 4 *The transition probability function* $p_\Phi(t, x; y)$ *coincides with the series defined in Lemma* 2.

Proof Let us first observe that $N_\Phi(0) = N(0)$ almost surely (since $E_\Phi(0) = 0$ almost surely), thus, by a simple conditioning argument, we have

$$p_\Phi(t, x; y) = \int_0^{+\infty} p(s, x; y) f_\Phi(s, t) ds.$$

Now let us recall, by Theorem 1, that

$$p(s, x; y) = m(x) \sum_{n \in E} e^{\lambda_n t} Q_n(x) Q_n(y)$$

thus we have

$$p_\Phi(t, x; y) = m(x) \int_0^{+\infty} \sum_{n \in E} e^{\lambda_n s} Q_n(x) Q_n(y) f_\Phi(s, t) ds.$$

If E is finite, we conclude the proof. So let us suppose $E = \mathbb{N}_0$. Recalling the notation in the proof of Lemma 2, let us define

$$n_0 = \lceil \max(\text{root}(x) \cup \text{root}(y)) \rceil + 1.$$

and let us split the summation in two parts. We have

$$p_\Phi(t, x; y) = m(x) \int_0^{+\infty} \sum_{n=0}^{n_0} Q_n(x) Q_n(y) e^{\lambda_n s} f_\Phi(s, t) \mathrm{d}s$$

$$+ m(x) \int_0^{+\infty} \sum_{n=n_0+1}^{+\infty} Q_n(x) Q_n(y) e^{\lambda_n s} f_\Phi(s, t) \mathrm{d}s.$$

Now let us observe that we can exchange the integral sign with the first summation by linearity of the integral, while in the second summation this can be done by Fubini's theorem, being the integrands of fixed sign (exactly the sign is determined by $(-1)^{x+y}$ since $Q_n(x) Q_n(y) = \widetilde{m}(n) P_x(n) P_y(n)$). Thus, we have

$$p_\Phi(t, x; y) = m(x) \sum_{n=0}^{n_0} Q_n(x) Q_n(y) \int_0^{+\infty} e^{\lambda_n s} f_\Phi(s, t) \mathrm{d}s$$

$$+ m(x) \sum_{n=n_0+1}^{+\infty} Q_n(x) Q_n(y) \int_0^{+\infty} e^{\lambda_n s} f_\Phi(s, t) \mathrm{d}s$$

$$= m(x) \sum_{n=0}^{+\infty} Q_n(x) Q_n(y) \int_0^{+\infty} e^{\lambda_n s} f_\Phi(s, t) \mathrm{d}s.$$

Finally, let us recall that, by definition

$$\int_0^{+\infty} e^{\lambda_n s} f_\Phi(s, t) \mathrm{d}s = \mathbb{E}[e^{\lambda_n E_\Phi(t)}] = \mathfrak{e}_\Phi(t; \lambda_n),$$

concluding the proof. □

By using this result, we can achieve some stochastic representation of the solutions of the Cauchy problems (14) and (23). Indeed we have the following result.

Proposition 5 *Let $N_\Phi(t)$ be the non-local solvable birth–death process associated to $N(t)$ and Φ. Suppose $N(t)$ admits state space E. Then*

1. *For $g \in \ell^2(\mathbf{m})$ the function $u(t, y) = \mathbb{E}_y[g(N_\Phi(t))]$ (where $\mathbb{E}_y[\cdot] = \mathbb{E}[\cdot | N_\Phi(0) = y]$) is strong solution of* (14);
2. *For $f/m \in \ell^2(\mathbf{m})$ such that $f \geq 0$ and $\sum_{x \in E} f(x) = 1$, denoting by $\mathbb{P}_f$ the probability measure obtained by $\mathbb{P}$ conditioning with the fact that $N_\Phi(0)$ admits distribution f, then the function $v(t, x) = \mathbb{P}_f(N_\Phi(t) = x)$ is strong solution of* (23).

Proof To show assertion 1, let us observe that

$$u(t, y) = \mathbb{E}_y[g(N_\Phi(t))] = \sum_{x \in E} g(x) p_\Phi(t, x; y),$$

obtaining that $u(t, y)$ is the strong solution of (13) by Theorem 2.

Concerning assertion 2, we have

$$v(t, x) = \mathbb{P}_f(N_\Phi(t) = x) = \sum_{y \in E} f(y) p_\Phi(t, x; y)$$

obtaining that $v(t, x)$ is the strong solution of (22) by Theorem 3. $\qquad \square$

Finally, this proposition allows us to obtain the invariant distribution of $N_\Phi(t)$ and show that it is also the limit distribution.

Corollary 3 *Let $N_\Phi(t)$ be the non-local solvable birth–death process associated to $N(t)$ and Φ. Suppose $N(t)$ admits state space E. Then*

1. *If $N_\Phi(0)$ admits distribution $\mathbf{m}$, then $N_\Phi(t)$ admits distribution $\mathbf{m}$ for any $t \geq 0$;*
2. *If $N_\Phi(0)$ admits distribution f such that $f/m \in \ell^2(\mathbf{m})$, then $\lim_{t \to +\infty} \mathbb{P}_f(N_\Phi(t) = x) = m(x)$.*

Proof Let us first show property (1). To do this, let us observe that $1 \in \ell^2(\mathbf{m})$ since $\mathbf{m}$ is a probability measure. Thus, recall, by Proposition 5, that $v(t, x) = \mathbb{P}_{\mathbf{m}}(N_\Phi(t) = x)$ is a strong solution of (22). Now we need to determine m_n such that $\sum_{n \in E} m_n Q_n(x) = 1$. Let us recall that $Q_0(x) = 1$ while $\deg(Q_n(x)) = n$ for any $n = 1, \ldots, n_E$, thus we have $m_0 = 1$ and $m_n = 0$ for any $n = 1, \ldots, n_E$. Moreover, let us recall that $\lambda_0 = 0$, since $1 \in Ker(\mathcal{G})$. Hence we have, by Theorem 3,

$$v(t, x) = m(x) \sum_{n \in E} \mathfrak{e}_\Phi(t; \lambda_n) Q_n(x) m_n = m(x)$$

and $N_\Phi(t)$ admits $\mathbf{m}$ as distribution.

Now let us suppose $N_\Phi(0)$ admits f as distribution with $f/m \in \ell^2(\mathbf{m})$. By Proposition 5 we have that $v(t, x) = \mathbb{P}_f(N_\Phi(t) = x)$ is a strong solution of (22) hence it holds

$$v(t, x) = m(x) \sum_{n \in E} \mathfrak{e}_\Phi(t; \lambda_n) Q_n(x) f_n = m(x) f_0 + m(x) \sum_{\substack{n \in E \\ n \geq 1}} \mathfrak{e}_\Phi(t; \lambda_n) Q_n(x) f_n.$$

Let us determine f_0. We have, by definition of scalar product in $\ell^2(\mathbf{m})$,

$$f_0 = \sum_{x \in E} m(x) \frac{f(x)}{m(x)} Q_0 = \sum_{x \in E} f(x) = 1$$

hence we have

$$v(t, x) = m(x) + m(x) \sum_{\substack{n \in E \\ n \geq 1}} \mathfrak{e}_\Phi(t; \lambda_n) Q_n(x) f_n.$$

Now let us consider the summation part. First of all, let us recall that $\mathbf{m}$ is a probability measure, hence $m(x) \leq 1$. We have

$$\sum_{x \in E} m(x) f^2(x) \leq \sum_{x \in E} m(x) \frac{f^2(x)}{m^2(x)} = \| f/m \|_{\ell^2(\mathbf{m})}$$

hence $f \in \ell^2(\mathbf{m})$. By Lemma 3 we know that $\sum_{\substack{n \in E \\ n \geq 1}} \mathfrak{e}_\Phi(t; \lambda_n) Q_n(x) f_n$ normally converges, hence we can take the limit as $t \to +\infty$ under the summation sign.

Now let us observe that $\lim_{t \to +\infty} E_\Phi(t) = +\infty$ almost surely. On the other hand, we have $e^{\lambda_n E_\Phi(t)} \leq 1$, thus we can use monotone convergence theorem to achieve

$$\lim_{t \to +\infty} \mathfrak{e}_\Phi(t; \lambda_n) = \mathbb{E}[\lim_{t \to +\infty} e^{\lambda_n E_\Phi(t)}] = 0.$$

Finally, by dominated convergence theorem, we have

$$\lim_{t \to +\infty} v(t, x) = m(x) + m(x) \lim_{t \to +\infty} \sum_{\substack{n \in E \\ n \geq 1}} \mathfrak{e}_\Phi(t; \lambda_n) Q_n(x) f_n$$

$$= m(x) + m(x) \sum_{\substack{n \in E \\ n \geq 1}} Q_n(x) f_n \lim_{t \to +\infty} \mathfrak{e}_\Phi(t; \lambda_n) = m(x),$$

concluding the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

6 Correlation Structure of Non-local Solvable Birth–Death Processes

Let us consider the potential measure $U_\Phi(t) = \mathbb{E}[E_\Phi(t)]$ of the subordinator $\sigma_\Phi(t)$. As a consequence of Corollary 2, we can directly apply [9, Theorem 2] to obtain and expression of the covariance of $N_\Phi(t)$ in terms of $\mathfrak{e}_\Phi(t; \lambda_1)$. In particular, with an easy refinement of the proof, by using the distributional derivative of the potential function $U_\Phi(t)$, we can get rid of some hypotheses of the aforementioned Theorem.

Proposition 6 *Let $N(t)$ be a solvable birth–death process with state space E, invariant measure $\mathbf{m}$ and family of associated classical orthogonal polynomials $(P_n)_{n \in E}$. Let us denote by $\iota : E \to E$ the identity function and suppose that $\iota = a_0 + a_1 Q_1$. Then it holds, for any $t \geq s \geq 0$*

$$Cov_m(N_\Phi(t), N_\Phi(s))$$
$$= a_1^2 \left(-\lambda_1 \int_0^s \mathfrak{e}_\Phi(t - \tau; \lambda_1) dU_\Phi(\tau) - 2 + 2\mathfrak{e}_\Phi(s; \lambda_1) + \mathfrak{e}_\Phi(t; \lambda_1) \right).$$

As we expected, since we are composing a stationary process $N(t)$ (if $N(0)$ admits $\mathbf{m}$ as distribution) with a non-stationary one $E_\Phi(t)$, $N_\Phi(t)$ is not second-order stationary. To introduce the notion of memory for our process $N_\Phi(t)$, we refer to the necessary

conditions given in [12, Lemmas 2.1 and 2.2], since, for our processes, it is easier to work with the auto-covariance function. In particular we focus on the *long memory with respect to the initial state*.

Definition 5 Given the function $\gamma(n) = Cov_m(N_\Phi(n), N_\Phi(0))$ for $n \in \mathbb{N}$:

- $N_\Phi(t)$ is said to exhibit long-range dependence if $\gamma(n) \sim \ell(n) n^{-k}$ where $\ell(n)$ is a slowly varying function and $k \in (0, 1)$;
- $N_\Phi(t)$ is said to exhibit short-range dependence if $\sum_{n=1}^{+\infty} |\gamma(n)| < +\infty$.

In the specific case of the initial state, we have the following Proposition.

Proposition 7 *Let $N(t)$ be a solvable birth–death process with state space E, invariant measure **m** and family of associated classical orthogonal polynomials $(P_n)_{n \in E}$. Let us denote by $\iota : E \to E$ the identity function and suppose that $\iota = a_0 + a_1 Q_1$. Then, $\lim_{t \to +\infty} Cov_m(N_\Phi(t), N_\Phi(0)) = 0$.*

Moreover, if $\mathfrak{e}_\Phi(t; \lambda_1) \sim C t^{-\alpha}$ as $t \to +\infty$ for some $\alpha \in (0, 1)$, then $N_\Phi(t)$ is long-range dependent.

Proof One has

$$Cov_m(N_\Phi(t), N_\Phi(0)) = a_1^2 \mathfrak{e}_\Phi(t; \lambda_1),$$

where this identity can be achieved by direct calculations, observing that $Cov_m(N(t), N(0)) = a_1^2 e^{\lambda_1 t}$. The second assertion easily follows from last identity. □

Remark 2 Let us observe that, actually, since $Cov_m(N(t), N(0)) = a_1^2 e^{\lambda_1 t}$, we can argue by direct calculations without using the regularity hypotheses on $\mathbb{P}(E_\Phi(t) \geq s)$.

Concerning the asymptotic behaviour of $\mathfrak{e}_\Phi(t; \lambda_1)$, one can obtain some information by the behaviour of Φ. Indeed one has the following result.

Proposition 8 *If Φ is regularly varying at 0^+ with order $\alpha \in (0, 1)$, then, for any fixed $\lambda > 0$, $\mathfrak{e}_\Phi(t; -\lambda)$ is regularly varying at ∞ with order $-\alpha$.*

Proof Let us consider $J(t) = \int_0^t \mathfrak{e}_\Phi(s; -\lambda) ds$ and let us observe that the Laplace-Stieltjes transform $\bar{J}(\eta)$ of $J(t)$ is given by

$$\bar{J}(\eta) = \frac{\Phi(\eta)}{\eta(\Phi(\eta) + \lambda)}.$$

Since λ is positive, one has that $\bar{J}(\eta)$ is regularly varying at 0^+ if and only if $\frac{\Phi(\eta)}{\eta}$ is (since $\Phi(0^+) = 0$ by the fact that $a = 0$, see, for instance, [45]). In particular, since Φ is regularly varying at 0^+ with order α, then $\bar{J}(\eta)$ is regularly varying at 0^+ with order $\alpha - 1$.

By Karamata's Tauberian theorem (see, for instance, [37] for a compact statement and [16] for the full proof), we know that $J(t)$ is regularly varying at infinity with order $1 - \alpha$. Now let us observe that $J'(t) = \mathfrak{e}_\Phi(t; -\lambda)$, that is monotone, hence, by Monotone density theorem (again, see [16,37]), we have that $\mathfrak{e}_\Phi(t; -\lambda)$ is regularly varying at ∞ of order $-\alpha$, concluding the proof. □

Remark 3 Let us observe that the results coincides with [36, Theorem 2.1−(4)], except for the fact that we are not asking Φ to be special or ν to be absolutely continuous and we are not considering any hypothesis on the existence or regularity of the density of the subordinator.

In particular we get the following Corollary.

Corollary 4 *Under the hypotheses of Proposition 7, if Φ is regularly varying at 0^+ with order $\alpha \in (0, 1)$, then $N_\Phi(t)$ is long-range dependent.*

Let us reconsider the examples given in Sect. 3 and study the asymptotic behaviour of the covariance.

– In the case $\Phi(\lambda) = \lambda^\alpha$ we can actually obtain an explicit formulation of the autocovariance function for $t \geq s > 0$:

$$
\begin{aligned}
&Cov_m(N_\Phi(t), N_\Phi(s)) \\
&= a_1^2 \left(E_\alpha(\lambda_1 t^\alpha) - \frac{\lambda_1 \alpha t^\alpha}{\Gamma(1+\alpha)} \int_0^{\frac{s}{t}} \frac{E_\alpha(\lambda_1 t^\alpha(1-z)^\alpha)}{z^{1-\alpha}} dz \right),
\end{aligned}
$$

where the proof of such formula is identical to the one of [31, Theorem 3.1]. Thus, we can use [31, Remark 3.3] to obtain directly long-range dependence of the process.

– As $\Phi(\lambda) = (\lambda + \theta)^\alpha - \theta^\alpha$, we have that $\Phi(\lambda)$ is actually regularly varying at 0^+ of order 1. Indeed we have $\Phi(\lambda)/\lambda \to \alpha\theta^{\alpha+1}$ as $\lambda \to 0^+$. In particular this means that the function $\overline{J}(\eta)$ defined in Proposition 8 is such that $\overline{J}(0+) = \frac{\alpha\theta^{\alpha+1}}{\lambda}$. By Karamata's Tauberian theorem, we still have $\lim_{t\to+\infty} J(t) = \frac{\alpha\theta^{\alpha+1}}{\lambda}$. This means in particular that $\mathfrak{e}_\Phi(s; -\lambda)$ is integrable in $(0, +\infty)$ and thus $\sum_{n=1}^{+\infty} \mathfrak{e}_\Phi(n; -\lambda) < +\infty$. From this observation we obtain that in such case $N_\Phi(t)$ is short-range dependent.

– If we consider $\Phi(\lambda) = \log(1+\lambda^\alpha)$ for $\alpha \in (0, 1)$, then $\Phi(\lambda)$ is regularly varying at 0^+ of order α, since $\lim_{\lambda\to 0^+} \frac{\log(1+\lambda^\alpha)}{\lambda^\alpha} = 1$. Thus, in particular, N_Φ is long-range dependent by Corollary 4.

– Finally, if $\Phi(\lambda) = \log(1+\lambda)$, we have $\lim_{\lambda\to 0^+} \frac{\log(1+\lambda)}{\lambda} = 1$ and then $\overline{J}(0+) = \frac{1}{\lambda}$. This means again that $\mathfrak{e}_\Phi(t; -\lambda)$ is integrable in $(0, +\infty)$ and then N_Φ is short-range dependent.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Afendras, G., Balakrishnan, N., Papadatos, N.: Orthogonal polynomials in the cumulative Ord family and its application to variance bounds. Statistics **52**(2), 364–392 (2018)
2. Albanese, C., Kuznetsov, A.: Affine lattice models. Int. J. Theor. Appl. Finance **8**(02), 223–238 (2005)
3. Arendt, W., Batty, C.J., Hieber, M., Neubrander, F.: Vector-Valued Laplace Transforms and Cauchy Problems. Springer, Berlin (2011)
4. Ascione, G.: Abstract Cauchy problems for the generalized fractional calculus (2021) NonLinear Analysis (Accepted)
5. Ascione, G., Leonenko, N., Pirozzi, E.: On the transient behaviour of $M/M/\infty$ fractional queues, nonlocal and fractional operators. SEMA SIMAI Springer Series, 26, in press, (2021)
6. Ascione, G., Leonenko, N., Pirozzi, E.: Fractional queues with catastrophes and their transient behaviour. Mathematics **6**(9), 159 (2018)
7. Ascione, G., Leonenko, N., Pirozzi, E.: Fractional Erlang queues. Stoch. Process. Appl. **130**(6), 3249–3276 (2020)
8. Ascione, G., Leonenko, N., Pirozzi, E.: Fractional immigration-death processes. J. Math. Anal. Appl. **495**(2), 124768 (2021)
9. Ascione, G., Toaldo, B.: A semi-Markov leaky integrate-and-fire model. Mathematics **7**(11), 1022 (2019)
10. Beghin, L., Orsingher, E., et al.: Fractional Poisson processes and related planar random motions. Electron. J. Probab. **14**, 1790–1826 (2009)
11. Beghin, L., Orsingher, E., et al.: Poisson-type processes governed by fractional and higher-order recursive differential equations. Electron. J. Probab. **15**, 684–709 (2010)
12. Beran, J., Feng, Y., Ghosh, S., Kulik, R.: Long-Memory Processes. Springer, Berlin (2016)
13. Bertoin, J.: Lévy Processes, vol. 121. Cambridge University Press, Cambridge (1996)
14. Bertoin, J.: Subordinators: Examples and Applications. In: Lectures on Probability Theory and Statistics, pp. 1–91. Springer (1999)
15. Bingham, N.H.: Limit theorems for occupation times of Markov processes. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **17**(1), 1–22 (1971)
16. Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular Variation, vol. 27. Cambridge University Press, Cambridge (1989)
17. Cahoy, D.O., Polito, F., Phoha, V.: Transient behavior of fractional queues and related processes. Methodol. Comput. Appl. Probab. **17**(3), 739–759 (2015)
18. Cannon, R.H.: Dynamics of Physical Systems. Courier Corporation (2003)
19. Forman, J.L., Sørensen, M.: The Pearson diffusions: a class of statistically tractable diffusion processes. Scand. J. Stat. **35**(3), 438–465 (2008)
20. Gajda, J., Wyłomańska, A.: Time-changed Ornstein-Uhlenbeck process. J. Phys. A Math. Theor. **48**(13), 135004 (2015)
21. Johnson, N.L., Kemp, A.W., Kotz, S.: Univariate Discrete Distributions, vol. 444. Wiley, Hoboken (2005)
22. Karlin, S., McGregor, J.: The classification of birth and death processes. Trans. Am. Math. Soc. **86**(2), 366–400 (1957)
23. Karlin, S., McGregor, J.: Linear growth, birth and death processes. J. Math. Mech. **7**, 643–662 (1958)
24. Karlin, S., McGregor, J.: Ehrenfest urn models. J. Appl. Probab. **2**(2), 352–376 (1965)
25. Karlin, S., McGregor, J.L.: The differential equations of birth-and-death processes, and the Stieltjes moment problem. Trans. Am. Math. Soc. **85**(2), 489–546 (1957)

26. Kleinrock, L.: Queueing Systems: Volume 1: Theory. A Wiley-Interscience Publication (1975)
27. Kochubei, A.N.: General fractional calculus, evolution equations, and renewal processes. Integr. Eqn. Oper. Theory **71**(4), 583–600 (2011)
28. Kochubei, A.N., Kondratiev, Y.: Growth equation of the general fractional calculus. Mathematics **7**(7), 615 (2019)
29. Kolokol'tsov, V.N.: Mixed fractional differential equations and generalized operator-valued Mittag-Leffler functions. Math. Notes **106**(5–6), 740–756 (2019)
30. Kumar, A., Vellaisamy, P.: Inverse tempered stable subordinators. Stat. Probab. Lett. **103**, 134–141 (2015)
31. Leonenko, N.N., Meerschaert, M.M., Sikorskii, A.: Correlation structure of fractional Pearson diffusions. Comput. Math. Appl. **66**(5), 737–745 (2013)
32. Leonenko, N.N., Meerschaert, M.M., Sikorskii, A.: Fractional Pearson diffusions. J. Math. Anal. Appl. **403**(2), 532–546 (2013)
33. Leonenko, N.N., Papić, I., Sikorskii, A., Šuvak, N.: Heavy-tailed fractional Pearson diffusions. Stoch. Process. Appl. **127**(11), 3512–3535 (2017)
34. Meerschaert, M.M., Scheffler, H.P.: Triangular array limits for continuous time random walks. Stoch. Process. Appl. **118**(9), 1606–1633 (2008)
35. Meerschaert, M.M., Straka, P.: Inverse stable subordinators. Math. Model. Nat. Phenom. **8**(2), 1–16 (2013)
36. Meerschaert, M.M., Toaldo, B.: Relaxation patterns and semi-Markov dynamics. Stoch. Process. Appl. **129**(8), 2850–2879 (2019)
37. Mikosch, T.: Regular Variation, Subexponentiality and Their Applications in Probability Theory. Eindhoven University of Technology (1999)
38. Nikiforov, A.F., Uvarov, V.B., Suslov, S.K.: Classical Orthogonal Polynomials of a Discrete Variable. Springer, Berlin (1991)
39. Nowak, M.A.: Evolutionary Dynamics: Exploring the Equations of Life. Harvard University Press, Cambridge (2006)
40. Orsingher, E., Polito, F., Sakhno, L.: Fractional non-linear, linear and sublinear death processes. J. Stat. Phys. **141**(1), 68–93 (2010)
41. Orsingher, E., Polito, F., et al.: Fractional pure birth processes. Bernoulli **16**(3), 858–881 (2010)
42. Orsingher, E., Polito, F., et al.: On a fractional linear birth-death process. Bernoulli **17**(1), 114–137 (2011)
43. Rudin, W.: Principles of Mathematical Analysis, vol. 3. McGraw-Hill, New York (1964)
44. Rudin, W.: Real and Complex Analysis. Tata McGraw-Hill Education, New York (2006)
45. Schilling, R.L., Song, R., Vondracek, Z.: Bernstein Functions: Theory and Applications, vol. 37. Walter de Gruyter, Berlin (2012)
46. Schoutens, W.: Stochastic Processes and Orthogonal Polynomials, vol. 146. Springer, Berlin (2012)
47. Šikić, H., Song, R., Vondraček, Z.: Potential theory of geometric stable processes. Probab. Theory Relat. Fields **135**(4), 547–575 (2006)
48. Simon, T., et al.: Comparing Fréchet and positive stable laws. Electron. J. Probab. **19**, 1–25 (2014)
49. Toaldo, B.: Convolution-type derivatives, hitting-times of subordinators and time-changed C^0-semigroups. Potential Anal. **42**(1), 115–140 (2015)