ALEA, Lat. Am. J. Probab. Math. Stat. 12 (1), 35–53 (2015)

The Mittag–Leffler process and a scaling limit for the block counting process of the Bolthausen–Sznitman coalescent

Martin Möhle

Mathematisches Institut
Eberhard Karls Universität Tübingen
Auf der Morgenstelle 10
72076 Tübingen, Germany
E-mail address: martin.moehle@uni-tuebingen.de
URL: http://www.math.uni-tuebingen.de/~moehle

Abstract. The Mittag–Leffler process $X = (X_t)_{t\geq 0}$ is introduced. This Markov process has the property that its marginal random variables X_t are Mittag–Leffler distributed with parameter e^{-t}, $t \in [0,\infty)$, and the semigroup $(T_t)_{t\geq 0}$ of X satisfies $T_t f(x) = \mathbb{E}(f(x^{e^{-t}} X_t))$ for all $x \geq 0$ and all bounded measurable functions $f : [0,\infty) \to \mathbb{R}$. Further characteristics of the process X are derived, for example an explicit formula for the joint moments of its finite-dimensional distributions. The Mittag–Leffler process turns out to be Siegmund dual to Neveu's continuous-state branching process. The main result states that the block counting process of the Bolthausen–Sznitman n-coalescent, properly scaled, converges in the Skorohod topology to the Mittag–Leffler process X as the sample size n tends to infinity. We provide an equivalent version of this convergence result involving stable distributions.

1. Introduction and main results

Exchangeable coalescent processes with multiple collisions are Markov processes with state space $\mathcal{P}$, the set of partitions of $\mathbb{N} := \{1, 2, \ldots\}$. During each transition blocks merge together to form a single block. These processes are characterized by a measure Λ on the unit interval $[0, 1]$. For more information on these processes we refer the reader to Pitman (1999) and Sagitov (1999). The Bolthausen–Sznitman coalescent (Bolthausen and Sznitman (1998)) is the particular Λ-coalescent $\Pi = (\Pi_t)_{t\geq 0}$ with Λ being the uniform distribution on $[0, 1]$. In this article we focus on

Received by the editors October 27, 2014; accepted January 13, 2015.
2010 *Mathematics Subject Classification.* 60F05, 60J27, 92D15, 97K60.
Key words and phrases. Block counting process; Bolthausen–Sznitman coalescent; continuous-state branching process; marginal distributions; Mittag–Leffler process; stable distribution; weak convergence.

the process $\Pi^{(n)} = (\Pi_t^{(n)})_{t\ge 0}$ of the Bolthausen–Sznitman coalescent Π restricted to a sample of size $n \in \mathbb{N}$. We are in particular interested in the process $N^{(n)} := (N_t^{(n)})_{t\ge 0}$, where $N_t^{(n)}$ denotes the number of blocks of $\Pi_t^{(n)}$. The process $N^{(n)}$ is called the block counting process of the Bolthausen–Sznitman n-coalescent. It is well known that $N^{(n)}$ is a time-homogeneous Markov chain with generator $Q = (q_{ij})_{1\le i,j\le n}$ having entries $q_{ij} := i/((i-j)(i-j+1))$ if $i > j$, $q_{ij} := 1-i$ if $i = j$ and $q_{ij} = 0$ if $i < j$. For $n \in \mathbb{N}$ and $t \in [0,\infty)$ define

$$X_t^{(n)} \; := \; \frac{N_t^{(n)}}{n^{e^{-t}}}. \tag{1.1}$$

We call $X^{(n)} := (X_t^{(n)})_{t\ge 0}$ the scaled block counting process of the Bolthausen–Sznitman n-coalescent. Similar power law scalings of the form n^α occur for example when studying the number of occupied boxes of certain urn models with infinitely many boxes. For more details we refer the reader to Section 10 of the survey of Gnedin et al. (2007). The scaling $n^{e^{-t}}$ in (1.1) is somewhat unusual since it involves not only the parameter n but also the time parameter t. Clearly, $X^{(n)}$ is a Markov process with state space $E := [0,\infty)$, however, since the scaling depends on t, $X^{(n)}$ is time-inhomogeneous. Our main result (Theorem 1.1 below) provides a distributional limiting result for $X^{(n)}$ as the sample size n tends to infinity. The arising limiting Markov process $X = (X_t)_{t\ge 0}$ we call the Mittag–Leffler process, since the marginal random variable X_t turns out to be Mittag–Leffler distributed with parameter e^{-t}. Note that the distribution of X_t is uniquely determined by its entire moments $\mathbb{E}(X_t^m) = m!/\Gamma(1+me^{-t})$, $m \in \mathbb{N}_0 := \{0,1,2,\ldots\}$. For detailed information on the Mittag–Leffler distribution and on the process X we refer the reader to Section 2, where the existence of X is established and fundamental properties of this process are derived. In order to wipe out possible confusion with processes in the literature having similar names we mention that the process X has nothing in common with the autoregressive Mittag–Leffler process studied for example by Jayakumar (2003) and Jayakumar and Pillai (1993). These processes are based on the (heavy-tailed) Mittag–Leffler distribution of the first type (see, for example, Mainardi and Gorenflo (2000) and Pillai (1990) for some related works), whereas the Mittag–Leffler distributed random variable X_t is of the second type and has finite moments of all orders. Let us now present our main convergence result.

Theorem 1.1. *For the Bolthausen–Sznitman coalescent the scaled block counting process $X^{(n)} = (X_t^{(n)})_{t\ge 0}$ defined via (1.1) converges in $D_E[0,\infty)$ as $n \to \infty$ to the Mittag–Leffler process $X = (X_t)_{t\ge 0}$ introduced in Section 2.*

Remark 1.2. 1. The proof of Theorem 1.1 is provided in Section 4. For an equivalent formulation of Theorem 1.1 involving α-stable distributions we refer the reader to Theorem 2.6.

2. Theorem 1.1 can be also stated logarithmically as follows. The process $(\log N_t^{(n)} - e^{-t}\log n)_{t\ge 0}$ converges in $D_{\mathbb{R}}[0,\infty)$ to the process $(\log X_t)_{t\ge 0}$ as $n \to \infty$. Neither X nor $(\log X_t)_{t\ge 0}$ is a Lévy process. Note that the logarithmic block counting process $(\log N_t^{(n)})_{t\ge 0}$ plays an important role in the problem of whether a coalescent process comes down from infinity, see, for example, Section 4 of Limic (2012).

3. Note that $X^{(n)}$ is time-inhomogeneous whereas the limiting process X is time-homogeneous. Thus, Theorem 1.1 in particular states that $X^{(n)}$ is asymptotically time-homogeneous.

The article is organized as follows. Section 2 is devoted to the Mittag–Leffler process X. We prove the existence of this process and derive fundamental properties of X, among them representations for the semigroup of X (see (2.6)) and an explicit formula for the joint moments (see Lemma 2.2) of the finite-dimensional distributions of X. We also discuss the Siegmund dual of the process X, which turns out to be Neveu's continuous-state branching process, leading to an equivalent formulation of Theorem 1.1 stated in Theorem 2.6. In Section 3 we provide some fundamental formulas (see Lemma 3.1 and Lemma 3.2) for certain moments of the block counting process $N^{(n)}$ of the Bolthausen–Sznitman n-coalescent. These results rely on the spectral decomposition (Möhle and Pitters (2014)) of the generator of the block counting process. Lemma 3.1 in particular shows that $N_t^{(n)}$ has mean

$$\mathbb{E}(N_t^{(n)}) \;=\; \frac{\Gamma(n+e^{-t})}{\Gamma(n)\Gamma(1+e^{-t})} \qquad n\in\mathbb{N}, t\in[0,\infty). \tag{1.2}$$

For large n (1.2) is asymptotically equal to $n^{e^{-t}}(\Gamma(1+e^{-t}))^{-1} = n^{e^{-t}}\mathbb{E}(X_t)$, which indicates that $n^{e^{-t}}$ is the appropriate scaling in order to obtain a non-degenerate limiting process for the scaled block counting process as n tends to infinity. In the final Section 4 this argument is made rigorous leading to a proof of Theorem 1.1. First the convergence of the finite-dimensional distributions is verified and afterwards the convergence in $D_E[0,\infty)$ is established.

Recently Baur and Bertoin (2015) independently obtained closely related results on essentially the same topic via fragmentation of recursive trees.

We leave it open for future work to establish convergence results in analogy to Theorem 1.1 for the block counting process $N^{(n)}$ of more general coalescent processes (that do not come down from infinity), for example for the $\beta(a,b)$-coalescent with $a\ge 1$ (and $b>0$).

2. The Mittag–Leffler process

Before we come to the Mittag–Leffler process let us briefly mention some well known results concerning the Mittag–Leffler distribution. Let $\eta=\eta(\alpha)$ be a random variable being Mittag–Leffler distributed with parameter $\alpha\in[0,1]$. Note that η has moments

$$\mathbb{E}(\eta^m) \;=\; \frac{\Gamma(1+m)}{\Gamma(1+m\alpha)}, \qquad m\in[0,\infty),$$

and that the entire moments $\mathbb{E}(\eta^m)$, $m\in\mathbb{N}_0$, uniquely determine the distribution of η. Clearly, η is standard exponentially distributed for $\alpha=0$ and $\mathbb{P}(\eta=1)=1$ for $\alpha=1$.

If $\alpha_n\to\alpha$, then the moments of $\eta(\alpha_n)$ converge to those of $\eta(\alpha)$, which implies the convergence $\eta(\alpha_n)\to\eta(\alpha)$ in distribution as $n\to\infty$. Thus, the map $\alpha\mapsto\mathbb{P}_{\eta(\alpha)}$ is a continuous function from $[0,1]$ to the space $\mathcal{P}(E)$ of probability measures on $E:=[0,\infty)$ equipped with the topology of convergence in distribution.

For $\alpha\in(0,1)$ the Mittag–Leffler distribution can be characterized in terms of an exponential integral of a particular subordinator as follows. Let $S=(S_t)_{t\ge 0}$

be a drift-free subordinator with killing rate $k := 1/\Gamma(1-\alpha)$ and Lévy measure ϱ having density

$$u \ \mapsto\ \frac{1}{\Gamma(1-\alpha)}\frac{e^{-u/\alpha}}{(1-e^{-u/\alpha})^{\alpha+1}}, \qquad u\in(0,\infty), \tag{2.1}$$

with respect to Lebesgue measure on $(0,\infty)$. It is readily checked (see Lemma 5.1 in the appendix) that S has Laplace exponent

$$\Phi(x) \ =\ \frac{\Gamma(1+\alpha x)}{\Gamma(1-\alpha+\alpha x)}, \qquad x\in[0,\infty). \tag{2.2}$$

The distribution of the exponential integral $I := \int_0^\infty e^{-S_t}\mathrm{d}t$ is uniquely determined (see Carmona et al. (1997)) via its entire moments

$$\mathbb{E}(I^m) \ =\ \frac{m!}{\Phi(1)\cdots\Phi(m)} \ =\ m!\prod_{j=1}^m \frac{\Gamma(1+(j-1)\alpha)}{\Gamma(1+j\alpha)} \ =\ \frac{\Gamma(1+m)}{\Gamma(1+m\alpha)}, \quad m\in\mathbb{N}.$$

Thus, I is Mittag–Leffler distributed with parameter α.

2.1. *Existence of the Mittag–Leffler process.* In this subsection we prove the existence of a particular Markov process $X=(X_t)_{t\ge 0}$ having sample paths in $D_E[0,\infty)$ such that every X_t is Mittag–Leffler distributed with parameter e^{-t}. Constructing Markov processes with given marginal distributions has attained some interest in the literature, mainly in the context of (semi)martingales. We exemplary refer the reader to Madan and Yor (2002) and the references therein. Note however, that the process X we are going to construct will be neither a supermartingale nor a submartingale.

For $t\in[0,\infty)$ let η_t be a random variable being Mittag–Leffler distributed with parameter e^{-t}. Define $p:[0,\infty)\times E\times\mathcal{B}(E)\to[0,1]$ via

$$p(t,x,B) \ :=\ \mathbb{E}(1_B(x^{e^{-t}}\eta_t)) \ =\ \mathbb{P}(x^{e^{-t}}\eta_t\in B). \tag{2.3}$$

The definition of p is such that for all $(t,x)\in[0,\infty)\times E$ the random variable $x^{e^{-t}}\eta_t$ has distribution $p(t,x,.)$. In particular, $p(t,x,.)$ has moments

$$\int_E y^m p(t,x,\mathrm{d}y) \ =\ \mathbb{E}((x^{e^{-t}}\eta_t)^m) \ =\ x^{me^{-t}}\frac{\Gamma(1+m)}{\Gamma(1+me^{-t})}, \quad m\in[0,\infty), \tag{2.4}$$

and the entire moments $\int_E y^m p(t,x,\mathrm{d}y)$, $m\in\mathbb{N}_0$, uniquely determine the distribution $p(t,x,.)$. In order to verify the Chapman–Kolmogorov property

$$p(s+t,x,B) \ =\ \int_E p(s,y,B)\,p(t,x,\mathrm{d}y), \quad s,t\in[0,\infty), x\in E, B\in\mathcal{B}(E), \tag{2.5}$$

fix $s,t\in[0,\infty)$ and $x\in E$. For all $B\in\mathcal{B}(E)$ define $\mu_1(B) := p(s+t,x,B)$ and $\mu_2(B) := \int_E p(s,y,B)\,p(t,x,\mathrm{d}y)$. Clearly, μ_1 and μ_2 are probability measures on E. By (2.4), μ_1 has moments

$$\int_E z^m\,\mu_1(\mathrm{d}z) \ =\ \int_E z^m p(s+t,x,\mathrm{d}z) \ =\ x^{me^{-(s+t)}}\frac{\Gamma(1+m)}{\Gamma(1+me^{-(s+t)})}, \quad m\in\mathbb{N}_0,$$

and these moments uniquely determine μ_1. By Fubini's theorem and (2.4), μ_2 has moments

$$\begin{aligned}
\int_E z^m \,\mu_2(\mathrm{d}z) &= \int_E z^m \int_E p(s,y,\mathrm{d}z)\,p(t,x,\mathrm{d}y)\\
&= \int_E \bigg(\int_E z^m p(s,y,\mathrm{d}z)\bigg)\,p(t,x,\mathrm{d}y)\\
&= \int_E y^{me^{-s}}\frac{\Gamma(1+m)}{\Gamma(1+me^{-s})}\,p(t,x,\mathrm{d}y)\\
&= \frac{\Gamma(1+m)}{\Gamma(1+me^{-s})}\int_E y^{me^{-s}}\,p(t,x,\mathrm{d}y)\\
&= \frac{\Gamma(1+m)}{\Gamma(1+me^{-s})}x^{me^{-s}e^{-t}}\frac{\Gamma(1+me^{-s})}{\Gamma(1+me^{-s}e^{-t})}\\
&= x^{me^{-(s+t)}}\frac{\Gamma(1+m)}{\Gamma(1+me^{-(s+t)})},
\end{aligned}$$

and these moments uniquely determine μ_2. Since the moments of μ_1 and μ_2 coincide, it follows that $\mu_1 = \mu_2$ and the Chapman–Kolmogorov property (2.5) is established. Thus, the family $(T_t)_{t\ge 0}$ of linear operators T_t, defined via

$$T_t f(x) \;:=\; \int_E f(y)\,p(t,x,\mathrm{d}y) \;=\; \mathbb{E}(f(x^{e^{-t}}\eta_t)),\; t\in[0,\infty), f\in B(E), x\in E, \quad (2.6)$$

defines a semigroup on $B(E)$, the set of bounded measurable functions $f : E \to \mathbb{R}$ equipped with the supremum norm $\|f\| := \sup_{x\in E}|f(x)|$. Note that (2.6) is also well defined for some unbounded functions, for example for all polynomials $f : E \to \mathbb{R}$. The semigroup $(T_t)_{t\ge 0}$ on $B(E)$ is clearly conservative, since $T_t 1 = 1$ for all $t \in [0,\infty)$. We have $\|T_t f\| = \sup_{x\in E}|\mathbb{E}(f(x^{e^{-t}}\eta_t))| \le \sup_{x\in E}\mathbb{E}(|f(x^{e^{-t}}\eta_t)|) \le \|f\|$ for all $t\in[0,\infty)$ and all $f \in B(E)$. Thus, $\|T_t\| \le 1$ for all $t \in [0,\infty)$, so the semigroup $(T_t)_{t\ge 0}$ is contracting. Moreover, $(T_t)_{t\ge 0}$ is obviously positive meaning that each operator T_t maps nonnegative functions (in $B(E)$) to nonnegative functions.

Let $\widehat{C}(E) \subseteq B(E)$ denote the Banach space of continuous functions $f : E \to \mathbb{R}$ vanishing at infinity. Using the dominated convergence theorem it is easily seen (see Lemma 5.3 in the appendix) that $T_t\widehat{C}(E) \subseteq \widehat{C}(E)$ for all $t \in [0,\infty)$. With some more effort (see again Lemma 5.3) it can be shown by exploiting the theorem of Heine that, for all $f \in \widehat{C}(E)$, $T_t f(x) \to f(x)$ as $t \to 0$ uniformly for all $x \in E$. Therefore, $(T_t)_{t\ge 0}$ is strongly continuous on $\widehat{C}(E)$, thus a Feller semigroup on $\widehat{C}(E)$. Hence (see, for example, Ethier and Kurtz (1986, p. 169, Theorem 2.7)) there exists a Markov process $X = (X_t)_{t\ge 0}$ corresponding to $(T_t)_{t\ge 0}$ with initial distribution $\mathbb{P}(X_0 = 1) = 1$ and sample paths in the space $D_E[0,\infty)$ of right continuous functions $x : [0,\infty) \to E$ with left limits equipped with the Skorohod topology. Note that $\mathbb{E}(f(X_{s+t})\,|\,X_u, u \le s) = T_t f(X_s)$ for all $f \in B(E)$ and all $s,t \in [0,\infty)$ and that

$$\mathbb{P}(X_{s+t}\in B\,|\,X_u, u\le s) \;=\; p(t,X_s,B), \qquad s,t\in[0,\infty), B\in\mathcal{B}(E).$$

From $\mathbb{E}(f(X_t)) = \mathbb{E}(f(X_t)\,|\,X_0=1) = T_t f(1) = \mathbb{E}(f(\eta_t))$, $f \in B(E)$, $t \in [0,\infty)$, we conclude that X_t has the same distribution as η_t, so X_t is Mittag–Leffler distributed with parameter e^{-t}. We therefore call X the *Mittag–Leffler process*.

Clearly, $X_t \to X_\infty$ in distribution as $t \to \infty$, where X_∞ is standard exponentially distributed. Thus, the stationary distribution of X is the standard exponential distribution.

Remark 2.1. The Chapman–Kolmogorov property holds whenever the random variable η_t introduced at the beginning of the construction in this subsection has moments of the form $\mathbb{E}(\eta_t^m) = h(m)/h(me^{-t})$ for some given function $h : [0,\infty) \to (0,\infty)$ and if these moments uniquely determine the distribution of η_t. We have carried out the construction for $h(x) := \Gamma(1+x)$ leading to the Mittag–Leffler process.

More generally, one may use other functions h, for example $h(x) := \Gamma(ax+b)$ for some constants $a, b \in (0,\infty)$, leading to a construction of a wider class of Markov processes $X = (X_t)_{t\geq 0}$. In this case X_t has moments $\mathbb{E}(X_t^m) = h(m)/h(me^{-t}) = \Gamma(am+b)/\Gamma(ame^{-t}+b)$, $m \in [0,\infty)$. Letting $t \to \infty$ it follows that the stationary distribution of X has moments $h(m)/h(0) = \Gamma(am+b)/\Gamma(b)$, $m \in \mathbb{N}_0$, and, hence, density $x \mapsto (a\Gamma(b))^{-1}x^{b/a-1}e^{-x^{1/a}}$, $x \in (0,\infty)$, with respect to Lebesgue measure on $(0,\infty)$.

A further simple example is $h(x) := a^x$ for some constant $a \in (0,\infty)$, leading to the deterministic process $X = (X_t)_{t\geq 0}$ with $X_t = a^{1-e^{-t}}$ for all $t \in [0,\infty)$.

2.2. *Further properties of the Mittag–Leffler process.* In this subsection we derive some further properties of the Mittag–Leffler process X. The following lemma provides a formula for the moments of the finite-dimensional distributions of X.

Lemma 2.2 (Moments of the finite-dimensional distributions of X). *Let $k \in \mathbb{N}$, $0 = t_0 \leq t_1 < t_2 < \cdots < t_k$ and $m_1, \ldots, m_k \in [0,\infty)$. For $j \in \{0,\ldots,k\}$ define $x_j := x_j(k) := \sum_{i=j+1}^k m_i e^{-(t_i - t_j)}$. Note that $x_k = 0$ and $x_0 = \sum_{i=1}^k m_i e^{-t_i}$. Then*

$$\mathbb{E}(X_{t_1}^{m_1} \cdots X_{t_k}^{m_k}) \;=\; \prod_{j=1}^k \frac{\Gamma(1+x_j+m_j)}{\Gamma(1+x_{j-1})} \tag{2.7}$$

and the entire moments $\mathbb{E}(X_{t_1}^{m_1} \cdots X_{t_k}^{m_k})$, $m_1, \ldots, m_k \in \mathbb{N}_0$, uniquely determine the distribution of $(X_{t_1}, \ldots, X_{t_k})$. In particular, $\mathbb{E}(X_t^m) = \Gamma(1+m)/\Gamma(1+me^{-t})$, $m \in \mathbb{N}_0$, $t \in [0,\infty)$, and, hence, $\mathbb{E}(X_t) = 1/\Gamma(1+e^{-t})$ and $\mathrm{Var}(X_t) = \mathbb{E}(X_t^2) - (\mathbb{E}(X_t))^2 = 2/\Gamma(1+2e^{-t}) - 1/(\Gamma(1+e^{-t}))^2$, $t \in [0,\infty)$.

Proof: Induction on k. Clearly, (2.7) holds for $k = 1$, since X_{t_1} is Mittag–Leffler distributed with parameter e^{-t_1}. The induction step from $k-1$ to k works as follows. We have

$$\begin{aligned} \mathbb{E}(X_{t_1}^{m_1} \cdots X_{t_k}^{m_k}) &= \mathbb{E}(\mathbb{E}(X_{t_1}^{m_1} \cdots X_{t_k}^{m_k} \,|\, X_{t_1}, \ldots, X_{t_{k-1}})) \\ &= \mathbb{E}(X_{t_1}^{m_1} \cdots X_{t_{k-1}}^{m_{k-1}} \mathbb{E}(X_{t_k}^{m_k} \,|\, X_{t_{k-1}})). \end{aligned}$$

Define $f_m(x) := x^m$ for convenience. Using the formula (2.6) for the semigroup operator T_t, the last conditional expectation is given by

$$\mathbb{E}(X_{t_k}^{m_k} \,|\, X_{t_{k-1}}) \;=\; (T_{t_k - t_{k-1}} f_{m_k})(X_{t_{k-1}}) \;=\; \frac{\Gamma(1+m_k)}{\Gamma(1+m_k e^{-(t_k - t_{k-1})})} X_{t_{k-1}}^{m_k e^{-(t_k - t_{k-1})}}.$$

We therefore obtain

$$\begin{aligned}
&\mathbb{E}(X_{t_1}^{m_1}\cdots X_{t_k}^{m_k})\\
&\quad= \frac{\Gamma(1+m_k)}{\Gamma(1+m_ke^{-(t_k-t_{k-1})})}\mathbb{E}(X_{t_1}^{m_1}\cdots X_{t_{k-2}}^{m_{k-2}}X_{t_{k-1}}^{m_{k-1}+m_ke^{-(t_k-t_{k-1})}})\\
&\quad= \frac{\Gamma(1+x_k+m_k)}{\Gamma(1+x_{k-1})}\mathbb{E}(X_{t_1}^{\tilde{m}_1}\cdots X_{t_{k-1}}^{\tilde{m}_{k-1}}),
\end{aligned}$$

where $\tilde{m}_j := m_j$ for $1 \le j \le k-2$ and $\tilde{m}_{k-1} := m_{k-1} + m_k e^{-(t_k-t_{k-1})}$. By induction,

$$\mathbb{E}(X_{t_1}^{\tilde{m}_1}\cdots X_{t_{k-1}}^{\tilde{m}_{k-1}}) \;=\; \prod_{j=1}^{k-1}\frac{\Gamma(1+y_j+\tilde{m}_j)}{\Gamma(1+y_{j-1})},$$

where $y_j := \sum_{i=j+1}^{k-1}\tilde{m}_i e^{-(t_i-t_j)}$ for all $j\in\{0,\ldots,k-1\}$. The result follows since, for $1\le j\le k-2$,

$$\begin{aligned}
y_j &= \sum_{i=j+1}^{k-2} m_ie^{-(t_i-t_j)} + (m_{k-1}+m_ke^{-(t_k-t_{k-1})})e^{-(t_{k-1}-t_j)}\\
&= \sum_{i=j+1}^{k} m_ie^{-(t_i-t_j)} \;=\; x_j
\end{aligned}$$

and $y_{k-1} = 0$ and, hence, $y_{k-1}+\tilde{m}_{k-1} = \tilde{m}_{k-1} = m_ke^{-(t_k-t_{k-1})} + m_{k-1} = x_{k-1}+m_{k-1}$. □

Remark 2.3. The mean $\mathbb{E}(X_t) = 1/\Gamma(1+e^{-t})$ is increasing for $t < t_0$ and decreasing for $t > t_0$, where $t_0 \approx 0.772987$ is the unique solution of the equation $\Psi(1+e^{-t_0}) = 0$ and $\Psi := \Gamma'/\Gamma$ denotes the logarithmic derivative of the gamma function. The process X has therefore neither non-increasing paths nor non-decreasing paths. In particular, we are not in the context of Haas and Miermont (2011), where essentially all considered processes have non-increasing paths.

Corollary 2.4. *The Mittag–Leffler process $X = (X_t)_{t\ge 0}$ is continuous in probability, i.e. $X_s \to X_t$ in probability as $s \to t$ for every $t \in [0,\infty)$.*

Proof: By Lemma 2.2, for all $s,t\in[0,\infty)$,

$$\mathbb{E}(X_s^2) \;=\; \frac{\Gamma(3)}{\Gamma(1+2e^{-s})} \;\to\; \frac{\Gamma(3)}{\Gamma(1+2e^{-t})} \;=\; \mathbb{E}(X_t^2), \qquad s\to t,$$

and

$$\mathbb{E}(X_sX_t) \;=\; \frac{\Gamma(2+e^{-|t-s|})}{\Gamma(1+e^{-s}+e^{-t})}\frac{\Gamma(2)}{\Gamma(1+e^{-|t-s|})} \;\to\; \frac{\Gamma(3)}{\Gamma(1+2e^{-t})} \;=\; \mathbb{E}(X_t^2), \quad s\to t.$$

It follows that $\mathbb{E}((X_s-X_t)^2) = \mathbb{E}(X_s^2) - 2\mathbb{E}(X_sX_t) + \mathbb{E}(X_t^2) \to 0$ as $s\to t$. Thus, for all $\varepsilon > 0$, $\mathbb{P}(|X_s - X_t| \ge \varepsilon) \le \mathbb{E}((X_s-X_t)^2)/\varepsilon^2 \to 0$ as $s \to t$. □

Remark 2.5. Note that (the Mittag–Leffler distributed random variable) X_t is not infinitely divisible. Moreover, X does not have independent increments. In particular, X is not a Lévy process. The process $(\log X_t)_{t\ge 0}$ is as well not a Lévy process, since this process does not have independent increments either. This can be also seen as follows. The Fourier transform $\phi_t(x) := \mathbb{E}(e^{ix\log X_t}) = \mathbb{E}(X_t^{ix}) = \Gamma(1+ix)/\Gamma(1+ixe^{-t})$, $x\in\mathbb{R}$, of $\log X_t$ is not the t-th power of $\phi_1(x)$.

We leave a possible construction of the Mittag–Leffler process via Lévy processes or subordinators, for example as a random time change and/or by taking the absolute value of a certain Lévy process, for future work. For related functionals of this type (local time processes, Bessel-type processes) we refer the reader to James (2010) and the references therein.

We finally provide in this subsection some information on the generator A of the Mittag–Leffler process X, but we will not use the generator A in our further considerations. Suppose that $f \in B(E)$ is infinitely often differentiable and that f satisfies $f(y) = \sum_{k=0}^\infty (f^{(k)}(x)/k!)(y-x)^k$ for all $x, y \in E$. Then

$$\frac{T_t f(x) - f(x)}{t} \;=\; \sum_{k=1}^\infty \frac{f^{(k)}(x)}{k!} \frac{\mathbb{E}((x^{e^{-t}}\eta_t - x)^k)}{t}.$$

Let $\Psi := \Gamma'/\Gamma$ denote the logarithmic derivative of the gamma function. Since

$$a_k(x) \;:=\; \lim_{t\searrow 0} \frac{\mathbb{E}((x^{e^{-t}}\eta_t - x)^k)}{t} \;=\; \begin{cases} x\Psi(2) - x\log x & \text{for } k = 1,\\ \dfrac{(-x)^k}{k-1} & \text{for } k \in \mathbb{N}\setminus\{1\}, \end{cases} \tag{2.8}$$

the generator A of X satisfies

$$Af(x) \;=\; \sum_{k=1}^\infty \frac{f^{(k)}(x)}{k!} a_k(x)$$

with $a_k(x)$ defined via (2.8). Since $f(xy) = \sum_{k=0}^\infty (f^{(k)}(x)/k!)(-x)^k(1-y)^k$ it follows that

$$Af(x) \;=\; a_1(x)f'(x) + \int_0^1 \frac{f(xy) - f(x) + x(1-y)f'(x)}{(1-y)^2}\,\mathrm{d}y.$$

The substitution $h = x(1-y)$ yields

$$Af(x) \;=\; a_1(x)f'(x) + x\int_0^x \frac{f(x-h) - f(x) + hf'(x)}{h^2}\,\mathrm{d}h.$$

2.3. *Duality to Neveu's continuous state branching process.* For every $t \in [0,\infty)$ and $y \in E$ the map $x \mapsto \mathbb{P}(X_{s+t} \le y \,|\, X_s = x) = \mathbb{E}(1_{[0,y]}(x^{e^{-t}}\eta_t)) = \mathbb{P}(x^{e^{-t}}\eta_t \le y)$ is non-increasing in $x \in E$, i.e. X is stochastically monotone. Via duality (Siegmund (1976)) it follows that there exists a Markov process $Y = (Y_t)_{t\ge 0}$ with state space $E = [0,\infty)$, which is dual to X with respect to the duality function $H : E^2 \to \{0,1\}$ defined via $H(x,y) := 1$ if $x \le y$ and $H(x,y) := 0$ otherwise. Note that Y has transition mechanism $\mathbb{P}(Y_t \ge x \,|\, Y_0 = y) = \mathbb{P}(X_t \le y \,|\, X_0 = x) = \mathbb{P}(x^{e^{-t}}\eta_t \le y)$, $t \in [0,\infty)$, $x, y \in E$. The semigroup $(S_t)_{t\ge 0}$ of Y is hence given by

$$S_t g(y) \;=\; \mathbb{E}(g(y^{e^t}/\eta_t^{e^t})) \;=\; \mathbb{E}(g(y^{e^t}\xi_t)), \qquad t \in [0,\infty), g \in B(E), y \in E, \tag{2.9}$$

where $\alpha := e^{-t}$ and $\xi_t := \eta_t^{-e^t} = \eta_t^{-1/\alpha}$. It is well known (see, for example, Feller (1971, p. 453)) that ξ_t is α-stable with Laplace transform $\lambda \mapsto \mathbb{E}(e^{-\lambda\xi_t}) = e^{-\lambda^\alpha}$, $\lambda \in [0,\infty)$. Choosing $g(y) := e^{-\lambda y}$, $y \in E$, in (2.9) it follows that Y_t, conditional on $Y_0 = y$, has Laplace transform $\mathbb{E}(e^{-\lambda Y_t} \,|\, Y_0 = y) = S_t g(y) = \mathbb{E}(g(y^{e^t}\xi_t)) = \mathbb{E}(e^{-\lambda y^{e^t}\xi_t}) = e^{-(\lambda y^{e^t})^\alpha} = e^{-y\lambda^\alpha}$, $\lambda \in [0,\infty)$. Thus, we identify Y as the continuous-state branching process (CSBP) of Neveu (1992) with Lamperti branching mechanism (see, for example, Bertoin and Le Gall (2000, Section

3)) $u \mapsto u \log u$. In summary we have just verified that Neveu's CSBP is the Siegmund dual of the Mittag–Leffler process. Note that, for all $m < \alpha = e^{-t}$, $\mathbb{E}(Y_t^m \,|\, Y_0 = y) = y^{me^t}\mathbb{E}(\xi_t^m) = y^{me^t}\Gamma(1-me^t)/\Gamma(1-m)$. On the state space $[0,\infty]$ (including the point ∞) Y_t converges in distribution as $t \to \infty$ to a random variable Y_∞ satisfying $\mathbb{P}(Y_\infty = 0 \,|\, Y_0 = y) = e^{-y}$ and $\mathbb{P}(Y_\infty = \infty \,|\, Y_0 = y) = 1-e^{-y}$.

Alternatively one may derive this duality result via Möhle (2013, Proposition 2.2) as follows. The set C_1 of all non-negative, non-increasing, left-continuous functions $f : E \to \mathbb{R}$ satisfying $f(0) = 1$ and $\lim_{x\to\infty} f(x) = 0$ is a cone of X. Moreover, each $f \in C_1$ has a unique integral representation over E with respect to H, namely $f(x) = \int_E H(x,y)\,Q_f(\mathrm{d}y)$, $x \in E$, where the probability measure Q_f on $(E, \mathcal{B}(E))$ is (uniquely) defined via $Q_f([x,\infty)) := f(x)$, $x \in E$. Thus, Proposition 2.2 of Möhle (2013) is applicable, which yields the existence of the desired dual Markov process Y.

If η is Mittag–Leffler distributed with parameter $\alpha \in (0,1]$, then (recall Feller (1971, p. 453)) $\xi := \eta^{-1/\alpha}$ is α-stable with Laplace transform $\lambda \mapsto \mathbb{E}(e^{-\lambda\xi}) = e^{-\lambda^\alpha}$, $\lambda \in [0,\infty)$. This observation leads to the following equivalent formulation of Theorem 1.1.

Theorem 2.6. *For the Bolthausen–Sznitman coalescent, the stochastic process* $Z^{(n)} := (Z_t^{(n)})_{t\ge 0}$, *defined via* $Z_t^{(n)} := (X_t^{(n)})^{-e^t} = n/(N_t^{(n)})^{e^t}$ *for all* $n \in \mathbb{N}$ *and* $t \in [0,\infty)$, *converges in* $D_E[0,\infty)$ *as* $n \to \infty$ *to the time-inhomogeneous process* $Z = (Z_t)_{t\ge 0}$ *with initial distribution* $\mathbb{P}(Z_0 = 1) = 1$ *and defined via* $Z_t := X_t^{-e^t}$ *for all* $t \in [0,\infty)$.

Theorem 2.6 follows from Theorem 1.1 by applying the continuous mapping theorem to the sequence of processes $X^{(n)} = (X_t^{(n)})_{t\ge 0}$, $n \in \mathbb{N}$, and the continuous map $h : D_{(0,\infty)}[0,\infty) \to D_E[0,\infty)$, defined via $h(x) := (x_t^{-e^t})_{t\ge 0}$ for all $x = (x_t)_{t\ge 0} \in D_{(0,\infty)}[0,\infty)$.

For every $t \in [0,\infty)$ the random variable Z_t has the same distribution as ξ_t and is hence α-stable with Laplace transform $\lambda \mapsto \mathbb{E}(e^{-\lambda Z_t}) = e^{-\lambda^\alpha}$, $\lambda \in [0,\infty)$, where $\alpha := e^{-t}$.

3. Moment calculations

In this section we provide formulas for certain moments of the block counting process $N^{(n)} = (N_t^{(n)})_{t\ge 0}$ of the Bolthausen–Sznitman n-coalescent. In the following we use for $x \in (0,\infty)$ and $m \in [0,\infty)$ the notation $[x]_m := \Gamma(x+m)/\Gamma(x)$. Note that for $m \in \mathbb{N}_0$ the symbol $[x]_m = x(x+1)\cdots(x+m-1)$ coincides with the ascending factorial. The following lemma provides an explicit formula for the expectation of $[N_t^{(n)}]_m$.

Lemma 3.1. *Fix* $n \in \mathbb{N}$ *and* $t \in [0,\infty)$. *For the Bolthausen–Sznitman coalescent the random variable* $N_t^{(n)}$ *satisfies for all* $m \in [0,\infty)$

$$
\begin{aligned}
\mathbb{E}([N_t^{(n)}]_m) &= \Gamma(m+1)\prod_{j=1}^{n-1}\frac{j+me^{-t}}{j}\\
&= \Gamma(m+1)\binom{n-1+me^{-t}}{n-1} \;=\; \frac{\Gamma(m+1)}{\Gamma(1+me^{-t})}[n]_{me^{-t}}.
\end{aligned}
$$

In particular,

$$\begin{aligned}\mathbb{E}(N_t^{(n)}) &= \prod_{j=1}^{n-1}\frac{j+e^{-t}}{j} = \binom{n-1+e^{-t}}{n-1}\\ &= \frac{1}{\Gamma(1+e^{-t})}[n]_{e^{-t}} = \frac{\Gamma(n+e^{-t})}{\Gamma(n)\Gamma(1+e^{-t})}\end{aligned}$$

and

$$\mathrm{Var}(N_t^{(n)}) = 2\prod_{j=1}^{n-1}\frac{j+2e^{-t}}{j} - \prod_{j=1}^{n-1}\frac{j+e^{-t}}{j} - \left(\prod_{j=1}^{n-1}\frac{j+e^{-t}}{j}\right)^2.$$

Proof: Fix $n \in \mathbb{N}$ and $t \in [0,\infty)$. The formula obviously holds for $m = 0$. Thus, we can assume that $m \in (0,\infty)$. Clearly, $\mathbb{E}([N_t^{(n)}]_m) = \sum_{j=1}^n [j]_m p_{nj}(t)$, where $p_{nj}(t) := \mathbb{P}(N_t^{(n)} = j)$. In the following $s(.,.)$ and $S(.,.)$ denote the Stirling numbers of the first and second kind respectively. Plugging in

$$p_{nj}(t) = (-1)^{n+j}\frac{\Gamma(j)}{\Gamma(n)}\sum_{k=j}^n e^{-(k-1)t}s(n,k)S(k,j)$$

(see Möhle and Pitters (2014, Corollary 1.3), Equation (1.3), corrected by an obviously missing sign factor $(-1)^{k+j}$) it follows that

$$\begin{aligned}\mathbb{E}([N_t^{(n)}]_m) &= \sum_{j=1}^n [j]_m(-1)^{n+j}\frac{\Gamma(j)}{\Gamma(n)}\sum_{k=j}^n e^{-(k-1)t}s(n,k)S(k,j)\\ &= \frac{(-1)^n}{\Gamma(n)}e^t\sum_{k=1}^n s(n,k)(e^{-t})^k\sum_{j=1}^k \Gamma(j+m)(-1)^jS(k,j).\end{aligned}$$

Since $\Gamma(j+m)(-1)^j = \Gamma(m)[m]_j(-1)^j = \Gamma(m)(-m)(-m-1)\cdots(-m-j+1) = \Gamma(m)(-m)_j$, where $(x)_j := x(x-1)\cdots(x-j+1)$, the last sum simplifies to $\sum_{j=1}^k \Gamma(j+m)(-1)^jS(k,j) = \Gamma(m)\sum_{j=1}^k(-m)_jS(k,j) = \Gamma(m)(-m)^k$. Thus,

$$\begin{aligned}\mathbb{E}([N_t^{(n)}]_m) &= \frac{(-1)^n}{\Gamma(n)}e^t\sum_{k=1}^n s(n,k)(e^{-t})^k\Gamma(m)(-m)^k\\ &= \Gamma(m)\frac{(-1)^n}{\Gamma(n)}e^t\sum_{k=1}^n s(n,k)(-me^{-t})^k = \Gamma(m)\frac{(-1)^n}{\Gamma(n)}e^t(-me^{-t})_n\\ &= \frac{\Gamma(m)}{\Gamma(n)}e^t[me^{-t}]_n = \Gamma(m+1)\prod_{j=1}^{n-1}\frac{j+me^{-t}}{j}\\ &= \Gamma(m+1)\binom{n-1+me^{-t}}{n-1} = \frac{\Gamma(m+1)}{\Gamma(1+me^{-t})}[n]_{me^{-t}}.\end{aligned}$$

Choosing $m = 1$ the formula for the mean of $N_t^{(n)}$ follows immediately. The formula for the variance of $N_t^{(n)}$ follows from $\mathrm{Var}(N_t^{(n)}) = \mathbb{E}([N_t^{(n)}]_2) - \mathbb{E}(N_t^{(n)}) - (\mathbb{E}(N_t^{(n)}))^2$. □

The following result (Lemma 3.2) is a generalization of Lemma 3.1. It will turn out to be quite useful later in order to verify the main convergence result (Theorem 1.1).

Lemma 3.2. *Let $k \in \mathbb{N}$, $0 = t_0 \le t_1 < t_2 < \cdots < t_k$ and $m_1, \ldots, m_k \in [0, \infty)$. For $j \in \{0, \ldots, k\}$ define $x_j := x_j(k) := \sum_{i=j+1}^k m_i e^{-(t_i - t_j)}$. Note that $0 = x_k \le x_{k-1} \le \cdots \le x_2 \le x_1 \le x_0 = \sum_{i=1}^k m_i e^{-t_i}$. Then*

$$\mathbb{E}\bigg(\prod_{j=1}^k [N_{t_j}^{(n)} + x_j]_{m_j}\bigg) \;=\; [n]_{x_0} \prod_{j=1}^k \frac{\Gamma(1 + x_j + m_j)}{\Gamma(1 + x_{j-1})}. \tag{3.1}$$

Proof: Induction on k. For $k = 1$ the assertion holds by Lemma 3.1. The induction step from $k - 1$ to k (≥ 2) works as follows. We have

$$\begin{aligned}\mathbb{E}\bigg(\prod_{j=1}^k [N_{t_j}^{(n)} + x_j]_{m_j}\bigg) &= \mathbb{E}\bigg(\mathbb{E}\bigg(\prod_{j=1}^k [N_{t_j}^{(n)} + x_j]_{m_j} \bigg| N_{t_1}^{(n)}, \ldots, N_{t_{k-1}}^{(n)}\bigg)\bigg) \\ &= \mathbb{E}\bigg(\prod_{j=1}^{k-1} [N_{t_j}^{(n)} + x_j]_{m_j} \mathbb{E}([N_{t_k}^{(n)}]_{m_k} \mid N_{t_{k-1}}^{(n)})\bigg), \end{aligned} \tag{3.2}$$

since $x_k = 0$. The process $N^{(n)} = (N_t^{(n)})_{t \ge 0}$ is time-homogeneous. Thus, for all $j \in \{1, \ldots, n\}$,

$$\begin{aligned}\mathbb{E}([N_{t_k}^{(n)}]_{m_k} \mid N_{t_{k-1}}^{(n)} = j) &= \mathbb{E}([N_{t_k - t_{k-1}}^{(j)}]_{m_k}) \\ &= \frac{\Gamma(1 + m_k)}{\Gamma(1 + m_k e^{-(t_k - t_{k-1})})} [j]_{m_k e^{-(t_k - t_{k-1})}}, \end{aligned}$$

where the last equality holds by Lemma 3.1. Thus,

$$\mathbb{E}([N_{t_k}^{(n)}]_{m_k} \mid N_{t_{k-1}}^{(n)}) \;=\; \frac{\Gamma(1 + m_k)}{\Gamma(1 + m_k e^{-(t_k - t_{k-1})})} [N_{t_{k-1}}^{(n)}]_{m_k e^{-(t_k - t_{k-1})}}.$$

Plugging this into (3.2) yields

$$\begin{aligned}&\mathbb{E}\bigg(\prod_{j=1}^k [N_{t_j}^{(n)} + x_j]_{m_j}\bigg) \\ &\quad= \frac{\Gamma(1 + m_k)}{\Gamma(1 + m_k e^{-(t_k - t_{k-1})})} \mathbb{E}\bigg(\bigg(\prod_{j=1}^{k-1} [N_{t_j}^{(n)} + x_j]_{m_j}\bigg) [N_{t_{k-1}}^{(n)}]_{m_k e^{-(t_k - t_{k-1})}}\bigg) \\ &\quad= \frac{\Gamma(1 + x_k + m_k)}{\Gamma(1 + x_{k-1})} \mathbb{E}\bigg(\prod_{j=1}^{k-1} [N_{t_j}^{(n)} + y_j]_{\tilde{m}_j}\bigg), \end{aligned}$$

where $y_j := x_j$ and $\tilde{m}_j := m_j$ for $0 \le j \le k-2$, $y_{k-1} := 0$ and $\tilde{m}_{k-1} := m_{k-1} + x_{k-1}$. By induction,

$$\mathbb{E}\bigg(\prod_{j=1}^{k-1} [N_{t_j}^{(n)} + y_j]_{\tilde{m}_j}\bigg) \;=\; [n]_{y_0} \prod_{j=1}^{k-1} \frac{\Gamma(1 + y_j + \tilde{m}_j)}{\Gamma(1 + y_{j-1})} \;=\; [n]_{x_0} \prod_{j=1}^{k-1} \frac{\Gamma(1 + x_j + m_j)}{\Gamma(1 + x_{j-1})},$$

and (3.1) follows immediately, which completes the induction. □

4. Proof of Theorem 1.1

The σ-algebra generated by $X_t^{(n)}$ coincides with the σ-algebra generated by $N_t^{(n)}$. Thus, the Markov property of the block counting process $N^{(n)}$ carries over to the scaled block counting process $X^{(n)}$. Note however that the process $X^{(n)}$ is time-inhomogeneous whereas $N^{(n)}$ is time-homogeneous.

As a warming up we first verify the convergence of the finite-dimensional distributions. Afterwards we turn to the convergence in $D_E[0,\infty)$. Since the proof of the convergence of the one-dimensional distributions turns out to be less technical, we start with a consideration of the one-dimensional distributions. Note that the convergence of the one-dimensional distributions has already been obtained by Pitman (2006), Theorem 5.19 in combination with Theorem 3.8.

Step 1. (Convergence of the one-dimensional distributions) Recall that $S(.,.)$ denote the Stirling numbers of the second kind. Fix $t\in[0,\infty)$. Applying the formula

$$x^m \;=\; \sum_{i=0}^m (-1)^{m-i}S(m,i)[x]_i, \qquad m\in\mathbb{N}_0, \tag{4.1}$$

it follows that

$$\mathbb{E}((X_t^{(n)})^m) \;=\; \frac{1}{n^{me^{-t}}}\mathbb{E}((N_t^{(n)})^m) \;=\; \sum_{i=0}^m (-1)^{m-i}S(m,i)\frac{\mathbb{E}([N_t^{(n)}]_i)}{n^{me^{-t}}},$$

$n\in\mathbb{N}$, $m\in\mathbb{N}_0$. By Lemma 3.1, $\mathbb{E}([N_t^{(n)}]_i) = (\Gamma(i+1)/\Gamma(1+ie^{-t}))[n]_{ie^{-t}} = \mathbb{E}(X_t^i)[n]_{ie^{-t}}$, leading to

$$\mathbb{E}((X_t^{(n)})^m) \;=\; \sum_{i=0}^m (-1)^{m-i}S(m,i)\mathbb{E}(X_t^i)\frac{[n]_{ie^{-t}}}{n^{me^{-t}}}, \qquad n\in\mathbb{N}, m\in\mathbb{N}_0.$$

Letting $n\to\infty$ shows that $\lim_{n\to\infty}\mathbb{E}((X_t^{(n)})^m) = \mathbb{E}(X_t^m)$ for all $m\in\mathbb{N}_0$. This convergence of moments implies (see, for example, Billingsley (1995, Theorems 30.1 and 30.2)) the convergence $X_t^{(n)}\to X_t$ in distribution as $n\to\infty$. Thus, the convergence of the one-dimensional distributions holds.

Step 2. (Convergence of the finite-dimensional distributions) Let us now turn to the convergence of the k-dimensional distributions, $k\in\mathbb{N}$. Fix $0=t_0\le t_1<t_2<\cdots<t_k<\infty$ and $m_1,\ldots,m_k\in[0,\infty)$. For $j\in\{0,\ldots,k\}$ define $x_j:=x_j(k):=\sum_{i=j+1}^k m_i e^{-(t_i-t_j)}$. Note that $x_k=0$ and that $x_0=\sum_{i=1}^k m_i e^{-t_i}$. We have

$$\prod_{j=1}^k\bigg(X_{t_j}^{(n)}+\frac{x_j}{n^{e^{-t_j}}}\bigg)^{m_j} \;=\; \prod_{j=1}^k \frac{(N_{t_j}^{(n)}+x_j)^{m_j}}{n^{m_je^{-t_j}}} \;=\; \frac{1}{n^{x_0}}\prod_{j=1}^k (N_{t_j}^{(n)}+x_j)^{m_j}.$$

Applying (4.1) it follows that

$$\begin{aligned}\prod_{j=1}^k\bigg(X_{t_j}^{(n)}+\frac{x_j}{n^{e^{-t_j}}}\bigg)^{m_j} &= \frac{1}{n^{x_0}}\prod_{j=1}^k\bigg(\sum_{i_j=0}^{m_j}(-1)^{m_j-i_j}S(m_j,i_j)[N_{t_j}^{(n)}+x_j]_{i_j}\bigg)\\ &= \frac{1}{n^{x_0}}\sum_{i_1\le m_1,\ldots,i_k\le m_k}\bigg(\prod_{j=1}^k(-1)^{m_j-i_j}S(m_j,i_j)\bigg)\bigg(\prod_{j=1}^k[N_{t_j}^{(n)}+x_j]_{i_j}\bigg).\end{aligned}$$

Taking expectation yields

$$\mathbb{E}\bigg(\prod_{j=1}^k\bigg(X_{t_j}^{(n)}+\frac{x_j}{n^{e^{-t_j}}}\bigg)^{m_j}\bigg)$$
$$= \sum_{i_1\le m_1,\ldots,i_k\le m_k}\bigg(\prod_{j=1}^k(-1)^{m_j-i_j}S(m_j,i_j)\bigg)\frac{1}{n^{x_0}}\mathbb{E}\bigg(\prod_{j=1}^k[N_{t_j}^{(n)}+x_j]_{i_j}\bigg). \quad (4.2)$$

By Lemma 3.2, the last expectation is $O(n^{\sum_{j=1}^k i_je^{-t_j}})$ and

$$\mathbb{E}(\prod_{j=1}^k[N_{t_j}^{(n)}+x_j]_{m_j}) \;=\; [n]_{x_0}\prod_{j=1}^k\frac{\Gamma(1+x_j+m_j)}{\Gamma(1+x_{j-1})} \;=\; [n]_{x_0}\mathbb{E}(X_{t_1}^{m_1}\cdots X_{t_k}^{m_k}),$$

where the last equality holds by Eq. (2.7) from Lemma 2.2. Thus, letting $n\to\infty$ in (4.2) yields

$$\lim_{n\to\infty}\mathbb{E}\bigg(\prod_{j=1}^k\bigg(X_{t_j}^{(n)}+\frac{x_j}{n^{e^{-t_j}}}\bigg)^{m_j}\bigg) \;=\; \mathbb{E}(X_{t_1}^{m_1}\cdots X_{t_k}^{m_k}). \quad (4.3)$$

In order to get rid of the disturbing fractions $x_j/n^{e^{-t_j}}$ on the left hand side in (4.3) one may use the binomial formula

$$\bigg(X_{t_j}^{(n)}+\frac{x_j}{n^{e^{-t_j}}}\bigg)^{m_j} \;=\; \sum_{l_j=0}^{m_j}\binom{m_j}{l_j}\bigg(\frac{x_j}{n^{e^{-t_j}}}\bigg)^{m_j-l_j}(X_{t_j}^{(n)})^{l_j}$$

and conclude from (4.3) by induction on $m:=m_1+\cdots+m_k\in\mathbb{N}_0$ that

$$\lim_{n\to\infty}\mathbb{E}((X_{t_1}^{(n)})^{m_1}\cdots(X_{t_k}^{(n)})^{m_k}) \;=\; \mathbb{E}(X_{t_1}^{m_1}\cdots X_{t_k}^{m_k}), \qquad m_1,\ldots,m_k\in\mathbb{N}_0. \quad (4.4)$$

This convergence of moments implies (see, for example, Billingsley (1995, Problems 30.5 and 30.6)) the convergence $(X_{t_1}^{(n)},\ldots,X_{t_k}^{(n)})\to(X_{t_1},\ldots,X_{t_k})$ in distribution as $n\to\infty$. Thus, the convergence of the finite-dimensional distributions holds.

Step 3. (Preparing the proof of the convergence in $D_E[0,\infty)$) Let $M(E)$ denote the set of all measurable functions $f:E\to\mathbb{R}$. Define $E_n(s):=\{j/n^{e^{-s}}: j\in\{1,\ldots,n\}\}$ for all $n\in\mathbb{N}$ and all $s\in[0,\infty)$ and

$$T_{s,t}^{(n)}f(x) \;:=\; \mathbb{E}(f(X_{s+t}^{(n)})\,|\,X_s^{(n)}=x), \quad n\in\mathbb{N}, s,t\in[0,\infty), f\in M(E), x\in E_n(s).$$

Note that $(T_{s,t}^{(n)})_{s,t\ge0}$ is the semigroup of the time-inhomogeneous Markov process $X^{(n)}$. Let us verify that, for all $s,t\in[0,\infty)$, all polynomials $p:E\to\mathbb{R}$ and all compact sets $K\subseteq E$,

$$\lim_{n\to\infty}\sup_{x\in E_n(s)\cap K}|T_{s,t}^{(n)}p(x)-T_tp(x)| \;=\; 0. \quad (4.5)$$

For $m\in\mathbb{N}_0$ let $p_m:E\to\mathbb{R}$ denote the m-th monomial defined via $p_m(x):=x^m$, $x\in E$. Fix $s,t\in[0,\infty)$ and a compact set $K\subseteq E$. For $n\in\mathbb{N}$, $m\in\mathbb{N}_0$ and $x\in E_n(s)$ we have

$$T_{s,t}^{(n)}p_m(x) \;=\; \mathbb{E}((X_{s+t}^{(n)})^m\,|\,X_s^{(n)}=x)$$
$$=\; \frac{\mathbb{E}((N_{s+t}^{(n)})^m\,|\,N_s^{(n)}=xn^{e^{-s}})}{n^{me^{-(s+t)}}} \;=\; \frac{\mathbb{E}((N_t^{(xn^{e^{-s}})})^m)}{n^{me^{-(s+t)}}},$$

where the last equality holds since the block counting process $N^{(n)} = (N_t^{(n)})_{t\ge 0}$ is time-homogeneous. By (4.1) and Lemma 3.1 it follows that

$$T_{s,t}^{(n)}p_m(x) \;=\; \sum_{i=0}^m (-1)^{m-i}S(m,i)\frac{\mathbb{E}([N_t^{(xn^{e^{-s}})}]_i)}{n^{me^{-(s+t)}}} \;=\; \sum_{i=0}^m (-1)^{m-i}S(m,i)\mathbb{E}(X_t^i)\frac{[xn^{e^{-s}}]_{ie^{-t}}}{n^{me^{-(s+t)}}}.$$

Since $T_tp_m(x) = \mathbb{E}(p_m(x^{e^{-t}}X_t)) = \mathbb{E}(X_t^m)x^{me^{-t}}$ it follows that

$$T_{s,t}^{(n)}p_m(x) - T_tp_m(x) \;=\; \mathbb{E}(X_t^m)\bigg(\frac{[xn^{e^{-s}}]_{me^{-t}}}{n^{me^{-(s+t)}}} - x^{me^{-t}}\bigg) + \sum_{i=0}^{m-1}(-1)^{m-i}S(m,i)\mathbb{E}(X_t^i)\frac{[xn^{e^{-s}}]_{ie^{-t}}}{n^{me^{-(s+t)}}}.$$

It is straightforward to check that this expression converges uniformly for all $x \in E_n(s)\cap K$ (even uniformly for all x in any compact subset of E) to zero as $n\to\infty$. Therefore, (4.5) holds for every monomial $p := p_m$, $m\in\mathbb{N}_0$, and, by linearity, for all polynomials $p : E\to\mathbb{R}$.

Step 4. (Convergence in $D_E[0,\infty)$) According to a time-inhomogeneous variant of Ethier and Kurtz (1986, p. 167, Theorem 2.5) it suffices to verify that for all $s,t\in[0,\infty)$ and all $f\in\widehat{C}(E)$,

$$\lim_{n\to\infty}\sup_{x\in E_n(s)} |T_{s,t}^{(n)}f(x) - T_tf(x)| \;=\; 0. \tag{4.6}$$

Fix $s,t\in[0,\infty)$ and $f\in\widehat{C}(E)$. Without loss of generality we may assume that $\|f\| > 0$. Let $\varepsilon > 0$. Since $f\in\widehat{C}(E)$ and $T_tf\in\widehat{C}(E)$, there exists a constant $x_0 = x_0(\varepsilon)\ge 1$ such that $|f(x)| < \varepsilon$ and $|T_tf(x)| < \varepsilon$ for all $x > x_0$. Moreover, since $\mathbb{P}(X_{s+t}\le x_0\,|\,X_s = x) = T_t1_{(-\infty,x_0]}(x) = \mathbb{P}(x^{e^{-t}}X_t\le x_0) = \mathbb{P}(X_t\le x_0/x^{e^{-t}})\to \mathbb{P}(X_t\le 0) = 0$ as $x\to\infty$, we can choose a real constant $L = L(\varepsilon)\ge x_0$ sufficiently large such that $\mathbb{P}(X_{s+t}\le x_0\,|\,X_s = x) < \varepsilon/\|f\|$ for all $x\ge L$. For all $n\in\mathbb{N}$ and all $x\in E_n(s)$ we have

$$\begin{aligned} |T_{s,t}^{(n)}f(x)| \;&\le\; \mathbb{E}(|f(X_{s+t}^{(n)})|\,|\,X_s^{(n)} = x)\\ &=\; \mathbb{E}(|f(X_{s+t}^{(n)})|\,1_{\{X_{s+t}^{(n)}>x_0\}}\,|\,X_s^{(n)} = x) + \mathbb{E}(|f(X_{s+t}^{(n)})|\,1_{\{X_{s+t}^{(n)}\le x_0\}}\,|\,X_s^{(n)} = x)\\ &\le\; \varepsilon + \|f\|\,\mathbb{P}(X_{s+t}^{(n)}\le x_0\,|\,X_s^{(n)} = x).\end{aligned}$$

By Step 2, the convergence of the two-dimensional distributions holds. In particular, for every $x\ge 1$, $\mathbb{P}(X_{s+t}^{(n)}\le x_0\,|\,X_s^{(n)} = \lfloor xn^{e^{-s}}\rfloor/n^{e^{-s}})$ converges to $\mathbb{P}(X_{s+t}\le x_0\,|\,X_s = x)$ as $n\to\infty$ pointwise for all $x\ge 1$. Since the map $x\mapsto\mathbb{P}(X_{s+t}\le x_0\,|\,X_s = x) = T_t1_{[0,x_0]}(x) = \mathbb{E}(1_{[0,x_0]}(x^{e^{-t}}X_t)) = \mathbb{P}(x^{e^{-t}}X_t\le x_0)$, $x\ge 1$, is continuous, non-increasing and bounded, this convergence holds even uniformly for all $x\ge 1$. [The proof of this uniform convergence works essentially the same as the proof that pointwise convergence of distribution functions holds even uniformly, if the limiting distribution function is continuous.] Thus, there exists $n_0 = n_0(\varepsilon)\in\mathbb{N}$ such that $\mathbb{P}(X_{s+t}^{(n)}\le x_0\,|\,X_s^{(n)} = x)\le\mathbb{P}(X_{s+t}\le x_0\,|\,X_s = x) + \varepsilon/\|f\|$ for all $n > n_0$ and all $x\in E_n(s)\cap[1,\infty)$. For all $n\in\mathbb{N}$ with $n > n_0$ and all $x\in E_n(s)\cap[L,\infty)$ it follows that

$$\begin{aligned}|T_{s,t}^{(n)}f(x)| &\le \varepsilon+\|f\|\left(\mathbb{P}(X_{s+t}\le x_0\,|\,X_s=x)+\frac{\varepsilon}{\|f\|}\right)\\ &= 2\varepsilon+\|f\|\,\mathbb{P}(X_{s+t}\le x_0\,|\,X_s=x)\ \le\ 3\varepsilon.\end{aligned}$$

Thus, for all $n>n_0$,

$$\begin{aligned}&\sup_{x\in E_n(s)\cap[L,\infty)}|T_{s,t}^{(n)}f(x)-T_tf(x)|\\ &\quad\le \sup_{x\in E_n(s)\cap[L,\infty)}|T_{s,t}^{(n)}f(x)|+\sup_{x\in E_n(s)\cap[L,\infty)}|T_tf(x)|\\ &\quad\le 3\varepsilon+\varepsilon\ =\ 4\varepsilon.\end{aligned}$$

Thus it is shown that

$$\lim_{n\to\infty}\sup_{x\in E_n(s)\cap[L,\infty)}|T_{s,t}^{(n)}f(x)-T_tf(x)|\ =\ 0.$$

Defining $K:=[0,L]$ it remains to verify that

$$\lim_{n\to\infty}\sup_{x\in E_n(s)\cap K}|T_{s,t}^{(n)}f(x)-T_tf(x)|\ =\ 0. \tag{4.7}$$

By the Tschebyscheff-Markov inequality, for all $y>0$ and all $x\in E_n(s)\cap K$,

$$\begin{aligned}T_t1_{(y,\infty)}(x) &= \mathbb{P}(X_{s+t}>y\,|\,X_s=x)\ \le\ \frac{1}{y}\mathbb{E}(X_{s+t}\,|\,X_s=x)\\ &= \frac{1}{y}\mathbb{E}(X_t)x^{e^{-t}}\ \le\ \frac{1}{y}\mathbb{E}(X_t)L^{e^{-t}}.\end{aligned}$$

Moreover, making again use of the Tschebyscheff-Markov inequality and using Lemma 3.1, for all $y>0$ and all $x\in E_n(s)\cap K$,

$$\begin{aligned}T_{s,t}^{(n)}1_{(y,\infty)}(x) &= \mathbb{P}(X_{s+t}^{(n)}>y\,|\,X_s^{(n)}=x)\\ &\le \frac{1}{y}\mathbb{E}(X_{s+t}^{(n)}\,|\,X_s^{(n)}=x)\ =\ \frac{1}{y}\mathbb{E}(X_t)\frac{[xn^{e^{-s}}]_{e^{-t}}}{n^{e^{-(s+t)}}}\\ &\le \frac{1}{y}\mathbb{E}(X_t)\frac{[Ln^{e^{-s}}]_{e^{-t}}}{n^{e^{-(s+t)}}}\ \sim\ \frac{1}{y}\mathbb{E}(X_t)L^{e^{-t}},\qquad n\to\infty.\end{aligned}$$

Thus, we can choose a real constant $y_0=y_0(\varepsilon)\ge x_0$ (which may depend on s, t and L but not on n) sufficiently large such that

$$T_t1_{(y_0,\infty)}(x)\ \le\ \varepsilon\quad\text{and}\quad T_{s,t}^{(n)}1_{(y_0,\infty)}(x)\ \le\ \varepsilon \tag{4.8}$$

for all $n\in\mathbb{N}$ and all $x\in E_n(s)\cap K$. With this choice of y_0 we are now able to verify (4.7) as follows. Since $|f(y)|<\varepsilon$ for all $y>x_0$ and, hence, for all $y>y_0$, we obtain for all $n\in\mathbb{N}$ and all $x\in E_n(s)$

$$|T_{s,t}^{(n)}f(x)-T_tf(x)|\ \le\ 2\varepsilon+|T_{s,t}^{(n)}g(x)-T_tg(x)|,$$

where $g:=f1_{[0,y_0]}$. By the Weierstrass approximation theorem we can approximate the continuous function f uniformly on the compact interval $[0,y_0]$ by a polynomial p. Hence, there exists a polynomial p such that $\|g-h\|<\varepsilon$, where $h:=p1_{[0,y_0]}$. Thus, for all $n\in\mathbb{N}$ and all $x\in E_n(s)$

$$\begin{aligned}|T_{s,t}^{(n)}f(x)-T_tf(x)| &\le 4\varepsilon+|T_{s,t}^{(n)}h(x)-T_th(x)|\\ &\le 4\varepsilon+|T_{s,t}^{(n)}p(x)-T_tp(x)|+|T_{s,t}^{(n)}r(x)|+|T_tr(x)|,\end{aligned}$$

where $r := p - h = p - p1_{[0,y_0]} = p1_{(y_0,\infty)}$. It is already shown in Step 3 that

$$\lim_{n\to\infty}\sup_{x\in E_n(s)\cap K}|T_{s,t}^{(n)}p(x)-T_tp(x)| \ = \ 0.$$

Thus it remains to treat $|T_{s,t}^{(n)}r(x)|$ and $|T_tr(x)|$. Applying the Hölder inequality and using (4.8) we obtain

$$|T_{s,t}^{(n)}r(x)|^2 \ \le \ T_{s,t}^{(n)}p^2(x)\,T_{s,t}^{(n)}1_{(y_0,\infty)}(x) \ \le \ \varepsilon T_{s,t}^{(n)}p^2(x)$$

for all $n\in\mathbb{N}$ and all $x\in E_n(s)\cap K$. Thus it remains to show that $T_{s,t}^{(n)}p^2(x)$ is bounded uniformly for all $x\in E_n(s)\cap K$. We have

$$\sup_{x\in E_n(s)\cap K}|T_{s,t}^{(n)}p^2(x)| \ \le \ \sup_{x\in E_n(s)\cap K}|T_{s,t}^{(n)}p^2(x)-T_tp^2(x)|+\sup_{x\in K}|T_tp^2(x)|.$$

Since p^2 is a polynomial, the first expression converges to zero as $n\to\infty$ by Step 3. The last supremum is obviously bounded, since T_tp^2 is continuous and hence bounded on the compact set K, i.e. $M := \sup_{x\in K}|T_tp^2(x)|<\infty$. Similarly, by the Hölder inequality and (4.8), $|T_tr(x)|^2\le T_tp^2(x)T_t1_{(y_0,\infty)}(x)\le\varepsilon T_tp^2(x)\le\varepsilon M$ for all $x\in E_n(s)\cap K$. In summary, (4.7) is established. The proof is complete. □

5. Appendix

In this appendix we collect essentially two results. The first result (Lemma 5.1) concerns the Laplace exponent of the subordinator S introduced at the beginning of Section 2. The second result (Lemma 5.3) concerns some fundamental properties of the semigroup $(T_t)_{t\ge0}$ defined via (2.6).

Lemma 5.1. *Fix $\alpha\in(0,1)$. The drift-free subordinator $S=(S_t)_{t\ge0}$ with killing rate $k:=1/\Gamma(1-\alpha)$ and Lévy measure ϱ with density (2.1) has Laplace exponent (2.2).*

Proof: By the Lévy-Khintchine representation, the subordinator S has Laplace exponent $\Phi(x)=k+\int_{(0,\infty)}(1-e^{-xu})\,\varrho(\mathrm{d}u)$, $x\in[0,\infty)$. Since ϱ has density (2.1) it follows that

$$\Phi(x) \ = \ k+\frac{1}{\Gamma(1-\alpha)}\int_0^\infty(1-e^{-xu})\frac{e^{-u/\alpha}}{(1-e^{-u/\alpha})^{\alpha+1}}\,\mathrm{d}u.$$

The substitution $y=1-e^{-u/\alpha}$ ($\Rightarrow u=-\alpha\log(1-y)$ and $\mathrm{d}u/\mathrm{d}y=\alpha/(1-y)$) leads to

$$\Phi(x) \ = \ k+\frac{1}{\Gamma(1-\alpha)}\int_0^1(1-(1-y)^{\alpha x})\frac{\alpha}{y^{\alpha+1}}\,\mathrm{d}y. \tag{5.1}$$

Partial integration with $u(y) := 1-(1-y)^{\alpha x}$ and $v(y) := -y^{-\alpha}$ turns the last integral into

$$\begin{aligned}\int_0^1 (1-(1-y)^{\alpha x})\frac{\alpha}{y^{\alpha+1}}\,\mathrm{d}y & \\ &= \big[(1-(1-y)^{\alpha x})(-y^{-\alpha})\big]_0^1 - \int_0^1 \alpha x(1-y)^{\alpha x-1}(-y^{-\alpha})\,\mathrm{d}y\\ &= -1+\alpha x\int_0^1 y^{-\alpha}(1-y)^{\alpha x-1}\,\mathrm{d}y\\ &= -1+\alpha x B(1-\alpha,\alpha x) \;=\; -1+\frac{\Gamma(1-\alpha)\Gamma(1+\alpha x)}{\Gamma(1-\alpha+\alpha x)}.\end{aligned}$$

Plugging this into (5.1) and noting that $k = 1/\Gamma(1-\alpha)$ yields $\Phi(x) = \Gamma(1+\alpha x)/\Gamma(1-\alpha+\alpha x)$, which is (2.2). □

Let $E := [0,\infty)$ and let $\widehat{C}(E)$ denote the set of continuous functions $f : E \to \mathbb{R}$ vanishing at infinity. The following result is well known, we nevertheless mention it since it will turn out to be useful to verify fundamental properties of the semigroup $(T_t)_{t\geq 0}$ defined via (2.6).

Lemma 5.2. *Every $f \in \widehat{C}(E)$ is uniformly continuous on E.*

Proof: Let $\varepsilon > 0$. Since f vanishes at infinity, there exists $x_0 \in [0,\infty)$ such that $|f(x)| < \varepsilon/2$ for all $x \in [x_0,\infty)$. By the theorem of Heine, f is uniformly continuous on $[0, x_0+1]$. Thus, there exists $\delta = \delta(\varepsilon) \in (0,1)$ such that $|f(x)-f(y)| < \varepsilon$ for all $x, y \in [0, x_0+1]$ with $|x-y| < \delta$. If $|x-y| < \delta$ but $x > x_0+1$ or $y > x_0+1$, then $x \geq x_0$ and $y \geq x_0$ and hence $|f(x)-f(y)| \leq |f(x)|+|f(y)| < \varepsilon/2+\varepsilon/2 = \varepsilon$. Thus, $|f(x)-f(y)| < \varepsilon$ for all $x, y \in E$ with $|x-y| < \delta$. □

Lemma 5.3. *For every $t \in [0,\infty)$ the operator T_t defined via (2.6) satisfies $T_t\widehat{C}(E) \subseteq \widehat{C}(E)$. Moreover, for every $f \in \widehat{C}(E)$, $\lim_{t\to 0} T_tf(x) = f(x)$ uniformly for all $x \in E$, so $(T_t)_{t\geq 0}$ is a strongly continuous semigroup on $\widehat{C}(E)$.*

Proof: For $t \in [0,\infty)$, $f \in \widehat{C}(E)$ and $x \in E$ we have

$$T_tf(x) \;=\; \int_E f(x^{e^{-t}}y)\mathbb{P}_{\eta_t}(\mathrm{d}y) \;\to\; 0, \qquad x \to \infty,$$

by dominated convergence and, similarly,

$$T_tf(x) - T_tf(x_0) \;=\; \int_E \big(f(x^{e^{-t}}y) - f(x_0^{e^{-t}}y)\big)\,\mathbb{P}_{\eta_t}(\mathrm{d}y) \;\to\; 0, \qquad x \to x_0,$$

again by dominated convergence. Thus $T_t\widehat{C}(E) \subseteq \widehat{C}(E)$ for all $t \in [0,\infty)$.

In order to prove the second statement fix $f \in \widehat{C}(E)$. Note that f is bounded, i.e. $\|f\| := \sup_{x\in E}|f(x)| < \infty$. For $\alpha \in [0,1]$ let Z_α denote a random variable being Mittag–Leffler distributed with parameter α. Since $T_tf(x) = \mathbb{E}(f(x^{e^{-t}}\eta_t))$, where η_t is Mittag–Leffler distributed with parameter $\alpha := e^{-t}$, we have to verify that $\lim_{\alpha\to 1}\mathbb{E}(f(x^\alpha Z_\alpha)) = f(x)$ uniformly for all $x \in E$, where without loss of generality we can assume that $\alpha \in [1/2, 1]$.

Fix $\varepsilon > 0$. Since f vanishes at infinity, there exists a constant $x_0 \in [1,\infty)$ such that $|f(x)| < \varepsilon$ for all $x \geq x_0$. Define $K := 4x_0^2$ $(\geq x_0 \geq 1)$. In the following the uniform convergence $\lim_{\alpha\to 1}\mathbb{E}(f(x^\alpha Z_\alpha)) = f(x)$ is verified by distinguishing the two situations $x \in [K, \infty)$ and $x \in [0, K]$. For $x \in [K, \infty)$ we essentially exploit the fact that f vanishes at infinity. For $x \in [0, K]$ the uniform continuity of f (Lemma 5.2) comes into play. Let us start with the case $x \in [K, \infty)$.

For all $x \geq K$ and all $z \geq 1/2$ we have $x^\alpha z \geq x^\alpha/2 \geq \sqrt{x}/2 \geq \sqrt{K}/2 = x_0$ and, hence, $|f(x^\alpha z)| < \varepsilon$. We therefore obtain uniformly for all $x \geq K$

$$\begin{aligned} |\mathbb{E}(f(x^\alpha Z_\alpha)) - f(x)| &\leq \int_E |f(x^\alpha z) - f(x)| \, \mathbb{P}_{Z_\alpha}(\mathrm{d}z) \\ &= \int_{[1/2,\infty)} \underbrace{|f(x^\alpha z) - f(x)|}_{\leq 2\varepsilon} \mathbb{P}_{Z_\alpha}(\mathrm{d}z) + \int_{[0,1/2)} \underbrace{|f(x^\alpha z) - f(x)|}_{\leq 2\|f\|} \mathbb{P}_{Z_\alpha}(\mathrm{d}z) \\ &\leq 2\varepsilon + 2\|f\| \, \mathbb{P}(Z_\alpha < 1/2) \ \to \ 2\varepsilon \end{aligned}$$

as $\alpha \to 1$, since $Z_\alpha \to Z_1 \equiv 1$ in distribution as $\alpha \to 1$.

Assume now that $x \in [0, K]$. By Lemma 5.2 the function f is uniformly continuous on E. Thus, there exists a constant $\delta = \delta(\varepsilon) > 0$ such that $|f(y) - f(x)| < \varepsilon$ for all $x, y \in E$ with $|y - x| < \delta$. Since x^α converges to x as $\alpha \to 1$ uniformly on $[0, K]$ we can choose $\alpha_0 = \alpha_0(\delta) = \alpha_0(\varepsilon) < 1$ sufficiently close to 1 such that $|x^\alpha - x| < \delta/2$ for all $\alpha \in [\alpha_0, 1]$ and all $x \in [0, K]$. For all $\alpha \in [\alpha_0, 1]$, $x \in [0, K]$ and all $z \in E$ with $|z - 1| < \gamma := \delta/(2K)$ we have

$$\begin{aligned} |x^\alpha z - x| &\leq |x^\alpha z - x^\alpha| + |x^\alpha - x| \ = \ x^\alpha |z - 1| + |x^\alpha - x| \\ &< K^\alpha |z - 1| + \frac{\delta}{2} \ \leq \ K|z - 1| + \frac{\delta}{2} \ < \ K\gamma + \frac{\delta}{2} \ = \ \delta, \end{aligned}$$

and, hence, $|f(x^\alpha z) - f(x)| < \varepsilon$. For all $\alpha \in [\alpha_0, 1]$ and all $x \in [0, K]$ it follows that

$$\begin{aligned} |\mathbb{E}(f(x^\alpha Z_\alpha)) - f(x)| &\leq \int_E |f(x^\alpha z) - f(x)| \, \mathbb{P}_{Z_\alpha}(\mathrm{d}z) \\ &= \int_{\{|z-1|<\gamma\}} \underbrace{|f(x^\alpha z) - f(x)|}_{\leq \varepsilon} \mathbb{P}_{Z_\alpha}(\mathrm{d}z) + \int_{\{|z-1|\geq\gamma\}} \underbrace{|f(x^\alpha z) - f(x)|}_{\leq 2\|f\|} \mathbb{P}_{Z_\alpha}(\mathrm{d}z) \\ &\leq \varepsilon + 2\|f\| \, \mathbb{P}(|Z_\alpha - 1| \geq \gamma) \ \to \ \varepsilon \end{aligned}$$

as $\alpha \to 1$, since $Z_\alpha \to Z_1 \equiv 1$ in probability as $\alpha \to 1$. In summary it is shown that $\lim_{\alpha\to 1} \mathbb{E}(f(x^\alpha Z_\alpha)) = f(x)$ uniformly for all $x \in E$. Thus, $\lim_{t\to 0} T_t f(x) = f(x)$ uniformly for all $x \in E$. □

References

E. Baur and J. Bertoin. The fragmentation process of an infinite recursive tree and Ornstein–Uhlenbeck type processes. *Preprint* (2015). arXiv:1501.01400.

J. Bertoin and J.-F. Le Gall. The Bolthausen–Sznitman coalescent and the genealogy of continuous-state branching processes. *Probab. Theory Related Fields* **117**, 249–266 (2000). MR1771663.

P. Billingsley. *Probability and Measure. 3rd edition.* Wiley, New York (1995). MR1324786.

E. Bolthausen and A.-S. Sznitman. On Ruelle's probability cascades and an abstract cavity method. *Commun. Math. Phys.* **197**, 247–276 (1998). MR1652734.

P. Carmona, F. Petit and M. Yor. On the distribution and asymptotic results for exponential functionals of Lévy processes. *Rev. Math. Iberoamericana* **14**, 311–367 (1997). MR1648657.

S. N. Ethier and T. G. Kurtz. *Markov Processes. Characterization and Convergence.* Wiley, New York (1986). MR0838085.

W. Feller. *An Introduction to Probability Theory and Its Applications. Vol. II.* Wiley, New York (1971). MR0270403.

A. Gnedin, B. Hansen and J. Pitman. Notes on the occupancy problem with infinitely many boxes: general asymptotics and power laws. *Probab. Surv.* **4**, 146–171 (2007). MR2318403.

B. Haas and G. Miermont. Self-similar scaling limits of non-increasing Markov chains. *Bernoulli* **17**, 1217–1247 (2011). MR2854770.

L. F. James. Lamperti-type laws. *Ann. Appl. Probab.* **20**, 1303–1340 (2010). MR2676940.

K. Jayakumar. Mittag–Leffler process. *Math. Comput. Modelling* **37**, 1427–1434 (2003). MR1996047.

K. Jayakumar and R. N. Pillai. The first-order autoregressive Mittag–Leffler process. *J. Appl. Probab.* **30**, 462–466 (1993). MR1212678.

V. Limic. Genealogies of regular exchangeable coalescents with applications to sampling. *Ann. Inst. Henri Poincaré Probab. Stat.* **48**, 706–720 (2012). MR2976560.

D. B. Madan and M. Yor. Making Markov martingales meet marginals: with explicit constructions. *Bernoulli* **8**, 509–536 (2002). MR1914701.

F. Mainardi and R. Gorenflo. On Mittag–Leffler-type functions in fractional evolution processes. *J. Comput. Appl. Math.* **118**, 283–299 (2000). MR1765955.

M. Möhle. Duality and cones of Markov processes and their semigroups. *Markov Process. Related Fields* **19**, 149–162 (2013). MR3088428.

M. Möhle and H. Pitters. A spectral decomposition for the block counting process of the Bolthausen–Sznitman coalescent. *Electron. Commun. Probab.* **19**, paper 47, 1–11 (2014). MR3246966.

J. Neveu. A continuous-state branching process in relation with the GREM model of spin glass theory. *Rapport interne no 267, Ecole Polytechnique* (1992).

N. Pillai. On Mittag–Leffler functions and related distributions. *Ann. Inst. Statist. Math.* **42**, 157–161 (1990). MR1054728.

J. Pitman. Coalescents with multiple collisions. *Ann. Probab.* **27**, 1870–1902 (1999). MR1742892.

J. Pitman. *Combinatorial Stochastic Processes. Lecture Notes in Mathematics* **1875**. Springer, Berlin (2006). MR2245368.

S. Sagitov. The general coalescent with asynchronous mergers of ancestral lines. *J. Appl. Probab.* **36**, 1116–1125 (1999). MR1742154.

D. Siegmund. The equivalence of absorbing and reflecting barrier problems for stochastically monotone Markov processes. *Ann. Probab.* **4**, 914–924 (1976). MR0431386.