# 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 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*: July 23, 2025.

## A. Proof of Theorem 4

A.I. **Joint transform characterization.** The proof is a suitable modification of the work done in [2]. We exploit the cluster representation provided in [1], Definition 2. First, we summarize the multidimensional notation as introduced in [2]. 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 $(T_r^{(0)}, J_r^{(0)}, j)$, we denote the $d$-dimensional cluster process it generates as $\boldsymbol{S}_j^{\boldsymbol{N}}(\cdot)$. 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

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

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

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

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

Using the clustering processes defined in (A.I) and (A.II), we can state distributional equalities for the component processes $N_i, Q_i, \lambda_i$. 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

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

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

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

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:

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

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 A.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*

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

*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 $j$th 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 $\boldsymbol{\mathcal{J}}_{\boldsymbol{X},\boldsymbol{Y}}(\cdot) \in \mathbb{J}^d$ defined by*

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

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

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 (A.III) to express $\mathcal{J}_{\boldsymbol{Q},\boldsymbol{\lambda}}(\cdot)$ in the joint transform of $(\boldsymbol{S}_j^{\boldsymbol{Q}}(\cdot), \boldsymbol{S}_j^{\boldsymbol{\lambda}}(\cdot))$

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

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

*Proof.* The proof proceeds as the proof of [2], Theorem 1, since that proof only depends on their distributional equalities [2], eqn. (13), which are the same as the equalities (A.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.

A.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 an 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

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

Next, we define a mapping $\phi$ 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 A.II.** *Let $\phi : \mathbb{J}^d \to \mathbb{J}^d$ be the mapping defined by*

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

*where for $j \in [d]$*

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

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

Below, when we prove that $\mathcal{G}(\cdot)$ is a fixed point of $\phi$, 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 $J, \omega$.

**Lemma A.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 $\left(\boldsymbol{S}_m^\star(u - T_k)\right)_{k\in[n]}$ 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*

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

*Proof.* The proof is analogous to the proof of [2], Lemma 3. Note that we use that $h_{ij,J}$ is a.s. piecewise continuous, for almost all realization of $J \sim J_j$, in order to be able to differentiate the probability $P_{ij,J,\omega}(v|u)$. □

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

*Proof.* The idea of this proof is the same as that of [2], 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., $\omega$-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 (A.V), and with $J \sim J_j$ the lifetime of the immigrant in coordinate $j$,

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

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 A.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 A.I. These times are distributed as $T_k$ if those were sampled by $K_{mj,J,\omega}$. This leads to

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

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 A.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,

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

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

**Lemma A.II.** *The mapping $\phi$ from Definition A.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 [2], Lemma 1. Our analogue of their Theorem 2 is Theorem A.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 $\mathcal{G}(u)$ 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

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

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

**Lemma A.III.** *The mapping $\phi$ 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 [2], 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

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

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

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 $\mathcal{J}_m(u-v, \boldsymbol{s}, \boldsymbol{z}) \leqslant 1$; 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 $\|\phi_j(\boldsymbol{\mathcal{J}}) - \phi_j(\tilde{\boldsymbol{\mathcal{J}}})\|_{\mathbb{J}} < \epsilon$. □

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

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

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

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

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

Theorem A.III describes how $\boldsymbol{\mathcal{J}}_{\boldsymbol{S}^{\boldsymbol{Q}},\boldsymbol{S}^{\lambda}}(u)$ can be approximated. Together with Theorem A.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 $(\boldsymbol{Q}(\cdot), \boldsymbol{\lambda}(\cdot))$: take some $\boldsymbol{\mathcal{J}}^{(0)}(\cdot) \in \mathbb{J}^d$, and approximate $\boldsymbol{\mathcal{J}}_{\boldsymbol{S}^{\boldsymbol{Q}},\boldsymbol{S}^{\lambda}}(u)$ for a grid of values $\{u_1, \dots, u_N\} \subset [0, T]$, by iterating $\phi$ 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 A.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 (2025). Delayed Hawkes birth-death processes. Preprint.

[2] 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`.