ONLINE SUPPLEMENT TO "DELAYED HAWKES BIRTH-DEATH PROCESSES"

JUSTIN BAARS, ROGER J. A. LAEVEN, AND MICHEL MANDJES

Abstract. In this online supplement to our paper "Delayed Hawkes birth-death processes", we provide the proofs of Section 4 and prove Theorem 4. For context, notation and definitions, see the main paper.

Email addresses. j.r.baars@uva.nl, r.j.a.laeven@uva.nl, and m.r.h.mandjes@math.leidenuniv.nl.

Date: October 21, 2026.

The purpose of this online supplement to our paper *Delayed Hawkes birth-death processes* [1] is twofold. The first part provides self-contained proofs of the scaling limits described in [1], Section 4. More specifically, we provide a proof of [1], Theorem 3, along with auxiliary results. Since those results partially mimic results from [2], we left the details out of the main text. In the second part, we prove [1], Theorem 4, which describes fixed points in the transform domain for Hawkes processes with sojourn-time dependent excitation. The proof is a modification of the work done in [4], and is therefore left out of the main text.

A. Proofs of Section 4: Scaling limits with stretched sojourn time

In [1], Section 4, the question arose whether we could show the difference between a "classical" and a delayed Hawkes process through some scaling limit. We consider a univariate unmarked delayed Hawkes process with i.i.d. sojourn times $(J_i)_i \in \mathbb{N}$, which we compare to its nondelayed counterpart. The idea is to consider the compensated process on an interval $[0,T]$ with sojourn times stretched out from J_i to $T^\alpha J_i$, for some $\alpha \in [0,1)$, after which we contract time by a factor of T, mapping $[0,T]$ onto $[0,1]$. After rescaling our counting process by $T^{-1/2}$ and letting $T \to \infty$, we obtain a nondegenerate limit. By considering sojourn-time stretching of degree α with $0 \leqslant \alpha < \frac{1}{2}$, we obtain the same scaling limit as given found in [2], Theorem 2, whereas if we set $\alpha = \frac{1}{2}$, the effect of the delays becomes visible.

To gain insight into the scaling limits, we modify the arguments from [2]. Let $(N_\alpha^T(v))_{v\in[0,1]}$ be equal to $(\tilde{N}_\alpha^T(Tv))_{v\in[0,1]}$, with $\tilde{N}_\alpha^T(\cdot)$ the process having sojourn times $(T^\alpha J_i)_{i\in\mathbb{N}}$. Those sojourn times correspond to the increasing sequence of arrival times $(\tau_i)_{i\in\mathbb{N}}$, where it is assumed that J_i is drawn at time τ_i. This process $N_\alpha^T(\cdot)$ has an arrival intensity $\Lambda_\alpha^T(\cdot)$ given by

$$\Lambda_\alpha^T(v) = \left(\lambda_0 + \sum_{\tau_i<t} h(Tv - \tau_i - T^\alpha J_i)\mathbf{1}\{Tv \geqslant \tau_i + T^\alpha J_i\}\right)\cdot T. \tag{A.I}$$

To derive our scaling limit, we impose the following three assumptions:

$$\|h\|_{L^1} < 1, \tag{A1}$$

$$\int_0^\infty t^{\frac{1}{2(1-\alpha)}} h(t)\,\mathrm{d}t < \infty, \tag{A2}$$

$$\mathbb{E}[J] < \infty. \tag{A3}$$

We assume (A1) throughout this section, whereas we use (A2)–(A3) only for $\alpha \leqslant \frac{1}{2}$.

In this section, we use the following notation for the (presumed) mean, standard deviation, and "shift" of the delayed Hawkes process with stretched sojourn times:

$$\mu := \frac{\lambda_0}{1 - \|h\|_{L^1}}, \quad \sigma^2 := \frac{\lambda_0}{(1 - \|h\|_{L^1})^3}, \quad \bar{\sigma} := \frac{\lambda_0 \mathbb{E}[J]\|h\|_{L^1}}{(1 - \|h\|_{L^1})^2}. \tag{A.II}$$

In the following, we use the function $\bar{h}^T(\cdot)$, which can be seen as an average of h over the past, weighed according to the stretched sojourn times:

$$\bar{h}^T(t) := \int_0^{T^{-\alpha}t} h(t - T^\alpha w)\, \mathrm{d}\bar{\mathscr{F}}(w). \tag{A.III}$$

We also define

$$\mathscr{H}^T := \sum_{k \geqslant 1} (\bar{h}^T)^{*k}, \tag{A.IV}$$

where f^{*k} denotes the k-fold convolution of a function f. In this document, we suppress the α-dependence in the notations $\bar{h}^T$ and $\mathscr{H}^T$ to make our notation more compact; the value of α will be clear from the context. Note that for any $T > 0$, $\alpha \in [0, 1)$,

$$\begin{aligned} \|\bar{h}^T\|_{L^1} &= \int_0^\infty \int_0^{T^{-\alpha}t} h(t - T^\alpha w)\, \mathrm{d}\bar{\mathscr{F}}(w)\, \mathrm{d}t = \int_0^\infty \int_{T^\alpha w}^\infty h(t - T^\alpha w)\, \mathrm{d}t\, \mathrm{d}\bar{\mathscr{F}}(w) \\ &= \int_0^\infty \int_0^\infty h(t)\, \mathrm{d}t\, \mathrm{d}\bar{\mathscr{F}}(w) = \|h\|_{L^1}, \end{aligned} \tag{A.V}$$

and therefore, using that $\|(\bar{h}^T)^{*k}\|_{L^1} = \|\bar{h}^T\|_{L^1}^k$, as can easily be proved by induction,

$$\|\mathscr{H}^T\|_{L^1} = \sum_{k \geqslant 1} \|(\bar{h}^T)^{*k}\|_{L^1} = \sum_{k \geqslant 1} \|h\|_{L^1}^k = \frac{\|h\|_{L^1}}{1 - \|h\|_{L^1}} < \infty. \tag{A.VI}$$

For the next lemmas, consider $\tilde{N}_\alpha^T$, a delayed Hawkes process with sojourn times stretched out by a factor T^α. We emphasize that we do *not* contract time, yet. The next three lemmas can be seen as suitable counterparts of [2], Lemmas 2, 4 and 5, respectively.

Lemma A.I. *Let $\alpha \in [0, 1)$, $T, t \geqslant 0$. For each a.s. finite stopping time S, we have*

$$\mathbb{E}[\tilde{N}_\alpha^T(S)] = \lambda_0 \mathbb{E}[S] + \mathbb{E}\left[\int_0^S \bar{h}^T(S - t)\tilde{N}_\alpha^T(t)\, \mathrm{d}t\right], \tag{A.VII}$$

$$\mathbb{E}[\tilde{N}_\alpha^T(S)] \leqslant \mu\, \mathbb{E}[S], \tag{A.VIII}$$

where μ is defined in (A.II).

Proof. The proof is a modification of the proof of [2], Lemma 2.

Let (τ_p) be the sequence jump times of $\tilde{N}_\alpha^T$, and set $S_p = S \wedge \tau_p$. Then it holds that

$$\begin{aligned} \mathbb{E}[\tilde{N}_\alpha^T(S_p)] &= \lambda_0 \mathbb{E}[S_p] + \mathbb{E}\left[\int_0^{S_p} \int_0^{T^{-\alpha}t} \int_0^{t - T^\alpha w} h(t - s - T^\alpha w)\, \mathrm{d}\tilde{N}_\alpha^T(s)\, \mathrm{d}\bar{\mathscr{F}}(w)\, \mathrm{d}t\right] \\ &= \lambda_0 \mathbb{E}[S_p] + \mathbb{E}\left[\int_0^{S_p} \int_0^t \int_0^{T^{-\alpha}(t-s)} h(t - s - T^\alpha w)\, \mathrm{d}\bar{\mathscr{F}}(w)\, \mathrm{d}\tilde{N}_\alpha^T(s)\, \mathrm{d}t\right] \\ &= \lambda_0 \mathbb{E}[S_p] + \mathbb{E}\left[\int_0^{S_p} \int_0^t \bar{h}^T(t - s)\, \mathrm{d}\tilde{N}_\alpha^T(s)\, \mathrm{d}t\right] \end{aligned}$$

$$= \lambda_0 \mathbb{E}[S_p] + \mathbb{E}\left[\int_0^{S_p} \int_0^{S_p - s} \bar{h}^T(t)\, \mathrm{d}t\, \mathrm{d}\tilde{N}_\alpha^T(s)\right]. \tag{A.IX}$$

Let $\bar{H}^T(t) := \int_0^t \bar{h}^T(s)\, \mathrm{d}s$ be the integral of $\bar{h}^T$. Note that $\bar{H}^T(0) = \tilde{N}_\alpha^T(0) = 0$. It follows by integration by parts that

$$\begin{aligned} 0 &= \bar{H}^T(0) N_\alpha^T(S_p) - \bar{H}^T(S_p) N_\alpha^T(0) \\ &= \int_0^{S_p} \bar{H}^T(S_p - t)\, \mathrm{d}\tilde{N}_\alpha^T(t) - \int_0^{S_p} \bar{h}^T(S_p - t) N_\alpha^T(t)\, \mathrm{d}t. \end{aligned} \tag{A.X}$$

We combine (A.IX) and (A.X), using that $\tilde{N}_\alpha^T(S_p) \uparrow \tilde{N}_\alpha^T(S)$ as $p \to \infty$, to obtain (A.VII).

Finally, the inequality (A.VIII) follows using the inductive argument from the proof of [2], Lemma 2, with the only change that the inductive step now starts from (A.VII). The stability condition is given by (A1) — use eqn. (A.V). □

Now consider the martingale $\tilde{M}_\alpha^T(t) = \tilde{N}_\alpha^T(t) - \int_0^t \tilde{\Lambda}_\alpha^T(s)\, \mathrm{d}s$, where $\tilde{\Lambda}_\alpha^T$ denotes the arrival intensity of $\tilde{N}_\alpha^T$.

Lemma A.II. *Let $\alpha \in [0, 1)$, $T, t \geqslant 0$. Then it holds that*

$$\mathbb{E}[\tilde{N}_\alpha^T(t)] = \lambda_0 t + \lambda_0 \int_0^t \mathscr{H}^T(t - s) s\, \mathrm{d}s, \tag{A.XI}$$

$$\tilde{N}_\alpha^T(t) - \mathbb{E}[\tilde{N}_\alpha^T(t)] = \tilde{M}_\alpha^T(t) + \int_0^t \mathscr{H}^T(t - s) \tilde{M}_\alpha^T(s)\, \mathrm{d}s. \tag{A.XII}$$

Proof. The result is derived from Lemma A.I in the same way as [2], Lemma 4 is derived from [2], Lemma 2; one simply replaces their φ by our $\bar{h}^T$ and their ψ by our $\mathscr{H}^T$. □

Lemma A.III. *Let $\alpha \in [0, 1)$, let $p \in [0, 1]$, and assume that $\int_0^\infty t^p h(t)\, \mathrm{d}t < \infty$. Let $\epsilon \in (0, 1)$. The following holds:*

- *If $p < 1$, then $T^{(1-\alpha)p} \left(T^{-1}\mathbb{E}[N_\alpha^T(v)] - \mu v\right) \to 0$, as $T \to \infty$, uniformly in $v \in [0, 1]$;*
- *If $p = 1$, then $T^{1-\alpha} \left(T^{-1}\mathbb{E}[N_\alpha^T(v)] - \mu v\right) \to -\bar{\sigma}$, as $T \to \infty$, uniformly in $v \in [\epsilon, 1]$,*

where μ and $\bar{\sigma}$ are defined in (A.II).

Proof. First, we calculate

$$\begin{aligned} \int_0^\infty t^p \bar{h}^T(t)\, \mathrm{d}t &= \int_0^\infty \int_0^{tT^{-\alpha}} t^p h(t - T^\alpha w)\, \mathrm{d}\bar{\mathscr{J}}(w)\, \mathrm{d}t = \int_0^\infty \int_{wT^\alpha}^\infty t^p h(t - T^\alpha w)\, \mathrm{d}t\, \mathrm{d}\bar{\mathscr{J}}(w) \\ &= \int_0^\infty \int_0^\infty (t + T^\alpha w)^p h(t)\, \mathrm{d}t\, \mathrm{d}\bar{\mathscr{J}}(w) \leqslant \int_0^\infty \int_0^\infty (t^p + T^{\alpha p} w^p) h(t)\, \mathrm{d}t\, \mathrm{d}\bar{\mathscr{J}}(w) \\ &\leqslant \int_0^\infty t^p h(t)\, \mathrm{d}t + T^{\alpha p}(1 + \mathbb{E}[J]) \|h\|_{L^1}, \end{aligned} \tag{A.XIII}$$

where for $p = 1$ we find the equality

$$\int_0^\infty t \bar{h}^T(t)\, \mathrm{d}t = \int_0^\infty t h(t)\, \mathrm{d}t + T^\alpha \mathbb{E}[J] \|h\|_{L^1}, \tag{A.XIV}$$

and where for $p = 0$, Assumptions (A2)–(A3) are not required.

Next, as in the proof of [2], Lemma 5, we find that

$$\int_0^\infty t^p \mathscr{H}^T(t)\,\mathrm{d}t \leqslant \frac{\int_0^\infty t^p \bar{h}^T(t)\,\mathrm{d}t}{(1-\|h\|_{L^1})^2}, \tag{A.XV}$$

again with equality if $p = 1$.

Now, consider a fixed $T > 0$, and scale the process $\tilde{N}_\alpha^T$ from Lemma A.II to N_α^T by contracting time by a factor T. Using (A.XI), it follows that

$$\frac{\lambda_0}{1-\|h\|_{L^1}} v - T^{-1}\mathbb{E}[N_\alpha^T(v)] = \lambda_0 v \int_{Tv}^\infty \mathscr{H}^T(s)\,\mathrm{d}s + \lambda_0 T^{-1}\int_0^{Tv} s\mathscr{H}^T(s)\,\mathrm{d}s. \tag{A.XVI}$$

We bound $T^{(1-\alpha)p}$ times the first term from (A.XVI) by

$$\begin{aligned}\lambda_0 v T^{(1-\alpha)p}\int_{Tv}^\infty \mathscr{H}^T(s)\,\mathrm{d}s &= \lambda_0 v^{1-p}T^{-\alpha p}\int_{Tv}^\infty (Tv)^p\mathscr{H}^T(s)\,\mathrm{d}s \\ &\leqslant \lambda_0 v^{1-p}T^{-\alpha p}\int_{Tv}^\infty s^p\mathscr{H}^T(s)\,\mathrm{d}s,\end{aligned} \tag{A.XVII}$$

which converges to 0 as $T \to \infty$, by invoking the bounds found in (A.XIII) and (A.XV). The convergence is uniform in $v \in [0,1]$ in case $p < 1$, whereas the convergence is uniform on $[\epsilon, 1]$ (for any $\epsilon \in (0,1)$) in case $p = 1$.

Next, we consider the second term from (A.XVI). Suppose first that $p < 1$. Since $T^{(1-\alpha)p}$ times the second term can be bounded by

$$\lambda_0 T^{-(1-(1-\alpha)p)}\int_0^T s\mathscr{H}^T(s)\,\mathrm{d}s, \tag{A.XVIII}$$

to prove convergence of the second term from (A.XVI) to 0, uniformly in $v \in [0,1]$, it suffices to prove that (A.XVIII) converges to 0. This can be proved in the same way as in [2], Lemma 5, applying integration by parts to $G(t) = \int_0^t s^{(1-\alpha)p}\mathscr{H}^T(s)\,\mathrm{d}s$.

Suppose now that $p = 1$. Using (A.XIV) and (A.XV), it follows that

$$T^{1-\alpha}T^{-1}\int_0^{Tv} s\mathscr{H}^T(s)\,\mathrm{d}s \to \frac{\mathbb{E}[J]\|h\|_{L^1}}{(1-\|h\|_{L^1})^2}, \tag{A.XIX}$$

as $T \to \infty$, uniformly in $v \in [\epsilon, 1]$. For the last limit, we really need $\alpha < 1$; otherwise the stretching factors are of the same order as the limit of integration Tv. The result follows. □

Lemma A.IV. *There exists a constant $C_{\lambda_0,h}$ such that for all $t, \Delta \geqslant 0$:*

$$\mathbb{E}\left[\sup_{t\leqslant s\leqslant t+\Delta}\left(\tilde{M}_\alpha^T(s) - \tilde{M}_\alpha^T(t)\right)^2\right] \leqslant C_{\lambda_0,h}\Delta. \tag{A.XX}$$

Proof. The proof closely follows the proof of [2], Lemma 6. Concretely, by Doob's maximal inequality,

$$\mathbb{E}\left[\sup_{t\leqslant s\leqslant t+\Delta}\left(\tilde{M}_\alpha^T(s) - \tilde{M}_\alpha^T(t)\right)^2\right] \leqslant 4\mathbb{E}\left[\left(\tilde{M}_\alpha^T(t+\Delta) - \tilde{M}_\alpha^T(t)\right)^2\right]. \tag{A.XXI}$$

Since the quadratic variation of $\tilde{M}_\alpha^T$ satisfies $[\tilde{M}_\alpha^T, \tilde{M}_\alpha^T]_t = \tilde{N}_\alpha^T(t)$, it follows that

$$\mathbb{E}\left[\left(\tilde{M}_\alpha^T(t+\Delta) - \tilde{M}_\alpha^T(t)\right)^2\right] = \mathbb{E}\left[\tilde{N}_\alpha^T(t+\Delta) - \tilde{N}_\alpha^T(t)\right]. \tag{A.XXII}$$

Applying (A.XI), we obtain

$$\mathbb{E}\left[\tilde{N}_\alpha^T(t+\Delta) - \tilde{N}_\alpha^T(t)\right] = \lambda_0\Delta + \lambda_0\left(\int_0^{t+\Delta} \mathscr{H}^T(t+\Delta-s)s\,\mathrm{d}s - \int_0^t \mathscr{H}^T(t-s)s\,\mathrm{d}s\right)$$
$$\leqslant \lambda_0\Delta\left(1 + \int_0^{t+\Delta} \mathscr{H}^T(s)\,\mathrm{d}s\right) \leqslant \lambda_0\Delta\left(1 + \int_0^{\infty} \mathscr{H}^T(s)\,\mathrm{d}s\right), \quad \text{(A.XXIII)}$$

where the first inequality follows by a change of variables $s' = s - \Delta$ in the first integral, and by rearranging. By applying (A.VI), the result follows from (A.XXI)–(A.XXIII). □

We are now equipped to establish an FLLN for $\left(N_\alpha^T(\cdot)\right)_{T\geqslant 0}$. After the FLLN, we present our FCLT for α-stretched sojourn times.

We previously defined the martingale $\tilde{M}_\alpha^T(t) = \tilde{N}_\alpha^T(t) - \int_0^t \tilde{\Lambda}_\alpha^T(s)\,\mathrm{d}s$. Now also consider the martingale $M_\alpha^T(t) = N_\alpha^T(t) - \int_0^t \Lambda_\alpha^T(s)\,\mathrm{d}s$ obtained from the time-scaled process N_α^T.

Theorem A.I (FLLN)**.** *Let $\alpha \in [0,1)$. It holds that $\tilde{N}_\alpha^T(t) \in L^2(\mathbb{P})$, for all $T, t \geqslant 0$, and we have*

$$\sup_{v\in[0,1]} \left|T^{-1}N_\alpha^T(v) - \mu\right| \to 0 \text{ in } L^2(\mathbb{P}), \text{ as } T \to \infty, \quad \text{(A.XXIV)}$$

where μ is as defined in (A.II)*.*

Proof. By Lemma A.III with $p = 0$, it suffices to prove that

$$T^{-1}\sup_{v\in[0,1]} \left|N_\alpha^T(v) - \mathbb{E}[N_\alpha^T(v)]\right| \to 0 \text{ in } L^2(\mathbb{P}), \text{ as } T \to \infty.$$

Noting that $N_\alpha^T(v) = \tilde{N}_\alpha^T(Tv)$, we use (A.XII) to obtain

$$\sup_{v\in[0,1]} \left|N_\alpha^T(v) - \mathbb{E}[N_\alpha^T(v)]\right| = \sup_{v\in[0,1]} \left|M_\alpha^T(v) + \int_0^{Tv} \mathscr{H}^T(Tv-s)M_\alpha^T(s/T)\,\mathrm{d}s\right|$$
$$= \sup_{v\in[0,1]} \left|M_\alpha^T(v) + \int_0^{Tv} \mathscr{H}^T(s)M_\alpha^T\left(\frac{Tv-s}{T}\right)\,\mathrm{d}s\right|$$
$$\leqslant \left(1 + \|\mathscr{H}^T\|_{L^1}\right)\sup_{v\in[0,1]} |M_\alpha^T(v)| = \left(1 + \|\mathscr{H}^T\|_{L^1}\right)\sup_{v\in[0,1]} |\tilde{M}_\alpha^T(Tv)|.$$

By Lemma A.IV, we know that $\mathbb{E}\left[\sup_{v\in[0,1]} \tilde{M}_\alpha^T(Tv)^2\right] \leqslant C_{\lambda_0,h}T$, from which the stated $L^2(\mathbb{P})$-convergence follows. □

Lemma A.V. *Let W be a standard Brownian motion, and let μ be as defined in* (A.II)*. Then $M_\alpha^{(T)} := (T^{-1/2}M_\alpha^T(v))_{v\in[0,1]} \to \sqrt{\mu}W$, as $T \to \infty$, weakly on $D[0,1]$ equipped with the Skorokhod J_1-topology.*

Proof. Since $M^{(T)}$ has uniformly bounded jumps, and since the quadratic variation satisfies

$$\left[M_\alpha^{(T)}, M_\alpha^{(T)}\right]_v = \frac{1}{T}\left[M_\alpha^T, M_\alpha^T\right]_v = vN_\alpha^T \to \mu v \quad \text{(A.XXV)}$$

in $L^2(\mathbb{P})$, as $T \to \infty$, by Theorem A.I, the result is a direct consequence of [3], Theorem VIII-3.11. □

Theorem A.II (FCLT)**.** *Let μ, σ, and $\bar{\sigma}$ be as in* (A.II)*. Let $\epsilon \in (0,1)$, and let B be a standard Brownian motion. For $\alpha = \frac{1}{2}$, we have*

$$\left(\frac{N^T_{1/2}(v) - \mu\, Tv}{\sqrt{T}}\right)_{v\in[\epsilon,1]} \longrightarrow (-\bar{\sigma} + \sigma\, B(v))_{v\in[\epsilon,1]}\,, \tag{A.XXVI}$$

as $T \to \infty$, weakly on $D[\epsilon,1]$ equipped with the Skorokhod J_1-topology. On the other hand, for $\alpha \in [0, \frac{1}{2})$, it holds that

$$\left(\frac{N^T_\alpha(v) - \mu\, Tv}{\sqrt{T}}\right)_{v\in[0,1]} \longrightarrow (\sigma\, B(v))_{v\in[0,1]}\,, \tag{A.XXVII}$$

as $T \to \infty$, weakly on $D[0,1]$ equipped with the Skorokhod J_1-topology.

Proof. Set

$$X^T_\alpha(v) = T^{-1/2}\left(N^T_\alpha(v) - \mathbb{E}[N^T_\alpha(v)]\right). \tag{A.XXVIII}$$

We obtain a limit result for $X^T_\alpha(v)$, after which we set $p = \frac{1}{2(1-\alpha)}$, and apply Lemma A.III for this value of p to replace $T^{-1/2}\mathbb{E}[N^T_\alpha(v)]$ by $T^{-1/2}\mu v$. To apply Lemma A.III with $p = \frac{1}{2(1-\alpha)}$, we invoke Assumption (A2).

By this application of Lemma A.III and by invoking Lemma A.V, it suffices to prove that

$$\sup_{v\in[\eta,1]}\left|X^T_\alpha(v) - \frac{1}{1-\|h\|_{L^1}} M^{(T)}_\alpha(v)\right| \xrightarrow{\mathbb{P}} 0, \quad \text{as } T\to\infty, \tag{A.XXIX}$$

where $\eta = 0$ if $\alpha < \frac{1}{2}$, and where $\eta = \epsilon$ when $\alpha = \frac{1}{2}$. By (A.XII) it holds that

$$\begin{aligned} X^T_\alpha(v) &= M^{(T)}_\alpha(v) + \int_0^{Tv} \mathscr{H}^T(s) M^{(T)}_\alpha(v - s/T)\,\mathrm{d}s \\ &= M^{(T)}_\alpha(v) + T\int_0^{v} \mathscr{H}^T(Ts) M^{(T)}_\alpha(v - s)\,\mathrm{d}s. \end{aligned} \tag{A.XXX}$$

Using the calculation (A.VI), we can prove (A.XXIX) by showing that

$$\sup_{v\in[\eta,1]}\left|T\int_0^v \mathscr{H}^T(Ts) M^{(T)}_\alpha(v-s)\,\mathrm{d}s - \int_0^\infty \mathscr{H}^T(t)\,\mathrm{d}t M^{(T)}_\alpha(v)\right| \xrightarrow{\mathbb{P}} 0, \quad \text{as } T\to\infty, \tag{A.XXXI}$$

which is proved by repeating the proof of [2], Theorem 2, replacing their Lemma 7 by our Lemma A.V. □

B. Proof of Theorem 4

B.I. **Joint transform characterization.** We exploit the cluster representation provided in [1], Definition 2. First, we summarize the multidimensional notation as introduced in [4]. We need this notation to gain insight in the clustering structure, and to state and prove the results we are after.

Before we come to that, we note that this cluster representation is useful for several reasons. First, modulo the time shift corresponding to the arrival times, clusters generated by immigrants in the same coordinate are i.i.d. Next, cluster processes are generated independently across source components. Finally, each event from the same source component generates offspring by the same iterative procedure, since each child itself determines a cluster, i.e., there is *self-similarity*.

We operationalize these ideas as follows. Let $j \in [d]$. For each immigrant $(T^{(0)}_r, J^{(0)}_r, j)$, we denote the d-dimensional cluster process it generates as $\boldsymbol{S}^N_j(\cdot)$. We also consider the *birth-death*

cluster $\boldsymbol{S}_j^{\boldsymbol{Q}}(\cdot)$ and the *rate cluster* $\boldsymbol{S}_j^{\boldsymbol{\lambda}}(\cdot)$ generated by this immigrant, which give the number of remaining offspring (including the parent) and remaining intensity increases caused by the immigrant arrival. From now on, for $t \geqslant T_r^{(0)}$, we interpret $u = t - T_r^{(0)}$ as the time elapsed since the arrival of the corresponding immigrant. We are dealing with d-dimensional cluster processes, of which we write the components as

$$\boldsymbol{S}_j^{\boldsymbol{N}}(u) = \begin{bmatrix} S_{1\leftarrow j}^{\boldsymbol{N}}(u) \\ \vdots \\ S_{d\leftarrow j}^{\boldsymbol{N}}(u) \end{bmatrix}, \quad \boldsymbol{S}_j^{\boldsymbol{Q}}(u) = \begin{bmatrix} S_{1\leftarrow j}^{\boldsymbol{Q}}(u) \\ \vdots \\ S_{d\leftarrow j}^{\boldsymbol{Q}}(u) \end{bmatrix}, \quad \boldsymbol{S}_j^{\boldsymbol{\lambda}}(u) = \begin{bmatrix} S_{1\leftarrow j}^{\boldsymbol{\lambda}}(u) \\ \vdots \\ S_{d\leftarrow j}^{\boldsymbol{\lambda}}(u) \end{bmatrix}. \tag{B.I}$$

Here, $S_{i\leftarrow j}^{\boldsymbol{N}}(u)$ records the number of events in component i up to time u with as oldest ancestor the immigrant generating $\boldsymbol{S}_j^{\boldsymbol{N}}(\cdot)$, including the immigrant itself when $i = j$. Similarly, $S_{i\leftarrow j}^{\boldsymbol{Q}}(u)$ records the number of non-expired events in component i up to time u with as oldest ancestor the immigrant generating $\boldsymbol{S}_j^{\boldsymbol{Q}}(\cdot)$, including the ancestor itself if $i = j$ and if the ancestor has not yet left the system. Finally, $S_{i\leftarrow j}^{\boldsymbol{\lambda}}(u)$ records aggregated change in the intensity of component i caused by jumps with excitation functions $h_{im,J,\omega}$, following arrivals in component m with lifetime J within the cluster $\boldsymbol{S}_j^{\boldsymbol{\lambda}}(\cdot)$ generated by an immigrant in component j. For each $S_{i\leftarrow j}^{\star}(\cdot)$, $\star \in \{\boldsymbol{N}, \boldsymbol{Q}, \boldsymbol{\lambda}\}$, note that changes in i within the cluster generated by an immigrant in j might propagate through other dimensions $m \in [d]$ due to the multivariate setting.

An immigration event in some coordinate j generates first-generation offspring in *all* coordinates, which in turn constitute clusters themselves, called *subclusters*, which are second-generation offspring of the immigrant. To analyze the self-similarity inherent in this process, for $\star \in \{\boldsymbol{N}, \boldsymbol{Q}, \boldsymbol{\lambda}\}$, we define the matrix process

$$\boldsymbol{S}^{\star}(\cdot) := \left[\boldsymbol{S}_1^{\star}(\cdot) \mid \cdots \mid \boldsymbol{S}_d^{\star}(\cdot)\right] = \begin{bmatrix} S_{1\leftarrow 1}^{\star}(\cdot) & \cdots & S_{1\leftarrow d}^{\star}(\cdot) \\ \vdots & \ddots & \vdots \\ S_{d\leftarrow 1}^{\star}(\cdot) & \cdots & S_{d\leftarrow d}^{\star}(\cdot) \end{bmatrix} =: \begin{bmatrix} \boldsymbol{S}_{(1)}^{\star}(\cdot) \\ \vdots \\ \boldsymbol{S}_{(d)}^{\star}(\cdot) \end{bmatrix}. \tag{B.II}$$

Note that the jth column $\boldsymbol{S}_j^{\star}(\cdot)$ of this matrix process corresponds to offspring events originating in coordinate j, while the ith row $\boldsymbol{S}_{(i)}^{\star}(\cdot)$ describes offspring events arriving in component i.

Using the clustering processes defined in (B.I) and (B.II), we can state distributional equalities for the component processes N_i, Q_i, λ_i. Indeed, letting $I_j(\cdot)$ be the immigration process in coordinate $j \in [d]$, i.e., a homogeneous Poisson process of rate $\lambda_{j,0}$, we have

$$\begin{aligned} N_i(t) &\overset{\mathcal{D}}{=} \sum_{j=1}^{d} \sum_{k=1}^{I_j(t)} S_{i\leftarrow j}^{\boldsymbol{N}}(t - T_k); \\ Q_i(t) &\overset{\mathcal{D}}{=} \sum_{j=1}^{d} \sum_{k=1}^{I_j(t)} S_{i\leftarrow j}^{\boldsymbol{Q}}(t - T_k); \\ \lambda_i(t) &\overset{\mathcal{D}}{=} \lambda_{i,0} + \sum_{j=1}^{d} \sum_{k=1}^{I_j(t)} S_{i\leftarrow j}^{\boldsymbol{\lambda}}(t - T_k). \end{aligned} \tag{B.III}$$

Similar distributional equations can be formulated for the cluster processes, using the observation that each cluster itself generates subclusters. To exploit this structure, letting $\boldsymbol{X}(\cdot)$ be an $\mathbb{R}_+^d$-valued

time-dependent process and $P \geqslant 0$, for $j \in [d]$ we define the functional

$$\mathcal{A}_j(P, \boldsymbol{X}(\cdot))(u) = P + \sum_{m=1}^{d} \sum_{k=1}^{K_{mj,J,\omega}(u)} X_m(u - T_k), \tag{B.IV}$$

where T_k are arrival times in component m, and where $K_{mj,J,\omega}(\cdot)$ denotes an inhomogeneous Poisson process of rate $h_{mj,J,\omega}$. Here, it is understood that $J \sim J_j$, which is the same for each target coordinate m, and it is understood that the excitation functions $h_{mj,J,\omega}$ are conditionally independent. Finally, whenever P is an expression of J, it is understood that J is again the lifetime of the immigrant in coordinate j under consideration, i.e., the same J as appearing in $K_{mj,J,\omega}(\cdot)$. Using this functional, we have the following distributional equalities for the cluster processes:

$$\begin{aligned}
S^{\boldsymbol{N}}_{i\leftarrow j}(u) &\stackrel{\mathcal{D}}{=} \mathcal{A}_j\left(\mathbf{1}\{i=j\}, \boldsymbol{S}^{\boldsymbol{N}}_{(i)}(\cdot)\right)(u);\\
S^{\boldsymbol{Q}}_{i\leftarrow j}(u) &\stackrel{\mathcal{D}}{=} \mathcal{A}_j\left(\mathbf{1}\{i=j\}\mathbf{1}\{J>u\}, \boldsymbol{S}^{\boldsymbol{Q}}_{(i)}(\cdot)\right)(u);\\
S^{\boldsymbol{\lambda}}_{i\leftarrow j}(u) &\stackrel{\mathcal{D}}{=} \mathcal{A}_j\left(h_{ij,J,\omega}(\cdot), \boldsymbol{S}^{\boldsymbol{\lambda}}_{(i)}(\cdot)\right)(u).
\end{aligned} \tag{B.V}$$

As in the Markovian case, to characterize the probabilistic behavior of the joint process $(\boldsymbol{Q}(\cdot), \boldsymbol{\lambda}(\cdot))$, we wish to characterize its joint Z- and Laplace transform. We define such a transform for general multivariate joint processes with first d-dimensional component $\mathbb{N}_0^d$-valued and second d-dimensional component $\mathbb{R}_+^d$-valued.

Definition B.I. *Let $(\boldsymbol{X}(\cdot), \boldsymbol{Y}(\cdot))$ be a stochastic process taking values in $\mathbb{N}_0^d \times \mathbb{R}_+^d$. For any $t \in \mathbb{R}_+$, the joint transform of $(\boldsymbol{X}(u), \boldsymbol{Y}(u))$ is defined by*

$$\mathcal{J}_{\boldsymbol{X},\boldsymbol{Y}}(u) \equiv \mathcal{J}_{\boldsymbol{X},\boldsymbol{Y}}(u, \boldsymbol{s}, \boldsymbol{z}) := \mathbb{E}\left[\boldsymbol{z}^{\boldsymbol{X}(u)} e^{-\boldsymbol{s}^\top \boldsymbol{Y}(u)}\right] = \mathbb{E}\left[\prod_{i=1}^{d} z_i^{X_i(u)} e^{-s_i Y_i(u)}\right], \tag{B.VI}$$

where $\boldsymbol{s} \in \mathbb{R}_+^d$ and $\boldsymbol{z} \in [-1,1]^d$. The expectation is w.r.t. the filtration at $t = 0$. We call the space of such transforms $\mathbb{J}$; we write $\mathcal{J}_{\boldsymbol{X},\boldsymbol{Y}}(\cdot) \in \mathbb{J}$.

Furthermore, when we have an $\mathbb{N}_0^{d\times d} \times \mathbb{R}_+^{d\times d}$-valued matrix stochastic process $(\boldsymbol{X}(\cdot), \boldsymbol{Y}(\cdot))$ with jth column processes $(\boldsymbol{X}_j(\cdot), \boldsymbol{Y}_j(\cdot))$, then we define $\mathbb{J}^d$ as the d-dimensional analogue of $\mathbb{J}$, with $\mathcal{J}_{\boldsymbol{X},\boldsymbol{Y}}(\cdot) \in \mathbb{J}^d$ defined by

$$\boldsymbol{\mathcal{J}}_{\boldsymbol{X},\boldsymbol{Y}}(u) := \begin{bmatrix} \mathcal{J}_{\boldsymbol{X}_1,\boldsymbol{Y}_1}(u) \\ \vdots \\ \mathcal{J}_{\boldsymbol{X}_d,\boldsymbol{Y}_d}(u) \end{bmatrix}, \tag{B.VII}$$

where $\boldsymbol{\mathcal{J}}_{\boldsymbol{X}_j,\boldsymbol{Y}_j}(\cdot) \in \mathbb{J}$.

As indicated before the definition, our aim is to characterize $\mathcal{J}_{\boldsymbol{Q},\boldsymbol{\lambda}}(\cdot)$ defined by $\mathcal{J}_{\boldsymbol{Q},\boldsymbol{\lambda}}(t) = \mathbb{E}\left[\boldsymbol{z}^{\boldsymbol{Q}(t)} e^{-\boldsymbol{s}^\top \boldsymbol{\lambda}(t)}\right]$, with initial conditions $\boldsymbol{Q}(0) = 0$ and $\boldsymbol{\lambda}(0) = \boldsymbol{\lambda}_0$. We start by using the distributional equalities (B.III) to express $\mathcal{J}_{\boldsymbol{Q},\boldsymbol{\lambda}}(\cdot)$ in the joint transform of $(\boldsymbol{S}^{\boldsymbol{Q}}_j(\cdot), \boldsymbol{S}^{\boldsymbol{\lambda}}_j(\cdot))$.

Theorem B.I. *The joint transform $\mathcal{J}_{\boldsymbol{Q},\boldsymbol{\lambda}}(\cdot)$ satisfies*

$$\mathcal{J}_{\boldsymbol{Q},\boldsymbol{\lambda}}(t, \boldsymbol{s}, \boldsymbol{z}) = \prod_{j=1}^{d} \exp\left(-\lambda_{j,0}\left(t + s_j - \int_0^t \mathcal{J}_{\boldsymbol{S}^{\boldsymbol{Q}}_j, \boldsymbol{S}^{\boldsymbol{\lambda}}_j}(u, \boldsymbol{s}, \boldsymbol{z})\, \mathrm{d}u\right)\right). \tag{B.VIII}$$

Proof. The proof proceeds as the proof of [4], Theorem 1, since that proof only depends on their distributional equalities [4], eqn. (13), which are the same as the equalities (B.III) in our case. □

This theorem shows that we can characterize the probabilistic behavior of $(\boldsymbol{Q}(\cdot), \boldsymbol{\Lambda}(\cdot))$ if we can characterize the joint transform of $(\boldsymbol{S}_j^{\boldsymbol{Q}}(\cdot), \boldsymbol{S}_j^{\boldsymbol{\lambda}}(\cdot))$ for any $j \in [d]$. This will be the subject of the next subsection, where we show that these transforms can be identified as a fixed point of a certain mapping. Furthermore, we will show that iterates of that mapping converge to the fixed point, for *any* starting point, thereby giving an iterative procedure to approximate those joint transforms.

B.II. **Fixed-point theorem and convergence results.** We expressed $\mathcal{J}_{\boldsymbol{Q},\boldsymbol{\lambda}}(\cdot)$ in terms of $\mathcal{J}_{\boldsymbol{S}_j^{\boldsymbol{Q}},\boldsymbol{S}_j^{\boldsymbol{\lambda}}}(\cdot)$, $j \in [d]$. Hence, to obtain a full characterization of $\mathcal{J}_{\boldsymbol{Q},\boldsymbol{\lambda}}(\cdot)$, we need a method to determine $\mathcal{J}_{\boldsymbol{S}_j^{\boldsymbol{Q}},\boldsymbol{S}_j^{\boldsymbol{\lambda}}}(\cdot)$. Write $\mathcal{G}_j(\cdot) = \mathcal{J}_{\boldsymbol{S}_j^{\boldsymbol{Q}},\boldsymbol{S}_j^{\boldsymbol{\lambda}}}(\cdot) \in \mathbb{J}$. We aim to find $\mathcal{G}_j$ for all $j \in [d]$. This is equivalent to finding the vector-valued transform

$$\mathbb{J}^d \ni \boldsymbol{\mathcal{G}}(\cdot) := \boldsymbol{\mathcal{J}}_{\boldsymbol{S}^{\boldsymbol{Q}},\boldsymbol{S}^{\boldsymbol{\lambda}}}(\cdot) = \begin{bmatrix} \mathcal{J}_{\boldsymbol{S}_1^{\boldsymbol{Q}},\boldsymbol{S}_1^{\boldsymbol{\lambda}}}(\cdot) \\ \vdots \\ \mathcal{J}_{\boldsymbol{S}_d^{\boldsymbol{Q}},\boldsymbol{S}_d^{\boldsymbol{\lambda}}}(\cdot) \end{bmatrix} = \begin{bmatrix} \mathcal{G}_1(\cdot) \\ \vdots \\ \mathcal{G}_d(\cdot) \end{bmatrix}. \tag{B.IX}$$

Next, we define a mapping ϕ for which we will prove that $\boldsymbol{\mathcal{G}}(\cdot)$ is a fixed point, and for which iterates of an arbitrary $\boldsymbol{\mathcal{J}}^{(0)}(\cdot) \in \mathbb{J}^d$ will converge to $\boldsymbol{\mathcal{G}}(\cdot)$.

Definition B.II. *Let $\phi : \mathbb{J}^d \to \mathbb{J}^d$ be the mapping defined by*

$$\boldsymbol{\mathcal{J}}(\cdot) = \begin{bmatrix} \mathcal{J}_1(\cdot) \\ \vdots \\ \mathcal{J}_d(\cdot) \end{bmatrix} \mapsto \begin{bmatrix} \phi_1(\mathcal{J}_1, \ldots, \mathcal{J}_d)(\cdot) \\ \vdots \\ \phi_d(\mathcal{J}_1, \ldots, \mathcal{J}_d)(\cdot) \end{bmatrix} = \begin{bmatrix} \phi_1(\boldsymbol{\mathcal{J}})(\cdot) \\ \vdots \\ \phi_d(\boldsymbol{\mathcal{J}})(\cdot) \end{bmatrix} = \phi(\boldsymbol{\mathcal{J}})(\cdot), \tag{B.X}$$

where for $j \in [d]$

$$\begin{aligned} \phi_j(\boldsymbol{\mathcal{J}})(u) &\equiv \phi_j(\boldsymbol{\mathcal{J}})(u, \boldsymbol{s}, \boldsymbol{z}) \\ &= \mathbb{E}_{J,\omega}\left[z_j^{\mathbf{1}\{J>u\}} \prod_{i=1}^{d} e^{-s_i h_{ij,J,\omega}(u)} \prod_{m=1}^{d} \exp\left(-\int_0^u h_{mj,J,\omega}(v)\,(1 - \mathcal{J}_m(u-v, \boldsymbol{s}, \boldsymbol{z}))\ \mathrm{d}v \right) \right]. \end{aligned} \tag{B.XI}$$

We need to prove that this mapping is well-defined, i.e., that for $\boldsymbol{\mathcal{J}}_{\boldsymbol{X},\boldsymbol{Y}}(\cdot) \in \mathbb{J}^d$, also $\phi(\boldsymbol{\mathcal{J}}_{\boldsymbol{X},\boldsymbol{Y}})(\cdot) \in \mathbb{J}^d$. We postpone this until after the next lemma and theorem, which prove that $\boldsymbol{\mathcal{G}}(\cdot)$ is a fixed point of ϕ. After we have proved that ϕ is well-defined, we will prove that it is continuous w.r.t. some appropriate topology. Thereafter, we show that iterates of an arbitrary $\boldsymbol{\mathcal{J}}^{(0)}(\cdot) \in \mathbb{J}^d$ under ϕ will converge to $\boldsymbol{\mathcal{G}}(\cdot)$. This final result describes a method to determine $\mathcal{J}_{\boldsymbol{S}_j^{\boldsymbol{Q}},\boldsymbol{S}_j^{\boldsymbol{\lambda}}}(\cdot)$ explicitly, for any $j \in [d]$, hence completing our characterization of $\mathcal{J}_{\boldsymbol{Q},\boldsymbol{\lambda}}(\cdot)$.

Below, when we prove that $\boldsymbol{\mathcal{G}}(\cdot)$ is a fixed point of ϕ, we need to specify when offspring events arrive exactly, given our knowledge that the offspring events arrive before time u, where u is the remaining time after the arrival of the source event. Remember that offspring events arrive according to an inhomogeneous Poisson process by the cluster representation [1], Definition 2. This implies that those offspring events are positioned in $[0, u]$ according to the normalized restriction to $[0, u]$ of the intensity measure corresponding to the inhomogeneous Poisson process. For $v \in [0, u]$, let $P_{ij,J,\omega}(v|u)$ be the probability that an offspring event in coordinate i caused by an

immigrant in coordinate j was already generated before time v, conditional on being generated before time u, and conditional on J, ω.

Lemma B.I. *Consider the cluster process $\boldsymbol{S}_j^\star$ for $\star \in \{\boldsymbol{N}, \boldsymbol{Q}, \boldsymbol{\lambda}\}$ generated by an immigrant event $(T^{(0)}, J^{(0)}, j)$ in component $j \in [d]$, and let $u = t - T^{(0)}$ be the time elapsed since its arrival. Then the following statements hold.*

(1) Subclusters are i.i.d. modulo the time shift: for each $m \in [d]$, modulo the time shifts T_k corresponding to the arrival times of the first generation events $(T_k^{(1)})_{k\in\mathbb{N}}$, the sequence $\big(\boldsymbol{S}_m^\star(u - T_k)\big)_{k\in[n]}$ is i.i.d., conditional on $\{K_{mj,J^{(0)}} = n\}$ for some $n \in \mathbb{N}$.

(2) For $v \leqslant u$ the probability $P_{ij,J,\omega}(v|u)$ is differentiable with derivative

$$p_{ij,J,\omega}(v|u) = \frac{h_{ij,J,\omega}(v)}{\int_0^u h_{ij,J,\omega}(s)\,\mathrm{d}s}. \tag{B.XII}$$

Proof. The proof is analogous to the proof of [4], Lemma 3. Note that we use that $h_{ij,J}$ is a.s. piecewise continuous, for almost all realizations of $J \sim J_j$, in order to be able to differentiate the probability $P_{ij,J,\omega}(v|u)$. □

Theorem B.II. *The vector of time-dependent joint transforms $\mathcal{G}(\cdot) := \mathcal{J}_{\boldsymbol{S}^{\boldsymbol{Q}},\boldsymbol{S}^{\boldsymbol{\lambda}}}(\cdot)$ satisfies $\mathcal{G}(\cdot) = \phi(\mathcal{G})(\cdot)$.*

Proof. The idea of this proof is the same as that of [4], Theorem 2. It suffices to prove, for arbitrary $j \in [d]$ and $u \geqslant 0$, that $\mathcal{G}_j(u) = \phi_j(\mathcal{G})(u)$. In the proof we keep $\boldsymbol{s}$ and $\boldsymbol{z}$ fixed. Write $\boldsymbol{K}_{j,J}(u) = \begin{bmatrix} K_{1j,J} & \cdots & K_{dj,J} \end{bmatrix}^\top$ for the random (i.e., ω-dependent) vector of Poisson processes of rate $h_{ij,J,\omega}$, where the random excitation functions $h_{ij,J,\omega}$ are assumed to be conditionally independent for $i \in [d]$. By using the distributional equalities (B.V), and with $J \sim J_j$ the lifetime of the immigrant in coordinate j,

$$\begin{aligned}
\mathcal{G}_j(u) &= \mathbb{E}_{J,\omega}\left[\mathbb{E}\left[\prod_{i=1}^d z_i^{S^{\boldsymbol{Q}}_{i\leftarrow j}(u)} e^{-s_i S^{\boldsymbol{\lambda}}_{i\leftarrow j}(u)} \middle| J, \omega\right]\right] \\
&= \mathbb{E}_{J,\omega}\left[\sum_{\boldsymbol{n}\in\mathbb{N}_0^d} \mathbb{E}\left[\prod_{i=1}^d z_i^{S^{\boldsymbol{Q}}_{i\leftarrow j}(u)} e^{-s_i S^{\boldsymbol{\lambda}}_{i\leftarrow j}(u)} \middle| \boldsymbol{K}_{j,J}(u) = \boldsymbol{n}\right] \mathbb{P}\left(\boldsymbol{K}_{j,J}(u) = \boldsymbol{n} \middle| J, \omega\right)\right] \\
&= \mathbb{E}_{J,\omega}\left[c(u) \sum_{\boldsymbol{n}\in\mathbb{N}_0^d} \mathbb{E}\left[\prod_{i=1}^d z_i^{\sum_{m=1}^d \sum_{k=1}^{n_m} S^{\boldsymbol{Q}}_{i\leftarrow m}(u-T_k)} e^{-s_i \sum_{m=1}^d \sum_{k=1}^{n_m} S^{\boldsymbol{\lambda}}_{i\leftarrow m}(u-T_k)}\right] \mathbb{P}\left(\boldsymbol{K}_{j,J}(u) = \boldsymbol{n} \middle| J, \omega\right)\right],
\end{aligned}$$

where

$$c(u) := z_j^{\mathbf{1}\{J>u\}} \prod_{i=1}^d e^{-s_i h_{ij,J,\omega}(u)}.$$

Now we use the i.i.d. nature of the subclusters as in Lemma B.I and the fact that next-generation offspring of a parent is distributed according to a Poisson process, to rewrite the inner expectation as a product over the source components of first-generation events. We let $T^{(mj)}$ be the r.v. with density $p_{mj,J,\omega}(v|u)$ as given in Lemma B.I. These times are distributed as T_k if those were sampled

by $K_{mj,J,\omega}$. This leads to

$$\mathbb{E}\left[\prod_{i=1}^{d} z_i^{\sum_{m=1}^{d}\sum_{k=1}^{n_m} S^{\boldsymbol{Q}}_{i\leftarrow m}(u-T_k)} e^{-s_i \sum_{m=1}^{d}\sum_{k=1}^{n_m} S^{\lambda}_{i\leftarrow m}(u-T_k)}\right]$$
$$= \prod_{m=1}^{d} \mathbb{E}\left[\prod_{i=1}^{d} z_i^{S^{\boldsymbol{Q}}_{i\leftarrow m}(u-T^{(mj)})} e^{-s_i S^{\lambda}_{i\leftarrow m}(u-T^{(mj)})}\right]^{n^m} = \prod_{m=1}^{d} \mathcal{G}_m\left(u-T^{(mj)}\right)^{n_m},$$

whence

$$\mathcal{G}_j(u) = \mathbb{E}_{J,\omega}\left[c(u)\sum_{\boldsymbol{n}\in\mathbb{N}_0^d}\prod_{m=1}^{d}\left(\int_0^u p_{mj,J,\omega}(v|u)\mathcal{G}_m\left(u-v\right)\,\mathrm{d}v\right)^{n_m}\mathbb{P}\left(\boldsymbol{K}_{j,J}(u)=\boldsymbol{n}\middle|J,\omega\right)\right].$$

By Lemma B.I we know that

$$p_{mj,J,\omega}(v|u) = \frac{h_{mj,J,\omega}(v)}{\int_0^u h_{mj,J,\omega}(s)\,\mathrm{d}s},$$

while using that $K_{mj}(\cdot)$ are Poisson processes with intensity $h_{mj,J,\omega}$, we calculate

$$\mathbb{P}\left(\boldsymbol{K}_{j,J}(u)=\boldsymbol{n}\middle|J,\omega\right) = \prod_{m=1}^{d}\frac{\left(\int_0^u h_{mj,J,\omega}(s)\,\mathrm{d}s\right)^{n_m}}{n_m!}\exp\left(-\int_0^u h_{mj,J,\omega}(s)\,\mathrm{d}s\right).$$

Combining the previous three displays,

$$\mathcal{G}_j(u) = \mathbb{E}_{J,\omega}\left[c(u)\sum_{\boldsymbol{n}\in\mathbb{N}_0^d}\prod_{m=1}^{d}\frac{1}{n_m!}\left(\int_0^u h_{mj,J,\omega}(v)\mathcal{G}_m\left(u-v\right)\right)\,\mathrm{d}v\right)^{n_m}\exp\left(-\int_0^u h_{mj,J,\omega}(s)\,\mathrm{d}s\right)\right]$$
$$= \mathbb{E}_{J,\omega}\left[c(u)\prod_{m=1}^{d}\exp\left(\int_0^u h_{mj,J,\omega}(s)\left(\mathcal{G}_m(u-v)-1\right)\,\mathrm{d}s\right)\right].$$

Plugging in the definition of $c(u)$ again, the theorem follows. □

Lemma B.II. *The mapping ϕ from Definition B.II is well-defined, i.e., for $\mathcal{J}_{\boldsymbol{X},\boldsymbol{Y}}(\cdot)\in\mathbb{J}^d$, we have $\phi(\mathcal{J}_{\boldsymbol{X},\boldsymbol{Y}})(\cdot)\in\mathbb{J}^d$.*

Proof. The proof is analogous to the proof of [4], Lemma 1. Our analogue of their Theorem 2 is Theorem B.II. Also, we should replace $B_{ij}g_{ij}$ in their proof by $h_{ij,J,\omega}$. □

For the fixed-point theorem, we wish to show that iterates $\phi^n(\mathcal{J}^{(0)})(u)$ for some arbitrary $\mathcal{J}^{(0)}(\cdot)\in\mathbb{J}^d$ converge to a unique limit, namely the value $\boldsymbol{\mathcal{G}}(u)$ that we are after. To this end, we need an appropriate notion of distance on $\mathbb{J}^d$. We define a norm $\|\cdot\|_{\mathbb{J}^d}$ as a uniform Euclidean norm by

$$\|\mathcal{J}\|_{\mathbb{J}^d} = \sup_{\substack{u\in[0,t],\boldsymbol{s}\in\mathbb{R}_+^d\\ \boldsymbol{z}\in[-1,1]^d}}\|\mathcal{J}(u,\boldsymbol{s},\boldsymbol{z})\| = \sup_{u,\boldsymbol{s},\boldsymbol{z}}\|\mathcal{J}(u,\boldsymbol{s},\boldsymbol{z})\|, \tag{B.XIII}$$

where $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^d$.

Lemma B.III. *The mapping ϕ is continuous w.r.t. $\|\cdot\|_{\mathbb{J}^d}$, if we work on a bounded interval $[0,t]$.*

Proof. The proof is similar to the proof of [4], Lemma 2. Take $\mathcal{J}(\cdot), \tilde{\mathcal{J}}(\cdot) \in \mathbb{J}^d$. It suffices to prove continuity in each coordinate separately, i.e., for each $j \in [d]$ we prove that given $\epsilon > 0$, we can find $\delta > 0$ such that

$$\|\mathcal{J} - \tilde{\mathcal{J}}\|_{\mathbb{J}^d} < \delta$$

implies that $\|\phi_j(\mathcal{J}) - \phi_j(\tilde{\mathcal{J}})\|_{\mathbb{J}} < \epsilon$. We have

$$\begin{aligned}
\|\phi_j(\mathcal{J}) - \phi_j(\tilde{\mathcal{J}})\|_{\mathbb{J}} &= \sup_{u,\boldsymbol{s},z} \left|\phi_j(\mathcal{J})(u,\boldsymbol{s},z) - \phi_j(\tilde{\mathcal{J}})(u,\boldsymbol{s},z)\right| \\
&\leqslant \sup_{u,\boldsymbol{s},z} \mathbb{E}_{J,\omega} \Bigg| \exp\left(\sum_{m=1}^{d}\left(s_m h_{mj,J,\omega}(u) + \int_0^u h_{mj,J,\omega}(v)\,(1 - \mathcal{J}_m(u-v,\boldsymbol{s},z))\;\mathrm{d}v\right)\right) \\
&\qquad - \exp\left(\sum_{m=1}^{d}\left(s_m h_{mj,J,\omega}(u) + \int_0^u h_{mj,J,\omega}(v)\,\left(1 - \tilde{\mathcal{J}}_m(u-v,\boldsymbol{s},z)\right)\;\mathrm{d}v\right)\right)\Bigg| \\
&\leqslant \sup_{u,\boldsymbol{s},z} \mathbb{E}_{J,\omega} \left|\sum_{m=1}^{d}\int_0^u h_{mj,J,\omega}(v)\,\left(\mathcal{J}_m(u-v,\boldsymbol{s},z) - \tilde{\mathcal{J}}_m(u-v,\boldsymbol{s},z)\right)\;\mathrm{d}v\right| \\
&\leqslant \sum_{m=1}^{d} \mathbb{E}_{J,\omega}\left[\int_0^t h_{mj,J,\omega}(v) \sup_{u,\boldsymbol{s},z}\left|\mathcal{J}_m(u-v,\boldsymbol{s},z) - \tilde{\mathcal{J}}_m(u-v,\boldsymbol{s},z)\right|\;\mathrm{d}v\right] \\
&\leqslant \sum_{m=1}^{d} \mathbb{E}_{J,\omega}\|h_{mj,J,\omega}\|_\infty \int_0^t \sup_{u,\boldsymbol{s},z}\left|\mathcal{J}_m(u,\boldsymbol{s},z) - \tilde{\mathcal{J}}_m(u,\boldsymbol{s},z)\right|\;\mathrm{d}v \\
&\leqslant d \max_{m,j\in[d]} \mathbb{E}_{J,\omega}\|h_{mj,J,\omega}\|_\infty t\delta,
\end{aligned}$$

where the first inequality follows by the triangle inequality and the fact that $|z_i| \leqslant 1$; the second by the mean value theorem applied to $x \mapsto e^x$, using that $\mathcal{J}_m(u-v,\boldsymbol{s},z) \leqslant 1$; the third by three more triangle inequalities and positivity of the integrand; the fourth since $h_{mj,J_\omega}(v) < \|h_{mj,J_\omega}\|_\infty$ a.e.; and the fifth is obvious. Note that $\max_{m,j\in[d]} \mathbb{E}_{J,\omega}\|h_{mj,J,\omega}\|_\infty < \infty$ by [1], Definition 2. It follows that if

$$\delta < \frac{\epsilon}{d \max_{m,j\in[d]} \mathbb{E}_{J,\omega}\|h_{mj,J,\omega}\|_\infty t},$$

then $\|\phi_j(\mathcal{J}) - \phi_j(\tilde{\mathcal{J}})\|_{\mathbb{J}} < \epsilon$. □

Remark B.I. *The restriction that we should work on bounded intervals* $[0,t]$ *in Lemma B.III is no obstacle. Whenever we want to use the bound appearing in the proof of this lemma, or want to use continuity of* ϕ*, we do this to find some value* $\mathcal{G}(u)$*. Then we can just take any* $t \geqslant u$*, and apply the lemma.*

We now state the convergence result. For some joint transform $\mathcal{J}^{(0)}(\cdot) \in \mathbb{J}^d$, we construct the sequence $(\mathcal{J}^{(n)}(\cdot))_{n\in\mathbb{N}_0}$ by setting $\mathcal{J}^{(n)}(\cdot) := \phi(\mathcal{J}^{(n-1)})(\cdot)$. By Lemma B.II, we know that $\mathcal{J}^{(n)}(\cdot) \in \mathbb{J}^d$ for all $n \in \mathbb{N}_0$.

Theorem B.III. *For any* $\mathcal{J}^{(0)}(\cdot) \in \mathbb{J}^d$*, the sequence* $(\mathcal{J}^{(n)}(u))_{n\in\mathbb{N}_0}$ *converges pointwise to the fixed point* $\mathcal{G}(u) := \mathcal{J}_{\boldsymbol{S}^Q,\boldsymbol{S}^\lambda}(u)$*. That is, as* $n \to \infty$*, for any* $u \leqslant t$*,*

$$\mathcal{J}^{(n)}(u) \equiv \mathcal{J}^{(n)}(u,\boldsymbol{s},z) \to \mathcal{J}_{\boldsymbol{S}^Q,\boldsymbol{S}^\lambda}(u,\boldsymbol{s},z) \equiv \mathcal{J}_{\boldsymbol{S}^Q,\boldsymbol{S}^\lambda}(u). \tag{B.XIV}$$

Proof. The proof is analogous to the proof of [4], Theorem 3. We should refer to Lemma B.III and Theorem B.II wherever they refer to their Lemma 2 and Theorem 2, respectively. □

Theorem B.III describes how $\mathcal{J}_{\boldsymbol{S}^{\boldsymbol{\varrho}},\boldsymbol{S}^{\boldsymbol{\lambda}}}(u)$ can be approximated. Together with Theorem B.I, this gives a full characterization of the time-dependent joint transform of $(\boldsymbol{Q}(\cdot), \boldsymbol{\lambda}(\cdot))$. This also gives rise to a numerical procedure to characterize $(\boldsymbol{Q}(\cdot), \boldsymbol{\lambda}(\cdot))$. Indeed, take some $\mathcal{J}^{(0)}(\cdot) \in \mathbb{J}^d$, and approximate $\mathcal{J}_{\boldsymbol{S}^{\boldsymbol{\varrho}},\boldsymbol{S}^{\boldsymbol{\lambda}}}(u)$ for a grid of values $\{u_1, \ldots, u_N\} \subset [0, T]$, by iterating ϕ until convergence obtains. Using this, it is possible to approximate $\mathcal{J}_{\boldsymbol{Q},\boldsymbol{\lambda}}(t, \boldsymbol{s}, \boldsymbol{z})$ with the aid of Theorem B.I. Thereafter, this joint transform can either be inverted numerically to determine the multivariate CDF, or it can be differentiated numerically to determine joint moments.

References

[1] J. Baars, R. J. A. Laeven, and M. Mandjes (2026). Delayed Hawkes birth-death processes. *Journal of/Advances in Applied Probability.*

[2] E. Bacry, S. Delattre, M. Hoffmann, and J. F. Muzy (2013). Some limit theorems for Hawkes processes and application to financial statistics. *Stochastic Processes and their Applications* **123**, pp. 2475–2499.

[3] J. Jacod and A. N. Shiryaev (1987). *Limit Theorems for Stochastic Processes*. Grundlehren der Mathematischen Wissenschaften **288**, Springer-Verlag, Berlin.

[4] R. Karim, R. J. A. Laeven, and M. Mandjes (2021). Exact and asymptotic analysis of general multivariate Hawkes processes and induced population processes. Preprint. Available at `https://arxiv.org/abs/2106.03560`.