1) A causal interpretation in terms of potential outcomes of target parameter in Rytgaard et al. (2022)

We consider a setting similar to the one of Ryalen (2024) and Rytgaard et al. (2022). As in Rytgaard et al. (2022), we consider some measure P on a probability space $(\Omega, \mathcal{F}, P)$. We consider a setting in which we observe a multivariate random measure $N = \left(N^y, N^a, N^\ell\right)$ which is defined on $(\Omega, \mathcal{F})$, where any two of the components do not jump at the same time. These processes are observed in $[0, T]$ for some $T > 0$. Here, N^y denotes an outcome process of interest Y (e.g., death), random measure N^a on $[0, T] \times \mathcal{A}$ for treatment A, where $\mathcal{A}$ is a measurable space; N^ℓ denotes a random measure for covariates L on $[0, T] \times \mathcal{L}$, where $\mathcal{A}$ and $\mathcal{L}$ are measurable spaces; for instance finite subsets of $\mathbb{R}$ and $\mathbb{R}^d$. Numerating these options, we can take

$$\mathcal{A} = \{a_1, ..., a_k\}$$
$$\mathcal{L} = \{l_1, ..., l_m\}.$$

Equivalently (in the sense that the natural filtrations are the same), we may work with the multivariate counting process

$$N(t) = \left(N^y((0, t]), N^a((0, t] \times \{a_1\}), ..., N^a((0, t] \times \{a_k\}), N^\ell((0, t] \times \{l_1\}), ..., N^\ell((0, t] \times \{l_m\})\right).$$

This proces gives rise to a filtration $(\mathcal{F}_t)_{t \geq 0}$, where $\mathcal{F}_t \coloneqq \sigma(N(s) \mid s \leq t)$. Further, we make the assumption of no explosion of N.

We concern ourselves with the hypothetical question if the treatment process N^a had been intervened upon such that treatment was given according to some treatment regime g^*. We will work with an intervention that specifies the treatment decisions but does not change timing of treatment visits. What this means precisely will be made clear below. We are interested in the outcome process Y under this intervention, which we denote by $\tilde{Y}$. Importantly, the intervention is defined as a static/dynamic intervention

$$N^{g^*}(\mathrm{d}t \times \mathrm{d}x) = \pi_t^*(\mathrm{d}x) N^a(\mathrm{d}t \times \mathcal{A})$$

where $\pi_t^*(\mathrm{d}x)$ is some kernel that specifies the treatment decision deterministically at time t in the sense that there are $\mathcal{F}_{T_{(k-1)}} \otimes \mathcal{B}([0, T])$-measurable functions g_k^* which return a treatment decision in $\mathcal{A}$ such that

$$\pi_t^*(\mathrm{d}x) = \sum_k \mathbb{1}\left\{T_{(k-1)} < t \leq T_{(k)}\right\} \delta_{g_k^*\left(\mathcal{F}_{T_{(k-1)}}, T_{(k)}\right)}(\mathrm{d}x),$$

This means that, critically, $N^{g^*}(\mathrm{d}t \times \mathrm{d}x)$ is also a random measure. Note that N^{g^*} has the compensator

$$\mathscr{L}\ (N)(\mathrm{d}t \times \mathrm{d}x) = \pi_t^*(\mathrm{d}x) \underbrace{\Lambda^a(\mathrm{d}t \times \mathcal{A})}_{=:\Lambda^a(\mathrm{d}t)},$$

where $\Lambda^a(\mathrm{d}t)$ is the P-$\mathcal{F}_t$-compensator of $N^a(\mathrm{d}t \times \mathcal{A})$ – also deemed the total P-$\mathcal{F}_t$-compensator of N^a. What this means is that

$$N^a((0, t] \times \mathrm{d}x) - \Lambda^a((0, t]) \times \mathrm{d}x)$$

is a local P-$\mathcal{F}_t$-martingale. We shall write similar notations for the other components of N. Let $\mathscr{L}$ denote the P-$\mathcal{F}_t$-canonical compensator of N^{g^*}. However, N^a generally has the compensator $\Lambda^a(\mathrm{d}t \times \mathrm{d}x) = \pi_t(\mathrm{d}x) \Lambda^a(\mathrm{d}t)$. Now define the time to deviation from the treatment regime as

$$\tau^{g^*} = \inf\{t \geq 0 \mid N^a((0,t] \times \{x\}) \neq N^{g^*}((0,t] \times \{x\}) \quad \text{for some} \quad x \in \mathcal{A}\},$$

when $\mathcal{A} = \{a_1, ..., a_k\}$ consists of a finite set of treatment options.

Definition 1.1: Let $\widetilde{\mathcal{F}}_t := \sigma\left(\widetilde{N}^y(\mathrm{d}s), \widetilde{N}^a(\mathrm{d}s \times \{x\}), \widetilde{N}^\ell(\mathrm{d}s \times \{y\}) \mid s \in (0,t], x \in \mathcal{A}, y \in \mathcal{L}\right)$. Let Λ denote the canonical P-$\mathcal{F}_t$-compensator of N.

A multivariate random measure $\widetilde{N} = \left(\widetilde{N}^y, \widetilde{N}^a, \widetilde{N}^\ell\right)$ is a **counterfactual random measure** under the intervention g^* if it satisfies the following conditions.

1. $\widetilde{N}^a$ has compensator $\mathscr{L}\left(\widetilde{N}\right)$ with respect to $\widetilde{\mathcal{F}}_t$.
2. $\widetilde{N}^x$, has the same compensator $\Lambda\left(\widetilde{N}\right)$ with respect to $\widetilde{\mathcal{F}}_t$ for $x \in \{y, \ell\}$.

Now let $\left(T_{(k)}\right)_k$ denote the ordered event times of N. The main outcome of interest is the counterfactual outcome process $\tilde{Y} := \widetilde{N}^y$; and we wish to identify $\mathbb{E}_P\left[\tilde{Y}_t\right]$.

Let $N^{a,x}(t) := N^a((0,t] \times \{x\})$ for $x \in \mathcal{A}$ and $M^{a,x}(t) := N^{a,x}(t) - \Lambda^{a,x}(t)$. Note that Equation 1 is the same likelihood ratio as in Rytgaard et al. (2022).

Theorem 1.1: If *all* of the following conditions hold:

- **Consistency**: $\tilde{Y}_\cdot \mathbb{1}\{\tau^{g^*} > \cdot\} = Y_\cdot \mathbb{1}\{\tau^{g^*} > \cdot\} \quad P - \text{a.s.}$
- **Exchangeability**: Define $\mathcal{H}_t := \mathcal{F}_t \vee \sigma\left(\tilde{Y}\right)$. Let $\Lambda^{a,a_j}(\mathrm{d}t) = \pi_t\left(\left\{a_j\right\}\right)\Lambda^a(\mathrm{d}t)$ denote the P-$\mathcal{F}_t$-compensator of N^{a,a_j} and $\Lambda_H^{a,a_j}(\mathrm{d}t) = \pi_t^H\left(\left\{a_j\right\}\right)\Lambda_H^a(\mathrm{d}t)$ denote the P-$\mathcal{H}_t$-compensator of N^{a,a_j}. π is indistinguishable from π^H, that is for all $j \in \{1, ..., k\}$ $P\left(\pi_t\left(\left\{a_j\right\}\right) = \pi_t^H\left(\left\{a_j\right\}\right), \forall t \in [0,T]\right) = 1$.
- **Positivity**:

$$W(t) := \prod_{j=1}^{N_t} \left(\prod_{i=1}^{k} \left(\frac{\pi^*_{T_{(j)}}\left(\{a_i\}; \mathcal{F}_{T_{(j-1)}}\right)}{\pi_{T_{(j)}}\left(\{a_i\}; \mathcal{F}_{T_{(j-1)}}\right)} \right)^{\mathbb{1}\left\{A\left(T_{(k)}\right)=a_i\right\}} \right)^{\mathbb{1}\left\{\Delta_{(j)}=a\right\}} \tag{1}$$

  is a uniformly integrable P-$\mathcal{F}_t$-martingale.

  Furthermore, assume that $K(t) = \int_0^t \sum_{j=1}^k \left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1\right) M^{a,a_j}(\mathrm{d}s)$ is a P-$\mathcal{F}_t$-martingale and that K is a process of **locally integrable variation**, meaning that $\mathbb{E}_P\left[\int_0^t |dK(s)|\right] < \infty$ for all $t > 0$.

Then,

$$\mathbb{E}_P\left[\tilde{Y}_t\right] = \mathbb{E}_P[Y_t W(t)]$$

and $W(t) = \mathcal{E}(K)_t$ is a uniformly integrable P-$\mathcal{F}_t$-martingale, where $\mathcal{E}$ denotes the Doléans-Dade exponential (Protter (2005)).

Proof: We shall use that the likelihood ratio solves a specific stochastic differential equation. To this end, note that

$$
\begin{aligned}
W(t) &= \mathcal{E}\left(\sum_{j=1}^{k}\left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})}-1\right)N^{a,a_j}(\mathrm{d}s)\right)_t \\
&\overset{(*)}{=} \mathcal{E}\left(\sum_{j=1}^{k}\left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})}-1\right)N^{a,a_j}(\mathrm{d}s)-\sum_{j=1}^{k}(\pi_s^*(\{a_j\})-\pi_s(\{a_j\}))\Lambda^{a}(\mathrm{d}s)\right)_t \\
&= \mathcal{E}\left(\sum_{j=1}^{k}\left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})}-1\right)N^{a,a_j}(\mathrm{d}s)-\sum_{j=1}^{k}\left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})}-1\right)\Lambda^{a,a_j}(\mathrm{d}s)\right)_t \\
&= \mathcal{E}\left(\sum_{j=1}^{k}\left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})}-1\right)M^{a,a_j}(\mathrm{d}s)\right)_t .
\end{aligned}
$$

In $(*)$, we use that $\sum_{j=1}^{k}\pi_s(\{a_j\})=\sum_{j=1}^{k}\pi_s^*(\{a_j\})=1$.

Thus, by properties of the product integral (e.g., Theorem II.6.1 of Andersen et al. (1993)),

$$
W(t) = 1 + \int_0^t W(s-)\sum_{j=1}^{k}\left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})}-1\right)M^{a,a_j}(\mathrm{d}s). \tag{2}
$$

We have that

$$
\zeta_t := \int_0^t W(s-)\sum_{j=1}^{k}\left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})}-1\right)M^{a,a_j}(\mathrm{d}s)
$$

is a zero mean P-$\mathcal{H}_t$-martingale by positivity. From this, we see that $\int_0^t \tilde{Y}_t\zeta(\mathrm{d}s)$ is also a uniformly integrable P-$\mathcal{H}_t$-martingale by Theorem 2.1.42 of Last & Brandt (1995). This implies that

$$
\begin{aligned}
\mathbb{E}_P[Y_t W(t)] &= \mathbb{E}_P[Y_t \mathbb{1}\{\tau^{g^*} > t\}W(t)] \\
&\overset{(**)}{=} \mathbb{E}_P\left[\tilde{Y}_t \mathbb{1}\{\tau^{g^*} > t\}W(t)\right] \\
&= \mathbb{E}_P\left[\tilde{Y}_t W(t)\right] \\
&= \mathbb{E}_P\left[\tilde{Y}_t \mathbb{E}_P[W(t) \mid \mathcal{H}_0]\right] \\
&= \mathbb{E}_P\left[\tilde{Y}_t W(0)\right] \\
&= \mathbb{E}_P\left[\tilde{Y}_t\right]
\end{aligned}
$$

where in $(**)$ we used consistency. □

Note that in the proof, it suffices that $W(t)$ is uniformly bounded because then it will also be a P-$\mathcal{H}_t$-martingale since it is a local, bounded P-$\mathcal{H}_t$-martingale.

It is also natural to ask oneself: how does our conditions relate to the ones of Ryalen (2024)? The condition of consistency is the same. However, the exchangeability condition and the positivity condition are different in general. We present slightly strengthened versions of the conditions as these are easier to compare. Let $\mathbb{N}_t^a = \mathbb{1}\{\tau^{g^*} \leq t\}$ and let $\mathbb{L}_t$ denote its P-$\mathcal{F}_t$-compensator.

- **Exchangeability**: Define $\mathcal{H}_t := \mathcal{F}_t \vee \sigma(\tilde{Y})$. The P-$\mathcal{F}_t$ compensator for $\mathbb{N}^a$ is also the P-$\mathcal{H}_t$ compensator.
- **Positivity**:

$$\widetilde{W}(t) := \frac{(\mathcal{E}(-\mathbb{N}^a))_t}{(\mathcal{E}(-\mathbb{L}^a))_t} = \mathcal{E}\left(\widetilde{K}\right)_t$$

is uniformly integrable, where $\widetilde{K}_t = \int_0^t \frac{1}{1-\Delta\mathbb{L}_s^a}(\mathbb{N}^a(\mathrm{d}s) - \mathbb{L}^a(\mathrm{d}s))$. Furthermore, $\widetilde{K}$ is a process of **locally integrable variation** and a P-$\mathcal{F}_t$-martingale.

It is unclear at this point whether there exist potential outcomes processes which fulfill the exchangeability condition and the consistency condition for any observed data distribution of N. We leave this question for future research.

1.a) Comparison with Rytgaard et al. (2022)

In Rytgaard et al. (2022), both an exchangeability condition and a positivity condition are presented, but no proof is given that these conditions imply that their target parameter is identified. Our proposal shows that under the conditions of Theorem 1.1, the g-formula given in Rytgaard et al. (2022) causally identifies the counterfactual mean outcome under the assumption that the other martingales are orthogonal to the treatment martingale. Lemma 1 of Ryalen (2024) then gives the desired target parameter. Note that this is weaker than the assumptions in Rytgaard et al. (2022), as they implicitly require that *all* martingales are orthogonal due to their factorization of the likelihood. This is because $\mathcal{E}(X)\mathcal{E}(Y) = \mathcal{E}(X+Y)$ if and only if $[X,Y] = 0$. This can be seen by applying Theorem 38, p. 130 of Protter (2005) and using that the stochastic exponential solves a specific stochastic differential equation.

Theorem 1.1.1 (g-formula): Let, further, $Q = W(T) \cdot P$ denote the probability measure defined by the likelihood ratio $W(T)$ given in Equation 1. Under positivity, then

1. The Q-$\mathcal{F}_t$ compensator of $N^a(\mathrm{d}t \times \mathrm{d}x)$ is $\pi_t^*(\mathrm{d}x)\Lambda_P^a(\mathrm{d}t)$.
2. The Q-$\mathcal{F}_t$ compensator of N^x is Λ_P^x for $x \in \{y, \ell\}$.

Proof: First note that for a local $\mathcal{F}_t$-martingale X in P, we have

$$\int_0^t \frac{1}{W_{s-}} \mathrm{d}\langle W, X\rangle_s^P = \langle K, X\rangle_t^P \tag{3}$$

since we have that $W_t = 1 + \int_0^t W_{s-}\, \mathrm{d}K_s$; whence

$$\langle W, X\rangle_t = \langle 1, X\rangle_t + \langle W_- \bullet K, X\rangle_t = W_{t-} \bullet \langle K, X\rangle_t$$

With $X = M^{a,x}$, we find

$$
\begin{aligned}
\langle K, M^{a,x}\rangle_t^P &= \int_0^t \sum_j \left(\frac{\pi_s^*(a_j)}{\pi_s(a_j)} - 1\right) \mathrm{d}\langle M_P^{a,a_j}, M_P^{a,x}\rangle_s^P \\
&= \int_0^t \left(\frac{\pi_s^*(x)}{\pi_s(x)} - 1\right) \mathrm{d}\langle M_P^{a,x}\rangle_s^P \\
&\quad + \sum_{j \neq x} \int_0^t \left(\frac{\pi_s^*(a_j)}{\pi_s(a_j)} - 1\right) \mathrm{d}\langle M_P^{a,a_j}, M_P^{a,x}\rangle_s^P \\
&= \int_0^t \left(\frac{\pi_s^*(x)}{\pi_s(x)} - 1\right) \pi_s(x)\Lambda_P^a(\mathrm{d}s) \\
&\quad - \int_0^t \left(\frac{\pi_s^*(x)}{\pi_s(x)} - 1\right) \Delta(\pi(x)\Lambda_P^a)_s \pi_s(x)\Lambda_P^a(\mathrm{d}s) \qquad (4) \\
&\quad - \sum_{j \neq x} \int_0^t \left(\frac{\pi_s^*(a_j)}{\pi_s(a_j)} - 1\right) \Delta(\pi(x)\Lambda_P^a)_s \pi_s(a_j)\Lambda_P^a(\mathrm{d}s) \\
&= \int_0^t (\pi_s^*(x) - \pi_s(x))\Lambda_P^a(\mathrm{d}s) \\
&\quad - \sum_j \int_0^t (\pi_s^*(a_j) - \pi_s(a_j))\Delta(\pi(x)\Lambda_P^a)_s \Lambda_P^a(\mathrm{d}s) \\
&= \int_0^t (\pi_s^*(x) - \pi_s(x))\Lambda_P^a(\mathrm{d}s).
\end{aligned}
$$

Girsanov's theorem (Theorem 41, p. 136 of Protter (2005)) together with Equation 3 and Equation 4 gives that

$$
N^a(\mathrm{d}t \times \mathrm{d}x) - \pi_t(\mathrm{d}x)\Lambda_P^a(\mathrm{d}t) - (\pi_t^*(\mathrm{d}x) - \pi_t(\mathrm{d}x))\Lambda_P^a(\mathrm{d}t) = N^a(\mathrm{d}t \times \mathrm{d}x) - \pi_t^*(\mathrm{d}x)\Lambda_P^a(\mathrm{d}t)
$$

is a Q-$\mathcal{F}_t$-local martingale. The second statement follows by noting that

$$
\begin{aligned}
[M^y, K]_t &= \int_0^t \Delta N_t^y \sum_j \left(\frac{\pi_s^*(a_j)}{\pi_s(a_j)} - 1\right) N^{a,a_j}(\mathrm{d}s) \\
&\quad - \int_0^t \Delta \Lambda_P^y(s) \sum_j \left(\frac{\pi_s^*(a_j)}{\pi_s(a_j)} - 1\right) M^{a,a_j}(\mathrm{d}s)
\end{aligned}
$$

where we apply the trick with adding and subtracting the treatment compensators in the second term. The first term is zero because no two counting processes jump at the same time. The second term is a local martingale. This implies $\langle M^y, K\rangle_t^P = 0$. For $x = \ell$ the argument is the same. $\square$

We now provide a sequential representation of the exchangeability condition. It aligns closely with the postulated exchangeability condition in Rytgaard et al. (2022); however, notably on the conditioning set, we include the k'th event time, which is not included in Rytgaard et al. (2022). We conclude that if we have independent marking for the treatment process, the condition in Rytgaard et al. (2022) is sufficient for causal identification.

Theorem 1.1.2: Suppose consistency and positivity holds as in Theorem 1.1. Then, we have

$$\mathbb{E}_P\left[\tilde{Y}_t\right] = \mathbb{E}_P[W_t Y_t],$$

for all $t \in [0, T]$, if for $k \in \mathbb{N}$ and $t \in [0, T]$ it holds that

$$\tilde{Y}_t \perp \mathbb{1}\left\{A\left(T_{(k)}\right) = g^*\left(\mathcal{F}_{T_{(k-1)}}, T_{(k)}\right)\right\} \mid \mathcal{F}^{g^*}_{T_{(k-1)}}, T_{(k)} \leq t, \Delta N^a\left(T_{(k)}\right) = 1,$$

where

$$\mathcal{F}^{g^*}_{T_{(k)}} = \sigma\left(L\left(T_{(k)}\right), \Delta_{(k)}, \mathbb{1}\left\{A\left(T_{(k)}\right) = g^*_k\left(\mathcal{F}_{T_{(k-1)}}, T_{(k)}\right)\right), ..., \mathbb{1}\{A(0) = g^*_0(L(0))\}, L(0)\right)$$

Proof: We see immediately that,

$$\begin{aligned}
&\int W_{s-} \mathbb{1}\left\{T_{(m)} < s < T_{(m+1)} \wedge t\right\} \mathrm{d}K_s \\
&= W_{T_{(m)}} \int \mathbb{1}\left\{T_{(m)} < s < T_{(m+1)} \wedge t\right\} \mathrm{d}K_s \\
&= W_{T_{(m)}} \mathbb{1}\left\{T_{(m)} < t\right\} \left(\sum_j \frac{\pi^*_{T_{(m)}}\left(\{a_j\}\right)}{\pi_{T_{(m)}}\left(\{a_j\}\right)} - 1\right) N^{a,a_j}\left(T_{(m+1)} \wedge t\right) \\
&= W_{T_{(m)}} \mathbb{1}\left\{T_{(m)} < t\right\} \left(\frac{\mathbb{1}\left\{A\left(T_{(m)}\right) = g^*_m\left(\mathcal{F}_{T_{(m-1)}}, T_{(m)}\right)\right\}}{\pi_{T_{(m)}}\left(\left\{g^*_m\left(\mathcal{F}_{T_{(m-1)}}, T_{(m)}\right)\right\}\right)} - 1\right) N^{a,a_j}\left(T_{(m+1)} \wedge t\right)
\end{aligned}$$

By consistency and positivity, the desired result is equivalent to

$$\sum_{m=0}^{\infty} \mathbb{E}_P\left[\int W_{s-} \mathbb{1}\left\{T_{(m)} < s < T_{(m+1)} \wedge t\right\} \mathrm{d}K_s \tilde{Y}_t\right] = 0$$

by Lemma 4 of Ryalen (2024), so

$$
\begin{aligned}
&\sum_{m=0}^{\infty} \mathbb{E}_P\left[\int W_{s-} \mathbb{1}\left\{T_{(m)} < s < T_{(m+1)} \wedge t\right\} \mathrm{d}K_s \tilde{Y}_t\right] \\
&= \sum_{m=0}^{\infty} \mathbb{E}_P\left[W_{T_{(m)}} \mathbb{1}\left\{T_{(m)} < t\right\}\left(\frac{\mathbb{1}\left\{A\left(T_{(m)}\right) = g_m^*\left(\mathcal{F}_{T_{(m)}}, T_{(m+1)}\right)\right\}}{\pi_{T_{(m)}}\left(\left\{g_m^*\left(\mathcal{F}_{T_{(m)}}, T_{(m+1)}\right)\right\}\right)} - 1\right) N^{a,a_j}\left(T_{(m+1)} \wedge t\right) \tilde{Y}_t\right] \\
&= \sum_{m=0}^{\infty} \mathbb{E}_P\Bigg[W_{T_{(m)}} \mathbb{1}\left\{T_{(m)} < t\right\} N^{a,a_j}\left(T_{(m+1)} \wedge t\right) \\
&\quad \times \mathbb{E}_P\left[\left(\frac{\mathbb{1}\left\{A\left(T_{(m)}\right) = g_m^*\left(\mathcal{F}_{T_{(m)}}, T_{(m+1)}\right)\right\}}{\pi_{T_{(m)}}\left(\left\{g_m^*\left(\mathcal{F}_{T_{(m-1)}}, T_{(m)}\right)\right\}\right)} - 1\right) \tilde{Y}_t \mid \mathcal{F}_{T_{(m)}}, T_{(m+1)} \leq \tau, \Delta N^a\left(T_{(k)}\right) = 1\right]\Bigg] \\
&= \sum_{m=0}^{\infty} \mathbb{E}_P\Bigg[W_{T_{(m)}} \mathbb{1}\left\{T_{(m)} < t\right\} N^{a,a_j}\left(T_{(m+1)} \wedge t\right) \\
&\quad \times \mathbb{E}_P\left[\left(\frac{\mathbb{1}\left\{A\left(T_{(m)}\right) = g_m^*\left(\mathcal{F}_{T_{(m)}}, T_{(m+1)}\right)\right\}}{\pi_{T_{(m)}}\left(\left\{g_m^*\left(\mathcal{F}_{T_{(m-1)}}, T_{(m)}\right)\right\}\right)} - 1\right) \mid \mathcal{F}_{T_{(m)}}, T_{(m+1)} \leq, \Delta N^a\left(T_{(k)}\right) = 1\right] \\
&\quad \times \mathbb{E}_P\left[\tilde{Y}_t \mid \mathcal{F}_{T_{(m)}}, T_{(m+1)} \leq \tau, \Delta N^a\left(T_{(k)}\right) = 1\right]\Bigg] \\
&= \sum_{m=0}^{\infty} \mathbb{E}_P[W_{T_{(m)}} \mathbb{1}\left\{T_{(m)} < t\right\} N^{a,a_j}\left(T_{(m+1)} \wedge t\right) \\
&\quad \times \mathbb{E}_P\left[\mathbb{E}_P\left[\left(\frac{\mathbb{1}\left\{A\left(T_{(m)}\right) = g_m^*\left(\mathcal{F}_{T_{(m)}}, T_{(m+1)}\right)\right\}}{\pi_{T_{(m)}}\left(\left\{g_m^*\left(\mathcal{F}_{T_{(m-1)}}, T_{(m)}\right)\right\}\right)} - 1\right) \mid \mathcal{F}_{T_{(m)}}, T_{(m+1)}, \Delta_{(m+1)} = a \mid \mathcal{F}_{T_{(m)}}, T_{(m+1)} \leq \tau, \Delta N^a\left(T_{(k)}\right) = 1\right]\right. \\
&\quad \times \mathbb{E}_P\left[\tilde{Y}_t \mid \mathcal{F}_{T_{(m)}}, T_{(m+1)} \leq \tau, \Delta N^a\left(T_{(k)}\right) = 1\right]\Bigg] \\
&= \sum_{m=0}^{\infty} \mathbb{E}_P\left[W_{T_{(m)}} \mathbb{1}\left\{T_{(m)} < t\right\} N^{a,a_j}\left(T_{(m+1)} \wedge t\right) \times (1-1) \mathbb{E}_P\left[\tilde{Y}_t \mid \mathcal{F}_{T_{(m)}}, T_{(m+1)} \leq \tau, \Delta N^a\left(T_{(k)}\right) = 1\right]\right] \\
&= 0.
\end{aligned}
$$

□

1.b) On the existence of counterfactual processes fulfilling consistency and exchangeability

It is natural to ask oneself whether there exist counterfactual processes for any given law of N such that consistency and exchangeability holds, that is can we extend the law of N to some possibly larger sample space? This question was already posed by Gill & Robins (2001) in the discrete time setting. If this does not hold, then, certainly, we would implicitly be imposing restrictions on the observed data law of N. As the theorem below shows, it is thus harmless to pretend that such counterfactual processes do in fact exist, as they cannot be ruled out by the observed data law of N.

Theorem 1.2.1: For any law of N, we can construct a probability space, wherein a counterfactual process $\widetilde{N}$ and N exists such that (strong) consistency and exchangeability holds. Here (strong) consistency means that

$$\widetilde{N}_t \mathbb{1}\{\tau^{g^*} \geq t\} = N_t \mathbb{1}\{\tau^{g^*} \geq t\} \quad P - \text{a.s.}$$

which implies that

$$\int_0^t \mathbb{1}\{s \leq \tau^{g^*}\} \, \mathrm{d}\widetilde{N}_t = \int_0^t \mathbb{1}\{s \leq \tau^{g^*}\} \, \mathrm{d}N_t,$$

or

$$\widetilde{N}_{\cdot \wedge \tau^{g^*}} = N_{\cdot \wedge \tau^{g^*}}$$

Proof: We provide an argument somewhat analogous to the one given in Section 6 of Gill & Robins (2001). We also consider only the case where $\mathcal{A} = \{0, 1\}$ and the static intervention $g^*\left(\mathcal{F}_{T_{(k-1)}}, T_{(k)}\right) = 1$ for all k. First, we let $\left(T_{(k)}, \Delta_{(k)}, A\left(T_{(k)}\right), L\left(T_{(k)}\right)\right)_{k \in \mathbb{N}}$ denote the marked point process corresponding to N and $(A(0), L(0))$ be the initial values of the treatment and covariate process. First, we generate $L(0)$ from its marginal distribution. Next, for $a_0 = 0, 1$, we generate $\left(L^{a_0}\left(T^{a_0}_{(1)}\right), T^{a_0}_{(1)}, \Delta^{a_0}_{(1)}\right) \sim L\left(T_{(1)}\right), T_{(1)}, \Delta_{(1)} \mid A(0) = a_j, L(0)$ (for each value of a_0, these can be generated independently). Next, for each $a_0 = 0, 1$ and each $a_1 = 0, 1$ where $\Delta^{a_0}_{(1)} = a$, we generate

$$\begin{aligned} &\left(L^{a_0, a_1}\left(T^{a_0, a_1}_{(2)}\right), T^{a_0, a_1}_{(2)}, \Delta^{a_0, a_1}_{(2)}\right) \\ &\sim L\left(T_{(2)}\right), T_{(2)}, \Delta_{(2)} \mid L\left(T_{(1)}\right) = l_1, A\left(T_{(1)}\right) = a_1, T_{(1)} = t_1, \Delta_{(1)} = s_1, A(0) = a_0, L(0) \end{aligned}$$

where $(l_1, t_1, s_1) = \left(L^{a_0}\left(T^{a_0}_{(1)}\right), T^{a_0}_{(1)}, \Delta^{a_0}_{(1)}\right)$ for $\Delta^{a_0}_{(1)} \neq y$ and $T^{a_0}_{(1)} < T$. If $\Delta^{a_0}_{(1)} = y$, put $\left(L^{a_0, a_1}\left(T^{a_0, a_1}_{(2)}\right), T^{a_0, a_1}_{(2)}, \Delta^{a_0, a_1}_{(2)}\right) = (\emptyset, \infty, \emptyset)$. Continue in this manner. Then, we define $\widetilde{N}$ by the marked point process $\left(L^{1, \ldots, 1}\left(T^{1, \ldots, 1}_{(k)}\right), 1, T^{1, \ldots, 1}_{(k)}, \Delta^{1, \ldots, 1}_{(k)}\right)_{k \in \mathbb{N}}$ with initial values in its filtration $(1, L(0))$. Next, we construct the observed data process N. We can generate the A's independently from all other considered random variables. Generate $A(0)$ from its conditional distribution given $L(0)$. Then, let

$$\left(L\left(T_{(1)}\right), T_{(1)}, \Delta_{(1)}\right) = \left(L^{A(0)}\left(T^{A(0)}_{(1)}\right), T^{A(0)}_{(1)}, \Delta^{A(0)}_{(1)}\right).$$

Then, again, generate $A\left(T_{(1)}\right)$ from its conditional distribution given $\mathcal{F}_0, T_{(1)}, \Delta_{(1)} = a$. Afterwards, let

$$\left(L\left(T_{(2)}\right), T_{(2)}, \Delta_{(2)}\right) = \left(L^{A(0), A\left(T_{(1)}\right)}\left(T^{A(0), A\left(T_{(1)}\right)}_{(2)}\right), T^{A(0), A\left(T_{(1)}\right)}_{(2)}, \Delta^{A(0), A\left(T_{(1)}\right)}_{(2)}\right).$$

Continue in this manner. By construction, we then have

$$\left(\widetilde{N}\right)_{t \in [0, T]} \perp A\left(T_{(k)}\right) \mid \mathcal{F}_{T_{(k-1)}}, T_{(k)}, \Delta_{(k)} = a$$

which suffices for exchangeability because it is stronger. Next, we show consistency. Define

$$\tilde{Y}_t = \sum_k \mathbb{1}\left\{T_{(k)}^{1,\ldots,1} \leq t, \Delta_{(k)}^{1,\ldots,1} = y\right\}$$

and

$$Y_t = \sum_k \mathbb{1}\left\{T_{(k)} \leq t, \Delta_{(k)} = y\right\}$$

Now note that

$$\begin{aligned}
Y_t \mathbb{1}\{\tau^{g^*} > t\} &= \sum_k \mathbb{1}\left\{T_{(k)} \leq t, \Delta_{(k)} = y\right\} \mathbb{1}\{\tau^{g^*} > t\} \\
&= \sum_k \mathbb{1}\left\{T_{(k)} \leq t, \Delta_{(k)} = y, A\left(T_{(j)}\right) = 1, \forall j < k\right\} \\
&= \sum_k \mathbb{1}\left\{T_{(k)}^{1,\ldots,1} \leq t, \Delta_{(k)}^{1,\ldots,1} = y, A\left(T_{(j)}\right) = 1, \forall j < k\right\} \\
&= \sum_k \mathbb{1}\left\{T_{(k)}^{1,\ldots,1} \leq t, \Delta_{(k)}^{1,\ldots,1} = y\right\} \mathbb{1}\{\tau^{g^*} > t\} \\
&= \tilde{Y}_t \mathbb{1}\{\tau^{g^*} > t\},
\end{aligned}$$

as desired. □

1.c) Comparison with Coarsening at Random (CAR) conditions of van der Vaart (2004)

Maybe they are the same. Let us define the process by $Z(t) = \left(N^y(t), N^\ell(t), L(t), N^a(t), A(t)\right)$. Consider also its potential outcome process $\tilde{Z} = \left(\tilde{N}^y, \tilde{N}^\ell, \tilde{L}, \tilde{N}^a, 1\right)$. These are both multivariate cadlag processes. Critically, we take $\mathcal{F}_t = \sigma(Z(s), s \leq t)$ – the natural filtration of the observed data process and $\mathcal{H}_t = \mathcal{F}_t \vee \sigma\left(\tilde{Z}(\cdot)\right)$. For simplicity, we consider only the static intervention $g^*\left(\mathcal{F}_{T_{(k-1)}}, T_{(k)}\right) = 1$ for all k.

van der Vaart (2004) introduces a Coarsening at Random (CAR) condition which in our setting can be stated as follows

$$p_{\tau^{g^*}}\left(t \mid \tilde{Z} = \tilde{z}\right) = h(t, \tilde{z}_{\cdot \wedge t}) \tag{5}$$

for some measurable function $h : [0, T] \times D_{[0,T]}\left(\mathbb{R}^d\right) \to [0, 1]$, where $D_{[0,T]}\left(\mathbb{R}^d\right)$ denotes the space of càdlàg functions from $[0, T]$ to $\mathbb{R}^d$. The conditional distributions exist since the sample space is Polish. It is assumed in van der Vaart (2004) that

$$P\left(\tau^{g^*} \in \mathrm{d}t \mid \tilde{Z} = \tilde{z}\right) = p_{\tau^{g^*}}\left(t \mid \tilde{Z} = \tilde{z}\right)\mu(\mathrm{d}t), \tag{6}$$

for some sigma-finite measure μ on $[0, T]$ that does not depend on $\tilde{z}$. Note that

$$P\left(\tau^{g^*} \leq t \mid \tilde{Z} = \tilde{z}\right) = \sum_k P\left(T_{(k)} \leq t, \Delta_{(k)} = a, A\left(T_{(k)}\right) = 0, A\left(T_{(k-1)}\right) = \ldots = A(0) = 1 \mid \tilde{Z} = \tilde{z}\right)$$

as a consequence, we have

$$
\begin{aligned}
P\left(\tau^{g^*} \in \mathrm{d}t \mid \tilde{Z} = \tilde{z}\right) &= \sum_k P\left(T_{(k)} \in \mathrm{d}t, \Delta_{(k)} = a, A\left(T_{(k)}\right) = 0, A\left(T_{(k-1)}\right) = \ldots = A(0) = 1 \mid \tilde{Z} = \tilde{z}\right) \\
&= \sum_k P\left(A\left(T_{(k)}\right) = 0 \mid T_{(k)} = t, \Delta_{(k)} = a, A\left(T_{(k-1)}\right) = \ldots = A(0) = 1, \tilde{Z} = \tilde{z}\right) \\
&\quad \times P\left(T_{(k)} \in \mathrm{d}t, \Delta_{(k)} = a \mid A\left(T_{(k-1)}\right) = \ldots = A(0) = 1, \tilde{Z} = \tilde{z}\right) \\
&= \sum_k P\left(A\left(T_{(k)}\right) = 0 \mid T_{(k)} = t, \Delta_{(k)} = a, A\left(T_{(k-1)}\right) = \ldots = A(0) = 1, \tilde{Z} = \tilde{z}\right) \\
&\quad \times \mathbb{1}(t_k \in \mathrm{d}t, \delta_k = a).
\end{aligned} \tag{7}
$$

This is because given everything else $T_{(k)} \in \mathrm{d}t, \Delta_{(k)} = a$ is a measurable function of $\tilde{Z}$ if deviation has not occured yet. However, due to Equation 7, we see that we do, in fact, not have Equation 6 (formal argument missing) if the t_k's do not have a discrete distribution. One may try to relax Equation 5 to the Markov kernel

$$
P\left(\tau^{g^*} \in \mathrm{d}t \mid \tilde{Z} = \tilde{z}\right) = k_{\tilde{z}}(\mathrm{d}t) = k_z^*(\mathrm{d}t)
$$

depends on $\tilde{z}$ only through z. However, is not clear what the gain of such a condition we are no longer guaranteed a result like Theorem 2.1 of van der Vaart (2004). What Equation 7 suggests is that we work with the following sequential condition:

$$
\tilde{Z} \perp A\left(T_{(k)}\right) \mid T_{(k)}, \Delta_{(k)} = a, \mathcal{F}^{a=1}_{T_{(k-1)}}, \tag{8}
$$

where $\mathcal{F}^{a=1}_t = \sigma\left(\left(N^y(s), N^\ell(s), L(s), N^a(s), 1\right), s \le t\right)$.

From now on, we work with this sequential condition Equation 8 instead of Equation 5.

Theorem 1.3.1: We have that Equation 8 if and only if $\mathbb{1}\{t \le \tau^{g^*}\}\pi_t(1)$, the Radon-Nikodym derivative $\pi_t(1)$ of $\Lambda^{a,1}$ with respect to Λ^a with respect to the filtration $\mathcal{H}_t$ is $\mathcal{F}_t$-adapted.

Proof: This Radon-Nikodym derivative has a version given by

$$
\pi_t(1) = \sum_k \mathbb{1}\left\{T_{(k-1)} < t \le T_{(k)}\right\} P\left(A\left(T_{(k)}\right) = 1 \mid \mathcal{F}_{T_{(k-1)}}, \Delta_{(k)} = a, T_{(k)} = t, \tilde{Z}\right).
$$

This follows by Theorem 4.1.11 of Last & Brandt (1995). Therefore, in general,

$$
\begin{aligned}
&\mathbb{1}\{t \le \tau^{g^*}\}\pi_t(1) \\
&= \sum_k \mathbb{1}\left\{T_{(k-1)} < t \le T_{(k)}\right\}\mathbb{1}\{t \le \tau^{g^*}\} P\left(A\left(T_{(k)}\right) = 1 \mid \mathcal{F}_{T_{(k-1)}}, \Delta_{(k)} = a, T_{(k)} = t, \tilde{Z}\right) \\
&= \sum_k \mathbb{1}\left\{T_{(k-1)} < t \le T_{(k)}\right\}\mathbb{1}\{t \le \tau^{g^*}\} P\left(A\left(T_{(k)}\right) = 1 \mid T_{(k)} = t, \Delta_{(k)} = a, \mathcal{F}^{a=1}_{T_{(k-1)}}, \tilde{Z}\right).
\end{aligned}
$$

From this, we see that the if-part follows immediately. For the only if part, we see that

$$
\begin{aligned}
&\mathbb{1}\left\{T_{(k-1)} < t\right\}\mathbb{1}\left\{t \le \tau^{g^*} \wedge T_{(k)}\right\}\pi_t(1) \\
&\mathbb{1}\left\{T_{(k-1)} < t \le T_{(k)}\right\}\mathbb{1}\{t \le \tau^{g^*}\}\pi_t(1) \\
&= \mathbb{1}\{t \le \tau^{g^*}\}\mathbb{1}\left\{T_{(k-1)} < t \le T_{(k)}\right\} P\left(A\left(T_{(k)}\right) = 1 \mid T_{(k)} = t, \Delta_{(k)} = a, \mathcal{F}^{a=1}_{T_{(k-1)}}, \tilde{Z}\right)
\end{aligned}
$$

Using that $\mathbb{1}\{t \le \tau^{g^*}\}\pi_t(1)$ is $\mathcal{F}_t$-measurable and hence $\mathbb{1}\left\{t \le \tau^{g^*} \wedge T_{(k)}\right\}\pi_t(1)$ is $\mathcal{F}_{T_{(k)}}$-measurable, we have by definition of stopping time σ-algebra that $\mathbb{1}\left\{T_{(k-1)} < t\right\}\mathbb{1}\left\{t \le \tau^{g^*} \wedge T_{(k)}\right\}\pi_t(1)$ is $\mathcal{F}_{T_{(k-1)}}$-measurable. We conclude that, too, $P\left(A\left(T_{(k)}\right) = 1 \mid T_{(k)} = t, \Delta_{(k)} = a, \mathcal{F}^{a=1}_{T_{(k-1)}}, \tilde{Z}\right)$ must be $\mathcal{F}_{T_{(k-1)}}$-measurable for all t with $T_{(k-1)} < t \le T_{(k)}$ and $t \le \tau^{g^*}$. This is all in the support (?). □

Note that if one exchangeability in terms of Theorem 1.1 holds, then both identification formulas are the same i.e.,

$$\mathbb{E}_P\left[\tilde{Y}_t\right] = \mathbb{E}_P[W_t Y_t] = \mathbb{E}_P\left[\widetilde{W}_t Y_t\right],$$

Theorem 1.3.2: Consider the static intervention with $g^*\left(\mathcal{F}_{T_{(k-1)}}, T_{(k)}\right) = 1$ for all k. and $\mathcal{A} = \{0, 1\}$. Suppose consistency for $\tilde{Y}_t$ as in Theorem 1.1. Let $\pi^{\mathcal{H}}_t(1)$, $\pi^{\mathcal{F}}_t(1)$ denote the Radon-Nikodym derivatives of $\Lambda^{a,1}$ with respect to Λ^a in the filtrations $\mathcal{H}_t$ and $\mathcal{F}_t$, respectively. Similarly, let $\Lambda^{a,\mathcal{H}}$, $\Lambda^{a,\mathcal{F}}$ denote the compensators of N^a in the filtrations $\mathcal{H}_t$ and $\mathcal{F}_t$, respectively. Suppose that a version of $\pi^{\mathcal{F}}_s)$ exists such that

$$\pi^{\mathcal{F}}_s(1, \omega) < 1 \tag{9}$$

for all $s \in [0, T]$ and $\omega \in \Omega^*$ with $P(\Omega^*) = 1$. Also, suppose that strong consistency holds for $\widetilde{N}^a_t$, i.e.,

$$\widetilde{N}^a_\cdot \mathbb{1}\{\tau^{g^*} \ge \cdot\} = N^a_\cdot \mathbb{1}\{\tau^{g^*} \ge \cdot\} \quad P-\text{a.s.}$$

Let $\mathcal{H}_t$ denote the initial expansion of $\mathcal{F}_t$ with $\sigma\left(\widetilde{N}^a_\cdot, \tilde{Y}_\cdot\right)$. If the compensator of $\mathbb{N}^a$ in the filtration $\mathcal{H}_t$ is $\mathcal{F}_t$-predictable, then $N^a\left(\cdot \wedge \tau^{g^*}\right)$ is $\mathcal{F}_t$-predictable.

Proof:

Note that, under the stated conditions

$$\int_0^\cdot \mathbb{1}\{s \le \tau^{g^*}\} N^a(\mathrm{d}s)$$

is a $\mathcal{H}_t$-predictable process. Therefore, it is its own compensator in $\mathcal{H}_t$. Now note that

$$\begin{aligned}
\mathbb{L}^{\mathcal{H}}_\cdot &= \int_0^\cdot (1 - \pi^{\mathcal{H}}_s(1))\mathbb{1}\{s \le \tau^{g^*}\}\Lambda^{a,\mathcal{H}}(\mathrm{d}s) \\
&= \int_0^\cdot (1 - \pi^{\mathcal{H}}_s(1))\mathbb{1}\{s \le \tau^{g^*}\}N^a(\mathrm{d}s).
\end{aligned}$$

On the other hand,

$$\mathbb{L}^{\mathcal{F}}_\cdot = \int_0^\cdot \left(1 - \pi^{\mathcal{F}}_s(1)\right)\mathbb{1}\{s \le \tau^{g^*}\}\Lambda^{a,\mathcal{F}}(\mathrm{d}s)$$

Conclude that if $\mathbb{L}^{\mathcal{H}}$ is indistinguishable from $\mathbb{L}^{\mathcal{F}}$, then

$$(1 - \pi^{\mathcal{H}}_\cdot(1))\mathbb{1}\{\cdot \le \tau^{g^*}\}\Delta N^a(\cdot) = \left(1 - \pi^{\mathcal{F}}_\cdot(1)\right)\mathbb{1}\{\cdot \le \tau^{g^*}\}\Delta\Lambda^{a,\mathcal{F}}(\cdot).$$

Therefore,

$$\mathbb{L}^{\mathcal{H}}_{\cdot} = \int_0^{\cdot} \big(1 - \pi^{\mathcal{F}}_s(1)\big) \Delta\Lambda^{a,\mathcal{F}}(s) \mathbb{1}\{s \leq \tau^{g^*}\} N^a(\mathrm{d}s)$$

on a set up to P-indistinguishability. If this were $\mathcal{F}_t$-predictable, then

$$\begin{aligned}\mathbb{L}^{\mathcal{H}}_{\cdot} &= \int_0^{\cdot} \big(1 - \pi^{\mathcal{F}}_s(1)\big) \Delta\Lambda^{a,\mathcal{F}}(s) \mathbb{1}\{s \leq \tau^{g^*}\} N^a(\mathrm{d}s) \\ &= \int_0^{\cdot} \big(1 - \pi^{\mathcal{F}}_s(1)\big) \Delta\Lambda^{a,\mathcal{F}}(s) \mathbb{1}\{s \leq \tau^{g^*}\} M^{a,\mathcal{F}}(\mathrm{d}s) + \int_0^{t} \big(1 - \pi^{\mathcal{F}}_s(1)\big) \Delta\Lambda^{a,\mathcal{F}}(s) \mathbb{1}\{s \leq \tau^{g^*}\} \Lambda^{a,\mathcal{F}}(\mathrm{d}s)\end{aligned}$$

Conclude that since $\int_0^{\cdot} \big(1 - \pi^{\mathcal{F}}_s(1)\big) \Delta\Lambda^{a,\mathcal{F}}(s) \mathbb{1}\{s \leq \tau^{g^*}\} M^{a,\mathcal{F}}(\mathrm{d}s)$ is a difference of two non-decreasing processes, and hence of finite variation, a local martingale, by p. 115 of Protter (2005) that

$$\int_0^{\cdot} \big(1 - \pi^{\mathcal{F}}_s(1)\big) \Delta\Lambda^{a,\mathcal{F}}(s) \mathbb{1}\{s \leq \tau^{g^*}\} N^a(\mathrm{d}s) = \int_0^{\cdot} \big(1 - \pi^{\mathcal{F}}_s(1)\big) \Delta\Lambda^{a,\mathcal{F}}(s) \mathbb{1}\{s \leq \tau^{g^*}\} \Lambda^{a,\mathcal{F}}(\mathrm{d}s).$$

Then, we have that $\mathbb{1}\{\cdot \leq \tau^{g^*}\} \Delta\Lambda^{a,\mathcal{F}}(T^a_k) = -1$ by Equation 9 for all jump times T^a_k of N^a. However, we also have that

$$\mathbb{1}\{s \leq \tau^{g^*}\} \Delta\Lambda^{a,\mathcal{F}}(s) = 0$$

whenever s is not a jump time of N^a by Equation 9. This shows the statement for the discrete part of the compensator. For the continuous part, we are now able to conclude from these two parts that

$$\int_0^{\cdot} \big(1 - \pi^{\mathcal{F}}_s(1)\big) \Delta\Lambda^{a,\mathcal{F}}(s) \mathbb{1}\{s \leq \tau^{g^*}\} \lambda^{a,\mathcal{F}}(s) \, \mathrm{d}s = 0$$

from which we conclude that

$$\mathbb{1}\{s \leq \tau^{g^*}\} \lambda^{a,\mathcal{F}}(s) = 0$$

for almost all s with respect to the Lebesgue measure and almost all $\omega \in \Omega^{**}$ with $P(\Omega^{**}) = 1$. However, these two restrictions exactly mean that $\mathbb{1}\{s \leq \tau^{g^*}\} N^a(\mathrm{d}s)$ is a $\mathcal{F}_t$-predictable process. □

This is because

$$\mathbb{L}^a_t = \int_0^t (1 - \pi_s(1)) \mathbb{1}\{s \leq \tau^{g^*}\} M^a(\mathrm{d}s) + \int_0^t (1 - \pi_s(1)) \mathbb{1}\{s \leq \tau^{g^*}\} \Lambda^a(\mathrm{d}s).$$

which implies under regularity conditions that $\mathbb{1}\{s \leq \tau^{g^*}\} N^a(\mathrm{d}s)$ is a predictable process in $\mathcal{F}_t$. Now, we calculate

$$
\begin{aligned}
\mathbb{K}_t^a &= -\int_0^t \frac{1}{1-\Delta \mathbb{L}_s^a} \,\mathrm{d}\mathbb{M}_s^a \\
&= -\int_0^t \frac{1}{1-(1-\pi_s(1))\mathbb{1}\{s \le \tau^{g^*}\}\Delta N^a(s)} \,\mathrm{d}\mathbb{M}_s^a \\
&= -\int_0^t \frac{1}{1-(1-\pi_s(1))\mathbb{1}\{s \le \tau^{g^*}\}\Delta N^a(s)} \,\mathrm{d}\mathbb{N}_s^a \\
&+ \int_0^t \frac{1}{1-(1-\pi_s(1))\mathbb{1}\{s \le \tau^{g^*}\}\Delta N^a(s)} \,\mathrm{d}\mathbb{L}_s^a \\
&= -\int_0^t \frac{1}{1-(1-\pi_s(1))\mathbb{1}\{s \le \tau^{g^*}\}\Delta N^a(s)} \,\mathrm{d}\mathbb{N}_s^a \\
&+ \int_0^t \frac{1}{1-(1-\pi_s(1))\mathbb{1}\{s \le \tau^{g^*}\}\Delta N^a(s)} (1-\pi_s(1))\mathbb{1}\{s \le \tau^{g^*}\} N^a(\mathrm{d}s) \\
&= -\int_0^t \frac{1}{1-(1-\pi_s(1))\mathbb{1}\{s \le \tau^{g^*}\}\Delta N^a(s)} \,\mathrm{d}\mathbb{N}_s^a \\
&+ \int_0^t \frac{1}{\pi_s(1)} (1-\pi_s(1))\mathbb{1}\{s \le \tau^{g^*}\} N^a(\mathrm{d}s) \\
&= K_t^*
\end{aligned}
$$

On the other hand, the same argument shows that if the other exchangeability condition holds then $\pi_s(1)\mathbb{1}\{s \le \tau^{g^*}\}$ can be chosen $\mathcal{F}_t$-adapted, since we can pick

$$
\pi_s(1) = \sum_k \mathbb{1}\left\{T_{(k-1)} < s \le T_{(k)}\right\}\left(1 - \Delta \mathbb{L}^a_{T_{(k-1)}}\right)
$$

By construction, $\pi_s(1)$ is the desired Radon-Nikodym derivative with respect to $\mathcal{H}_t$. Moreover, it is $\mathcal{F}_t$-adapted since $\Delta \mathbb{L}^a_{T_{(k-1)}}$ is $\mathcal{F}_{T_{(k-1)}}$-measurable.x Therefore, again, the weight $W(t)$ is a P-$\mathcal{H}_t$-martingale as well and $\mathcal{F}_t$-adapted under integrability conditions. Therefore, we have that

$$
\mathbb{E}_P\left[\tilde{Y}_t\right] = \mathbb{E}_P[W_t Y_t]
$$

also. Now calculate

$$
\begin{aligned}
K_t^* &= \int_0^t \mathbb{1}\{s \le \tau^{g^*}\}\left(\frac{1}{\pi_s(1)} - 1\right) N^{a,1}(\mathrm{d}s) + \int_0^t \mathbb{1}\{s \le \tau^{g^*}\}(-1) N^{a,0}(\mathrm{d}s) \\
&= \int_0^t \mathbb{1}\{s \le \tau^{g^*}\} \sum_k \mathbb{1}\left\{T_{(k-1)} < s \le T_{(k)}\right\}\left(\frac{1}{\pi_s(1)} - 1\right) N^{a,1}(\mathrm{d}s) + \int_0^t \mathbb{1}\{s \le \tau^{g^*}\}(-1) N^{a,0}(\mathrm{d}s) \\
&= \sum_k \int_{T_{(k-1)} \wedge t}^{T_{(k)} \wedge t} \mathbb{1}\{s \le \tau^{g^*}\}\left(\frac{1}{1-\Delta \mathbb{L}^a_{T_{(k-1)}}} - 1\right) N^{a,1}(\mathrm{d}s) + \int_0^t \mathbb{1}\{s \le \tau^{g^*}\}(-1) N^{a,0}(\mathrm{d}s) \\
&= \sum_k \int_{T_{(k-1)} \wedge t}^{T_{(k)} \wedge t} \mathbb{1}\{s \le \tau^{g^*}\} \frac{\Delta \mathbb{L}^a_{T_{(k-1)}}}{1-\Delta \mathbb{L}^a_{T_{(k-1)}}} N^{a,1}(\mathrm{d}s) + \int_0^t \mathbb{1}\{s \le \tau^{g^*}\}(-1) N^{a,0}(\mathrm{d}s) \\
&= \mathbb{K}_t^a.
\end{aligned}
$$

1.d) Comparison between Ryalen (2024) and Rytgaard et al. (2022)

Consider a simple example where $N^a(t) \leq 1$ for all t, and consists of the multivariate counting process $N = (N^y, N^{a,a_0}, N^{a,a_1})$. We consider the intervention $\pi_t^*(a_1) = 1$ for all t. Suppose that $(N^y, N^{a,a_0}, N^{a,a_1})$ has compensator

$$\begin{aligned} \Lambda^y(\mathrm{d}t)(P) &= \lambda_t^y \,\mathrm{d}t, \\ \Lambda^{a,a_j}(\mathrm{d}t)(P) &= \pi_t(a_j)\lambda_t^a \,\mathrm{d}t, j = 0, 1. \end{aligned}$$

with respect to $\mathcal{F}_t$ in P. In P, note that

$$\begin{aligned} \mathbb{E}_P[Y_t] &= \mathbb{E}_P[N^y(t)] \\ &= \mathbb{E}_P\Big[\mathbb{1}\big\{T_{(1)} \leq t, \Delta_{(1)} = y\big\} \\ &+\mathbb{1}\big\{\Delta_{(1)} = a, A\big(T_{(1)}\big) = 1, T_{(2)} \leq t, \Delta_{(2)} = y\big\} \\ &+\mathbb{1}\big\{\Delta_{(1)} = a, A\big(T_{(1)}\big) = 0, T_{(2)} \leq t, \Delta_{(2)} = y\big\}\Big] \\ &= \int_0^t \exp\left(-\int_0^s (\lambda_u^y + \lambda_u^{a,a_1} + \lambda_u^{a,a_0}) \,\mathrm{d}u\right) \lambda_s^y \,\mathrm{d}s \\ &+ \int_0^t \exp\left(-\int_0^s (\lambda_u^y + \lambda_u^{a,a_1} + \lambda_u^{a,a_0}) \,\mathrm{d}u\right) \lambda_s^{a,a_1} \\ &\quad \times \int_s^t \exp\left(-\int_s^v \lambda_u^y\right) \mathrm{d}u) \lambda_v^y \,\mathrm{d}v \,\mathrm{d}s \\ &+ \int_0^t \exp\left(-\int_0^s (\lambda_u^y + \lambda_u^{a,a_1} + \lambda_u^{a,a_0}) \,\mathrm{d}u\right) \lambda_s^{a,a_0} \\ &\quad \times \int_s^t \exp\left(-\int_s^v \lambda_u^y\right) \mathrm{d}u) \lambda_v^y \,\mathrm{d}v \,\mathrm{d}s \end{aligned}$$

Then, in Q, we have

$$\begin{aligned} \Lambda^y(\mathrm{d}t)(Q) &= \lambda_t^y \,\mathrm{d}t, \\ \Lambda^{a,a_0}(\mathrm{d}t)(Q) &= 0, \\ \Lambda^{a,a_1}(\mathrm{d}t)(Q) &= \lambda_t^a \,\mathrm{d}t. \end{aligned}$$

Therefore,

$$
\begin{aligned}
\mathbb{E}_Q[Y_t] &= \mathbb{E}_Q[N^y(t)] \\
&= \mathbb{E}_Q\Big[\mathbb{1}\Big\{T_{(1)} \leq t, \Delta_{(1)} = y\Big\} \\
&+ \mathbb{1}\Big\{\Delta_{(1)} = a, A\Big(T_{(1)}\Big) = 1, T_{(2)} \leq t, \Delta_{(2)} = y\Big\}\Big] \\
&= \int_0^t \exp\left(-\int_0^s (\lambda_u^y + \lambda_u^a)\,\mathrm{d}u\right)\lambda_s^y\,\mathrm{d}s \\
&+ \int_0^t \exp\left(-\int_0^s (\lambda_u^y + \lambda_u^a)\,\mathrm{d}u\right)\lambda_s^a \\
&\quad \times \int_s^t \exp\left(-\int_s^v \lambda_u^y\,\mathrm{d}u\right)\lambda_v^y\,\mathrm{d}v\,\mathrm{d}s \\
&= \int_0^t \exp\left(-\int_0^s (\lambda_u^y + \lambda_u^a)\,\mathrm{d}u\right)\lambda_s^y\,\mathrm{d}s \\
&+ \int_0^t \exp\left(-\int_0^s (\lambda_u^y + \lambda_u^a)\,\mathrm{d}u\right)\lambda_s^a \\
&\quad \times \left(1 - \exp\left(-\int_s^t \lambda_u^y\,\mathrm{d}u\right)\right)\mathrm{d}s \\
&= 1 - \exp\left(-\int_0^t (\lambda_u^y + \lambda_u^a)\,\mathrm{d}u\right) - \int_0^t \exp\left(-\int_0^s (\lambda_u^y + \lambda_u^a)\,\mathrm{d}u\right)\lambda_s^a\,\mathrm{d}s \\
&+ \int_0^t \exp\left(-\int_0^s (\lambda_u^y + \lambda_u^a)\,\mathrm{d}u\right)\lambda_s^a \\
&\quad \times \left(1 - \exp\left(-\int_s^t \lambda_u^y\,\mathrm{d}u\right)\right)\mathrm{d}s \\
&= 1 - \exp\left(-\int_0^t (\lambda_u^y + \lambda_u^a)\,\mathrm{d}u\right) \\
&- \int_0^t \exp\left(-\int_0^s (\lambda_u^y + \lambda_u^a)\,\mathrm{d}u\right)\lambda_s^a \exp\left(-\int_s^t \lambda_u^y\,\mathrm{d}u\right)\mathrm{d}s \\
&\overset{!}{=} 1 - \exp\left(-\int_0^t (\lambda_u^y + \lambda_u^a)\,\mathrm{d}u\right) \\
&- \exp\left(-\int_0^t \lambda_u^y\,\mathrm{d}u\right)\int_0^t \exp\left(-\int_0^s (\lambda_u^a)\,\mathrm{d}u\right)\lambda_s^a\,\mathrm{d}s \\
&= 1 - \exp\left(-\int_0^t (\lambda_u^y + \lambda_u^a)\,\mathrm{d}u\right) \\
&- \exp\left(-\int_0^t \lambda_u^y\,\mathrm{d}u\right)\left(1 - \exp\left(-\int_0^t (\lambda_u^a)\,\mathrm{d}u\right)\right) \\
&= 1 - \exp\left(-\int_0^t \lambda_u^y\,\mathrm{d}u\right)
\end{aligned}
$$

In (!) use that $\lambda^y(u)$ does not depend on s at all (jump time a). With constant intensities and constant treatment probabilities, this reduces to

$$
\begin{aligned}
\mathbb{E}_Q[Y_t] = & \int_0^t \exp(-(\lambda^y + \lambda^a)s)\lambda^y \,\mathrm{d}s \\
& + \int_0^t \exp(-(\lambda^y + \lambda^a)s)\lambda^a \\
& \times (1 - \exp(-\lambda^y(t-s)))\,\mathrm{d}s \\
& = \left(\frac{\lambda^y}{\lambda^y + \lambda^a}\right)(1 - \exp(-(\lambda^y + \lambda^a)t)) \\
& + \left(\frac{\lambda^a}{\lambda^y + \lambda^a}\right)(1 - \exp(-(\lambda^y + \lambda^a)t)) - \lambda^a \frac{\exp(-\lambda^y t)}{\lambda^a}(1 - \exp(-\lambda^a t)) \\
& = (1 - \exp(-(\lambda^y + \lambda^a)t)) - \exp(-\lambda^y t)(1 - \exp(-\lambda^a t)) \\
& = 1 - \exp(-\lambda^y t)
\end{aligned}
$$

Now consider "a going to the Tivoli example" in which the visit itself directly affects the probability of dying.

$$
\begin{aligned}
\mathbb{E}_Q[Y_t] = & \int_0^t \exp(-(\lambda^y + \lambda^a)s)\lambda^y \,\mathrm{d}s \\
& + \int_0^t \exp(-(\lambda^y + \lambda^a)s)\lambda^a \\
& \times (1 - \exp(-\lambda^{y,2}(t-s)))\,\mathrm{d}s \\
& = \left(\frac{\lambda^y}{\lambda^y + \lambda^a}\right)(1 - \exp(-(\lambda^y + \lambda^a)t)) \\
& + \left(\frac{\lambda^a}{\lambda^y + \lambda^a}\right)(1 - \exp(-(\lambda^y + \lambda^a)t)) - \lambda^a \frac{\exp(-\lambda^{y,2} t)}{\lambda^a + \lambda^y + \lambda^{y,2}}(1 - \exp(-(\lambda^a + \lambda^y + \lambda^{y,2})t)) \\
& = 1 - \exp(-(\lambda^y + \lambda^a)t) - \lambda^a \frac{\exp(-\lambda^{y,2} t)}{\lambda^a + \lambda^y - \lambda^{y,2}}(1 - \exp(-(\lambda^a + \lambda^y - \lambda^{y,2})t))
\end{aligned}
$$

However, in $\tilde{Q}$ (Ryalen), we have

$$
\begin{aligned}
\Lambda^y(\mathrm{d}t)\left(\tilde{Q}\right) &= \frac{\lambda_t^y \,\mathrm{d}t - 0}{1 - 0}, \\
\Lambda^{a,a_0}(\mathrm{d}t)\left(\tilde{Q}\right) &= \frac{0 - 0}{1 - 0} \\
\Lambda^{a,a_1}(\mathrm{d}t)\left(\tilde{Q}\right) &= \frac{\lambda_t^a \,\mathrm{d}t - (1 - \pi_{t \wedge \tau^{g^*}}(a_1))\lambda_{t \wedge \tau^{g^*}}^a \,\mathrm{d}t}{1 - 0} = \pi_{t \wedge \tau^{g^*}}(a_1)\lambda_{t \wedge \tau^{g^*}}^a \,\mathrm{d}t = \pi_t(a_1)\lambda_t^a \,\mathrm{d}t.
\end{aligned}
$$

since $\tau^{g^*} = \infty$ almost surely in $\tilde{Q}$. With constant intensities, we get the same as above plugging in $\pi\lambda^a$ instead of λ^a. Therefore,

$$
\begin{aligned}
\mathbb{E}_{\widetilde{Q}}[Y_t] &= \mathbb{E}_{\widetilde{Q}}[N^y(t)] \\
&= \mathbb{E}_{\widetilde{Q}}\Big[\mathbb{1}\Big\{T_{(1)} \leq t, \Delta_{(1)} = y\Big\} \\
&+\mathbb{1}\Big\{\Delta_{(1)} = a, A\Big(T_{(1)}\Big) = 1, T_{(2)} \leq t, \Delta_{(2)} = y\Big\}\Big] \\
&= \int_0^t \exp\left(-\int_0^s (\lambda_u^y + \pi_u(a_1)\lambda_u^a)\,\mathrm{d}u\right)\lambda_s^y\,\mathrm{d}s \\
&+ \int_0^t \exp\left(-\int_0^s (\lambda_u^y + \pi_u(a_1)\lambda_u^a)\,\mathrm{d}u\right)\pi_s(a_1)\lambda_s^a \\
&\quad \times \int_s^t \exp\left(-\int_s^v \lambda_u^y\,\mathrm{d}u\right)\lambda_v^y\,\mathrm{d}v\,\mathrm{d}s \\
&= \int_0^t \exp\left(-\int_0^s (\lambda_u^y + \pi_u(a_1)\lambda_u^a)\,\mathrm{d}u\right)\lambda_s^y\,\mathrm{d}s \\
&+ \int_0^t \exp\left(-\int_0^s (\lambda_u^y + \pi_u(a_1)\lambda_u^a)\,\mathrm{d}u\right)\pi_s(a_1)\lambda_s^a \\
&\quad \times \left(1 - \exp\left(-\int_s^t \lambda_u^y\,\mathrm{d}u\right)\right)\mathrm{d}s
\end{aligned}
$$

We now consider a more natural example, where the difference is due to time-varying confounding. Påls functional can be written as

$$
\begin{aligned}
&\int_0^t \exp\left(-\int_0^s \left(\lambda_u^y + \pi_u(a_1)\lambda_u^a + \lambda_u^\ell\right) \mathrm{d}u\right) \lambda_s^y \,\mathrm{d}s \\
&+ \int_0^t \exp\left(-\int_0^s \left(\lambda_u^y + \pi_u(a_1)\lambda_u^a + \lambda_u^\ell\right) \mathrm{d}u\right) \pi_s(a_1)\lambda_s^a \\
&\quad \times \int_s^t \exp\left(-\int_s^v \left(\lambda_u^y + \lambda_u^\ell\right) \mathrm{d}u\right) \left(\lambda_v^y + \lambda_v^\ell \int_v^t \exp\left(-\int_v^w \lambda_u^y(\ell_2; v) \,\mathrm{d}u\right) \lambda_w^y(\ell_2; v) \,\mathrm{d}w\right) \mathrm{d}v \,\mathrm{d}s \\
&+ \int_0^t \exp\left(-\int_0^s \left(\lambda_u^y + \pi_u(a_1)\lambda_u^a + \lambda_u^\ell\right) \mathrm{d}u\right) \lambda_s^\ell \\
&\quad \times \int_s^t \exp\left(-\int_s^v \left(\lambda_u^y(\ell_1; s) + \pi_u(a_1)(\ell_1; s)\lambda_u^a(\ell_1; s)\right) \mathrm{d}u\right) \\
&\quad \times \left(\lambda_v^y(\ell_1; s) + \pi_v(a_1)(\ell_1; s)\lambda_v^a(\ell_1; s) \int_v^t \exp\left(-\int_v^w \lambda_u^y(\ell_1; s) \,\mathrm{d}u\right) \lambda_w^y(\ell_1; s) \,\mathrm{d}w\right) \mathrm{d}v \,\mathrm{d}s \\
&= \int_0^t \exp\left(-\int_0^s \left(\lambda_u^y + \pi_u(a_1)\lambda_u^a + \lambda_u^\ell\right) \mathrm{d}u\right) \lambda_s^y \,\mathrm{d}s \\
&+ \int_0^t \exp\left(-\int_0^s \left(\lambda_u^y + \pi_u(a_1)\lambda_u^a + \lambda_u^\ell\right) \mathrm{d}u\right) \pi_s(a_1)\lambda_s^a \\
&\quad \times \int_s^t \exp\left(-\int_s^v \left(\lambda_u^y + \lambda_u^\ell\right) \mathrm{d}u\right) \left(\lambda_v^y + \lambda_v^\ell \left(1 - \exp\left(-\int_v^t \lambda_u^y(\ell_2; v) \,\mathrm{d}u\right)\right)\right) \mathrm{d}v \,\mathrm{d}s \\
&+ \int_0^t \exp\left(-\int_0^s \left(\lambda_u^y + \pi_u(a_1)\lambda_u^a + \lambda_u^\ell\right) \mathrm{d}u\right) \lambda_s^\ell \\
&\quad \times \int_s^t \exp\left(-\int_s^v \left(\lambda_u^y(\ell_1; s) + \pi_u(a_1)(\ell_1; s)\lambda_u^a(\ell_1; s)\right) \mathrm{d}u\right) \\
&\quad \times \left(\lambda_v^y(\ell_1; s) + \pi_v(a_1)(\ell_1; s)\lambda_v^a(\ell_1; s)\left(1 - \exp\left(-\int_v^t \lambda_u^y(\ell_1; s) \,\mathrm{d}u\right)\right)\right) \mathrm{d}v \,\mathrm{d}s
\end{aligned}
$$

Can we conclude this result does not depend on $\pi_t(a_1)$? Now, note that

$$
\begin{aligned}
&\int_0^t \exp\left(-\int_0^s \left(\lambda_u^y + \pi_u(a_1)\lambda_u^a + \lambda_u^\ell\right) \mathrm{d}u\right) \lambda_s^y \,\mathrm{d}s \\
&= \left[-\exp\left(-\int_0^s \left(\lambda_u^y + \pi_u(a_1)\lambda_u^a + \lambda_u^\ell\right) \mathrm{d}u\right)\right]_0^t - \int_0^t \exp\left(-\int_0^s \left(\lambda_u^y + \pi_u(a_1)\lambda_u^a + \lambda_u^\ell\right) \mathrm{d}u\right) \left(\pi_s(a_1)\lambda_s^a + \lambda_s^\ell\right) \mathrm{d}s \\
&= 1 - \exp\left(-\int_0^t \left(\lambda_u^y + \pi_u(a_1)\lambda_u^a + \lambda_u^\ell\right) \mathrm{d}u\right) - \int_0^t \exp\left(-\int_0^s \left(\lambda_u^y + \pi_u(a_1)\lambda_u^a + \lambda_u^\ell\right) \mathrm{d}u\right) \left(\pi_s(a_1)\lambda_s^a + \lambda_s^\ell\right) \mathrm{d}s
\end{aligned}
$$

Conclude that the previous is equal to

$$
\begin{aligned}
&1-\exp\left(-\int_0^t \left(\lambda^y_u+\pi_u(a_1)\lambda^a_u+\lambda^\ell_u\right)\mathrm{d}u\right)\\
&-\int_0^t \exp\left(-\int_0^s \left(\lambda^y_u+\pi_u(a_1)\lambda^a_u+\lambda^\ell_u\right)\mathrm{d}u\right)\pi_s(a_1)\lambda^a_s\exp\left(-\int_s^t \left(\lambda^y_u+\lambda^\ell_u\right)\mathrm{d}u\right)\mathrm{d}s\\
&-\int_0^t \exp\left(-\int_0^s \left(\lambda^y_u+\pi_u(a_1)\lambda^a_u+\lambda^\ell_u\right)\mathrm{d}u\right)\lambda^\ell_s\exp\left(-\int_s^t (\lambda^y_u(\ell_1;s)+\pi_u(a_1)(\ell_1;s)\lambda^a_u(\ell_1;s))\right)\mathrm{d}s\\
&-\int_0^t \exp\left(-\int_0^s \left(\lambda^y_u+\pi_u(a_1)\lambda^a_u+\lambda^\ell_u\right)\mathrm{d}u\right)\pi_s(a_1)\lambda^a_s\\
&\quad\times\int_s^t \exp\left(-\int_s^v \left(\lambda^y_u+\lambda^\ell_u\right)\mathrm{d}u\right)\lambda^\ell_v\exp\left(-\int_v^t \lambda^y_u(\ell_2;v)\,\mathrm{d}u\right)\mathrm{d}v\,\mathrm{d}s\\
&-\int_0^t \exp\left(-\int_0^s \left(\lambda^y_u+\pi_u(a_1)\lambda^a_u+\lambda^\ell_u\right)\mathrm{d}u\right)\lambda^\ell_s\\
&\quad\times\int_s^t \exp\left(-\int_s^v (\lambda^y_u(\ell_1;s)+\pi_u(a_1)(\ell_1;s)\lambda^a_u(\ell_1;s))\,\mathrm{d}u\right)\\
&\quad\times\pi_v(a_1)(\ell_1;s)\lambda^a_v(\ell_1;s)\exp\left(-\int_v^t \lambda^y_u(\ell_1;s)\,\mathrm{d}u\right)\mathrm{d}v\,\mathrm{d}s\\
&=1-\exp\left(-\int_0^t \left(\lambda^y_u+\pi_u(a_1)\lambda^a_u+\lambda^\ell_u\right)\mathrm{d}u\right)\\
&-\int_0^t \exp\left(-\int_0^s \left(\lambda^y_u+\pi_u(a_1)\lambda^a_u+\lambda^\ell_u\right)\mathrm{d}u\right)\pi_s(a_1)\lambda^a_s\exp\left(-\int_s^t \left(\lambda^y_u+\lambda^\ell_u\right)\mathrm{d}u\right)\mathrm{d}s\\
&-\int_0^t \exp\left(-\int_0^s \left(\lambda^y_u+\pi_u(a_1)\lambda^a_u+\lambda^\ell_u\right)\mathrm{d}u\right)\lambda^\ell_s\exp\left(-\int_s^t (\lambda^y_u(\ell_1;s)+\pi_u(a_1)(\ell_1;s)\lambda^a_u(\ell_1;s))\right)\mathrm{d}s\\
&-\int_0^t \exp\left(-\int_0^s \left(\lambda^y_u+\pi_u(a_1)\lambda^a_u+\lambda^\ell_u\right)\mathrm{d}u\right)\pi_s(a_1)\lambda^a_s\\
&\quad\times\int_s^t \exp\left(-\int_s^v \left(\lambda^y_u+\lambda^\ell_u\right)\mathrm{d}u\right)\lambda^\ell_v\exp\left(-\int_v^t \lambda^y_u(\ell_2;v)\,\mathrm{d}u\right)\mathrm{d}v\,\mathrm{d}s\\
&-\int_0^t \exp\left(-\int_0^s \left(\lambda^y_u+\pi_u(a_1)\lambda^a_u+\lambda^\ell_u\right)\mathrm{d}u\right)\lambda^\ell_s\\
&\quad\times\exp\left(-\int_s^t \lambda^y_u(\ell_1;s)\,\mathrm{d}u\right)\left(1-\exp\left(-\int_s^t (\pi_u(a_1)(\ell_1;s)\lambda^a_u(\ell_1;s))\,\mathrm{d}u\right)\right)\mathrm{d}s
\end{aligned}
$$

$$
\begin{aligned}
&1 - \exp\left(- \int_0^t \left(\lambda_u^y + \pi_u(a_1)\lambda_u^a + \lambda_u^\ell\right) \mathrm{d}u \right) \\
&- \int_0^t \exp\left(- \int_0^s \left(\lambda_u^y + \pi_u(a_1)\lambda_u^a + \lambda_u^\ell\right) \mathrm{d}u \right) \pi_s(a_1)\lambda_s^a \exp\left(- \int_s^t \left(\lambda_u^y + \lambda_u^\ell\right) \mathrm{d}u \right) \mathrm{d}s \\
&- \int_0^t \exp\left(- \int_0^s \left(\lambda_u^y + \pi_u(a_1)\lambda_u^a + \lambda_u^\ell\right) \mathrm{d}u \right) \pi_s(a_1)\lambda_s^a \\
&\quad \times \int_s^t \exp\left(- \int_s^v \left(\lambda_u^y + \lambda_u^\ell\right) \mathrm{d}u \right) \lambda_v^\ell \exp\left(- \int_v^t \lambda_u^y(\ell_2; v) \, \mathrm{d}u \right) \mathrm{d}v \, \mathrm{d}s \\
&- \int_0^t \exp\left(- \int_0^s \left(\lambda_u^y + \pi_u(a_1)\lambda_u^a + \lambda_u^\ell\right) \mathrm{d}u \right) \lambda_s^\ell \exp\left(- \int_s^t \lambda_u^y(\ell_1; s) \, \mathrm{d}u \right) \mathrm{d}s \\
&= 1 - \exp\left(- \int_0^t \left(\lambda_u^y + \lambda_u^\ell\right) \mathrm{d}u \right) \\
&- \int_0^t \exp\left(- \int_0^s \left(\lambda_u^y + \pi_u(a_1)\lambda_u^a + \lambda_u^\ell\right) \mathrm{d}u \right) \pi_s(a_1)\lambda_s^a \\
&\quad \times \int_s^t \exp\left(- \int_s^v \left(\lambda_u^y + \lambda_u^\ell\right) \mathrm{d}u \right) \lambda_v^\ell \exp\left(- \int_v^t \lambda_u^y(\ell_2; v) \, \mathrm{d}u \right) \mathrm{d}v \, \mathrm{d}s \\
&- \int_0^t \exp\left(- \int_0^s \left(\lambda_u^y + \pi_u(a_1)\lambda_u^a + \lambda_u^\ell\right) \mathrm{d}u \right) \lambda_s^\ell \exp\left(- \int_s^t \lambda_u^y(\ell_1; s) \, \mathrm{d}u \right) \mathrm{d}s
\end{aligned}
$$

Partiel integration. Let there be constant intensities, i.e.,

$$
\begin{gathered}
\lambda_u^y = \lambda^y, \\
\lambda_u^y(\ell_i; s) = \lambda^{y,*}(\ell), \\
\pi_u(a_1) = \pi, \\
\lambda_u(a_1) = \lambda^a, \\
\lambda_u^\ell = \lambda^\ell.
\end{gathered}
$$

This simplifies further to:

$$
\begin{aligned}
&1-\exp\bigl(-\bigl(\lambda^y+\lambda^\ell\bigr)t\bigr) \\
&-\int_0^t \exp\bigl(-\bigl(\lambda^y+\pi\lambda^a+\lambda^\ell\bigr)s\bigr)\pi\lambda^a \\
&\quad\times\int_s^t \exp\bigl(-\bigl(\lambda^y+\lambda^\ell\bigr)(v-s)\bigr)\lambda^\ell\exp(-\lambda^{y,*}(\ell)(t-v))\,\mathrm{d}v\,\mathrm{d}s \\
&-\int_0^t \exp\bigl(-\bigl(\lambda^y+\pi\lambda^a+\lambda^\ell\bigr)s\bigr)\lambda^\ell\exp(-\lambda^{y,*}(\ell)(t-s))\,\mathrm{d}s \\
&=1-\exp\bigl(-\bigl(\lambda^y+\lambda^\ell\bigr)t\bigr) \\
&-\int_0^t \exp\bigl(-\bigl(\lambda^y+\pi\lambda^a+\lambda^\ell\bigr)s\bigr)\pi\lambda^a \\
&\quad\times\exp\bigl(\bigl(\lambda^y+\lambda^\ell\bigr)s\bigr)\exp(-\lambda^{y,*}(\ell)(t))\lambda^\ell\int_s^t \exp\bigl(-\bigl(\lambda^y-\lambda^{y,*}(\ell)+\lambda^\ell\bigr)v\bigr)\,\mathrm{d}v\,\mathrm{d}s \\
&-\exp(-\lambda^{y,*}(\ell)t)\lambda^\ell\left(1-\frac{1}{(\lambda^y-\lambda^{y,*}+\pi\lambda^a+\lambda^\ell)}\exp\bigl(-\bigl(\lambda^y-\lambda^{y,*}+\pi\lambda^a+\lambda^\ell\bigr)\bigr)\right) \\
&=1-\exp\bigl(-\bigl(\lambda^y+\lambda^\ell\bigr)t\bigr) \\
&-\int_0^t \exp\bigl(-\bigl(\lambda^y+\pi\lambda^a+\lambda^\ell\bigr)s\bigr)\pi\lambda^a \\
&\quad\times\exp\bigl(\bigl(\lambda^y+\lambda^\ell\bigr)s\bigr)\exp(-\lambda^{y,*}(\ell)(t))\lambda^\ell\frac{1}{\lambda^y-\lambda^{y,*}(\ell)+\lambda^\ell}\bigl(\exp\bigl(-\bigl(\lambda^y-\lambda^{y,*}(\ell)+\lambda^\ell\bigr)s\bigr)-\exp\bigl(-\bigl(\lambda^y-\lambda^{y,*}(\ell)+\lambda^\ell\bigr)t\bigr)\bigr)\,\mathrm{d}s \\
&-\exp(-\lambda^{y,*}(\ell)t)\lambda^\ell\frac{1}{(\lambda^y-\lambda^{y,*}+\pi\lambda^a+\lambda^\ell)}\bigl(1-\exp\bigl(-\bigl(\lambda^y-\lambda^{y,*}+\pi\lambda^a+\lambda^\ell\bigr)\bigr)\bigr)
\end{aligned}
$$

Note that

$$
\begin{aligned}
&\int_0^t \exp\bigl(-\bigl(\lambda^y+\pi\lambda^a+\lambda^\ell\bigr)s\bigr)\pi\lambda^a \\
&\quad\times\exp\bigl(\bigl(\lambda^y+\lambda^\ell\bigr)s\bigr)\exp(-\lambda^{y,*}(\ell)(t))\lambda^\ell\frac{1}{\lambda^y-\lambda^{y,*}(\ell)+\lambda^\ell}\bigl(\exp\bigl(-\bigl(\lambda^y-\lambda^{y,*}(\ell)+\lambda^\ell\bigr)s\bigr)-\exp\bigl(-\bigl(\lambda^y-\lambda^{y,*}(\ell)+\lambda^\ell\bigr)t\bigr)\bigr)\,\mathrm{d}s \\
&=\exp(-\lambda^{y,*}(\ell)(t))\lambda^\ell\frac{1}{\lambda^y-\lambda^{y,*}(\ell)+\lambda^\ell}\int_0^t \exp(-\pi\lambda^a s)\pi\lambda^a \\
&\quad\times\bigl(\exp\bigl(-\bigl(\lambda^y-\lambda^{y,*}(\ell)+\lambda^\ell\bigr)s\bigr)-\exp\bigl(-\bigl(\lambda^y-\lambda^{y,*}(\ell)+\lambda^\ell\bigr)t\bigr)\bigr)\,\mathrm{d}s \\
&=\pi\lambda^a\exp(-\lambda^{y,*}(\ell)(t))\lambda^\ell\frac{1}{\lambda^y-\lambda^{y,*}(\ell)+\lambda^\ell}\int_0^t \exp\bigl(-\bigl(\lambda^y-\lambda^{y,*}(\ell)+\pi\lambda^a+\lambda^\ell\bigr)s\bigr) \\
&-\pi\lambda^a\exp\bigl(-\bigl(\lambda^y+\lambda^\ell\bigr)(t)\bigr)\lambda^\ell\frac{1}{\lambda^y-\lambda^{y,*}(\ell)+\lambda^\ell}\int_0^t \exp(-\pi\lambda^a s) \\
&=\pi\lambda^a\exp(-\lambda^{y,*}(\ell)(t))\lambda^\ell\frac{1}{\lambda^y-\lambda^{y,*}(\ell)+\lambda^\ell}\frac{1}{(\lambda^y-\lambda^{y,*}(\ell)+\pi\lambda^a+\lambda^\ell)}(1-\exp\bigl(-\bigl(\lambda^y-\lambda^{y,*}(\ell)+\pi\lambda^a+\lambda^\ell\bigr)t\bigr) \\
&-\pi\lambda^a\exp\bigl(-\bigl(\lambda^y+\lambda^\ell\bigr)(t)\bigr)\lambda^\ell\frac{1}{\lambda^y-\lambda^{y,*}(\ell)+\lambda^\ell}\frac{1}{\pi\lambda^a}(1-\exp(-\pi\lambda^a t))
\end{aligned}
$$

Collecting the terms

$$
\begin{aligned}
&1-\exp\bigl(-\bigl(\lambda^{y}+\lambda^{\ell}\bigr)t\bigr)\\
&-\int_{0}^{t}\exp\bigl(-\bigl(\lambda^{y}+\pi\lambda^{a}+\lambda^{\ell}\bigr)s\bigr)\pi\lambda^{a}\\
&\quad\times\int_{s}^{t}\exp\bigl(-\bigl(\lambda^{y}+\lambda^{\ell}\bigr)(v-s)\bigr)\lambda^{\ell}\exp(-\lambda^{y,*}(\ell)(t-v))\,\mathrm{d}v\,\mathrm{d}s\\
&-\int_{0}^{t}\exp\bigl(-\bigl(\lambda^{y}+\pi\lambda^{a}+\lambda^{\ell}\bigr)s\bigr)\lambda^{\ell}\exp(-\lambda^{y,*}(\ell)(t-s))\,\mathrm{d}s\\
&=1-\exp\bigl(-\bigl(\lambda^{y}+\lambda^{\ell}\bigr)t\bigr)\\
&\quad-\pi\lambda^{a}\exp(-\lambda^{y,*}(\ell)(t))\lambda^{\ell}\frac{1}{\lambda^{y}-\lambda^{y,*}(\ell)+\lambda^{\ell}}\frac{1}{(\lambda^{y}-\lambda^{y,*}(\ell)+\pi\lambda^{a}+\lambda^{\ell})}(1-\exp\bigl(-\bigl(\lambda^{y}-\lambda^{y,*}(\ell)+\pi\lambda^{a}+\lambda^{\ell}\bigr)t\bigr)\\
&+\exp\bigl(-\bigl(\lambda^{y}+\lambda^{\ell}\bigr)(t)\bigr)\lambda^{\ell}\frac{1}{\lambda^{y}-\lambda^{y,*}(\ell)+\lambda^{\ell}}(1-\exp(-\pi\lambda^{a}t))\\
&-\exp(-\lambda^{y,*}(\ell)t)\lambda^{\ell}\frac{1}{(\lambda^{y}-\lambda^{y,*}+\pi\lambda^{a}+\lambda^{\ell})}\bigl(1-\exp\bigl(-\bigl(\lambda^{y}-\lambda^{y,*}+\pi\lambda^{a}+\lambda^{\ell}\bigr)\bigr)\bigr)\\
&=1-\exp\bigl(-\bigl(\lambda^{y}+\lambda^{\ell}\bigr)t\bigr)\left(1-\lambda^{\ell}\frac{1}{\lambda^{y}-\lambda^{y,*}(\ell)+\lambda^{\ell}}\right)\\
&-\exp\bigl(-\bigl(\lambda^{y}+\pi\lambda^{a}+\lambda^{\ell}\bigr)(t)\bigr)\lambda^{\ell}\left(\frac{1}{\lambda^{y}-\lambda^{y,*}(\ell)+\lambda^{\ell}}-\frac{1}{(\lambda^{y}-\lambda^{y,*}+\pi\lambda^{a}+\lambda^{\ell})}\right.\\
&\quad\left.-\pi\lambda^{a}\frac{1}{\lambda^{y}-\lambda^{y,*}(\ell)+\lambda^{\ell}}\frac{1}{(\lambda^{y}-\lambda^{y,*}(\ell)+\pi\lambda^{a}+\lambda^{\ell})}\right)\\
&-\exp(-\lambda^{y,*}(\ell)t)\lambda^{\ell}\frac{1}{(\lambda^{y}-\lambda^{y,*}+\pi\lambda^{a}+\lambda^{\ell})}\left(1+\pi\lambda^{a}\frac{1}{\lambda^{y}-\lambda^{y,*}(\ell)+\lambda^{\ell}}\right)\\
&=1-\exp\bigl(-\bigl(\lambda^{y}+\lambda^{\ell}\bigr)t\bigr)\left(1-\lambda^{\ell}\frac{1}{\lambda^{y}-\lambda^{y,*}(\ell)+\lambda^{\ell}}\right)\\
&-\exp(-\lambda^{y,*}(\ell)t)\lambda^{\ell}\frac{1}{(\lambda^{y}-\lambda^{y,*}+\lambda^{\ell})}
\end{aligned}
$$

which is constant in π. If $\lambda_u^y(\ell_1;s)=\lambda_u^y(\ell_2;v)=\lambda_u^y$, then this reduces to

$$
\begin{aligned}
&1-\exp\left(-\int_0^t \left(\lambda_u^y+\lambda_u^\ell\right)\mathrm{d}u\right)\\
&+\int_0^t \exp\left(-\int_0^s \left(\lambda_u^y+\pi_u(a_1)\lambda_u^a+\lambda_u^\ell\right)\mathrm{d}u\right)\pi_s(a_1)\lambda_s^a\left(\exp\left(-\int_s^t \left(\lambda_u^y+\lambda_u^\ell\right)\mathrm{d}u\right)-\exp\left(-\int_s^t \left(\lambda_u^y\right)\mathrm{d}u\right)\right)\mathrm{d}s\\
&-\int_0^t \exp\left(-\int_0^s \left(\lambda_u^y+\pi_u(a_1)\lambda_u^a+\lambda_u^\ell\right)\mathrm{d}u\right)\lambda_s^\ell \exp\left(-\int_s^t \lambda_u^y\,\mathrm{d}u\right)\mathrm{d}s\\
&=1-\exp\left(-\int_0^t \left(\lambda_u^y+\lambda_u^\ell\right)\mathrm{d}u\right)\\
&+\exp\left(-\int_0^t \lambda_u^y\,\mathrm{d}u\right)\int_0^t \exp\left(-\int_0^s \left(\pi_u(a_1)\lambda_u^a+\lambda_u^\ell\right)\mathrm{d}u\right)\pi_s(a_1)\lambda_s^a\left(\exp\left(-\int_s^t \left(\lambda_u^\ell\right)\mathrm{d}u\right)-1\right)\mathrm{d}s\\
&-\exp\left(-\int_0^t \lambda_u^y\,\mathrm{d}u\right)\int_0^t \exp\left(-\int_0^s \left(\pi_u(a_1)\lambda_u^a+\lambda_u^\ell\right)\mathrm{d}u\right)\lambda_s^\ell\,\mathrm{d}s\\
&=1-\exp\left(-\int_0^t \left(\lambda_u^y+\lambda_u^\ell\right)\mathrm{d}u\right)\\
&+\exp\left(-\int_0^t \lambda_u^y\,\mathrm{d}u\right)\exp\left(-\int_0^t \lambda_u^\ell\,\mathrm{d}u\right)\int_0^t \exp\left(-\int_0^s \left(\pi_u(a_1)\lambda_u^a\right)\mathrm{d}u\right)\pi_s(a_1)\lambda_s^a\,\mathrm{d}s\\
&-\exp\left(-\int_0^t \lambda_u^y\,\mathrm{d}u\right)\int_0^t \exp\left(-\int_0^s \left(\pi_u(a_1)\lambda_u^a+\lambda_u^\ell\right)\mathrm{d}u\right)\left(\lambda_s^\ell+\pi_s(a_1)\lambda_s^a\right)\mathrm{d}s\\
&=1-\exp\left(-\int_0^t \left(\lambda_u^y+\lambda_u^\ell\right)\mathrm{d}u\right)\\
&+\exp\left(-\int_0^t \lambda_u^y\,\mathrm{d}u\right)\exp\left(-\int_0^t \lambda_u^\ell\,\mathrm{d}u\right)\left(1-\exp\left(-\int_0^t \left(\pi_u(a_1)\lambda_u^a\right)\mathrm{d}u\right)\right)\\
&-\exp\left(-\int_0^t \lambda_u^y\,\mathrm{d}u\right)(1-\exp\left(-\int_0^t \left(\pi_u(a_1)\lambda_u^a+\lambda_u^\ell\right)\mathrm{d}u\right)\\
&=1-\exp\left(-\int_0^t \left(\lambda_u^y+\lambda_u^\ell\right)\mathrm{d}u\right)\\
&+\exp\left(-\int_0^t \lambda_u^y\,\mathrm{d}u\right)\exp\left(-\int_0^t \lambda_u^\ell\,\mathrm{d}u\right)\\
&-\exp\left(-\int_0^t \lambda_u^y\,\mathrm{d}u\right)
\end{aligned}
$$

If we could find an example where the above is non-zero, then we would by the construction in "existence of counterfactuals" have found an example where Påls exchangeability does not hold and mine does.

2) More general exchangeability conditions

We now consider more general exchangeability conditions.

Theorem 2.1: Let $Q_\kappa = \mathcal{E}(\kappa)_T \cdot P$ where κ is a local P-$\mathcal{F}_t$ martingale with $\Delta\kappa_t \geq -1$. If

1. Consistency holds as in Theorem 1.1.
2. $\mathcal{E}(\kappa)_t \mathcal{E}(-\mathbb{N}^a)_t = \mathcal{E}(\kappa)_t$ for all $t \in [0, T]$ P-a.s.
3. Q_κ is a uniformly integrable P-$\mathcal{F}_t$-martingale and P-$\mathcal{H}_t$-martingale, where $\mathcal{H}_t = \mathcal{F}_t \vee \sigma\left(\tilde{Y}\right)$.

Then,

$$\mathbb{E}_P\left[\tilde{Y}_t\right] = \mathbb{E}_P[Y_t \mathcal{E}(\kappa)_t]$$

Proof: The proof is the same as in Theorem 1.1 – mutatis mutandis. □

We provide an equivalent characterization of condition 2 in the above theorem which gives direct interpretability of that condition in the sense that it should induce a probability measure Q_κ under which the time to deviation from the treatment regime is infinite almost surely.

Lemma 1: $Q_\kappa\left(\tau^{g^*} = \infty\right) = 1$ if and only if $\mathcal{E}(\kappa)_t \mathcal{E}(-\mathbb{N}^a)_t = \mathcal{E}(\kappa)_t$ for all $t \in [0, T]$ P-a.s.

Proof: "If" part:

$$\begin{aligned} Q_\kappa\left(\tau^{g^*} = \infty\right) &= \mathbb{E}_P\left[\mathcal{E}(\kappa)_T \mathbb{1}\left\{\tau^{g^*} = \infty\right\}\right] \\ &= \mathbb{E}_P\left[\lim_{t \to \infty} \mathcal{E}(\kappa)_t \mathbb{1}\left\{\tau^{g^*} > t\right\}\right] \\ &= \mathbb{E}_P\left[\lim_{t \to \infty} \mathcal{E}(\kappa)_t\right] \\ &= \mathbb{E}_P[\mathcal{E}(\kappa)_T] = 1. \end{aligned}$$

"Only if" part: Suppose that $Q_\kappa\left(\tau^{g^*} = \infty\right) = 1$. Then for every $t \in [0, T]$,

$$\begin{aligned} \mathbb{E}_P\left[\mathcal{E}(\kappa)_t \mathbb{1}\left\{\tau^{g^*} > t\right\}\right] &= \mathbb{E}_P\left[\mathbb{E}_P[\mathcal{E}(\kappa)_T \mid \mathcal{F}_t] \mathbb{1}\left\{\tau^{g^*} > t\right\}\right] \\ &= \mathbb{E}_P\left[\mathcal{E}(\kappa)_T \mathbb{1}\left\{\tau^{g^*} > t\right\}\right] \\ &= Q\left(\tau^{g^*} > t\right) \\ &= 1. \end{aligned}$$

On the other hand, $\mathbb{E}_P[\mathcal{E}(\kappa)_t] = \mathbb{E}_P[\mathbb{E}_P[\mathcal{E}(\kappa)_T \mid \mathcal{F}_t]] = \mathbb{E}_P[\mathcal{E}(\kappa)_T] = 1$. Conclude that

$$\mathbb{E}_P\left[\mathcal{E}(\kappa)_t\left(1 - \mathbb{1}\left\{\tau^{g^*} > t\right\}\right)\right] = 0.$$

The integrand on the left hand side is non-negative and so it must be zero P-a.s. □

Theorem 2.2: Let κ_t be a (local) martingale with $\Delta\kappa_t \geq -1$ and $\Delta\kappa_t^* > -1$ if $t < \tau^{g^*}$. Then,

$$W_t^* := \mathcal{E}(\kappa)_t = \mathcal{E}(\kappa)_t \mathcal{E}(-\mathbb{N}^a)_t \quad P - \text{a.s.}$$

if and only if

$$\mathbb{1}\left\{\tau^{g^*} < \infty\right\} \kappa_{\tau^{g^*}} = -\mathbb{1}\left\{\tau^{g^*} < \infty\right\} \quad P - \text{a.s.}$$

Proof: Using the well-known formula $\mathcal{E}(X)\mathcal{E}(Y) = \mathcal{E}(X + Y + [X, Y])$, we have

$$\mathcal{E}(\kappa) = \mathcal{E}(\kappa - \mathbb{N}^a - [\kappa, \mathbb{N}^a])$$

This holds if and only if

$$1 + \int_0^t W_{t-} \, \mathrm{d}\kappa_s = 1 + \int_0^t W_{t-} \, \mathrm{d}\left(\kappa_s - \mathbb{N}_s^a - [\kappa, \mathbb{N}^a]_s\right)$$

if and only if

$$\int_0^t W_{t-} \Delta\kappa_s \, \mathrm{d}\mathbb{N}_s^a = -\int_0^t W_{t-} \, \mathrm{d}\mathbb{N}_s^a$$

and this is

$$\mathbb{1}\{\tau^{g^*} \leq t\} W_{\tau^{g^*}-} \Delta\kappa_{\tau^{g^*}} = -\mathbb{1}\{\tau^{g^*} \leq t\} W_{\tau^{g^*}-}$$

By assumption, $W_{\tau^{g^*}-} > 0$ (looking at the explicit solution of the SDE) and so the above holds if and only if

$$\mathbb{1}\{\tau^{g^*} \leq t\} \Delta\kappa_{\tau^{g^*}} = -\mathbb{1}\{\tau^{g^*} \leq t\}$$

Taking $t \to \infty$ gives the desired result. On the other hand, if the result holds then,

$$\begin{aligned}
\mathbb{1}\{\tau^{g^*} \leq t\} \Delta\kappa_{\tau^{g^*}} &= \mathbb{1}\{\tau^{g^*} \leq t\} \mathbb{1}\{\tau^{g^*} < \infty\} \Delta\kappa_{\tau^{g^*}} \\
&= \mathbb{1}\{\tau^{g^*} \leq t\} \mathbb{1}\{\tau^{g^*} < \infty\}(-1) = -\mathbb{1}\{\tau^{g^*} \leq t\}
\end{aligned}$$

□

Now, we consider only κ's of the form

$$\kappa(t) = \int_0^t \sum_{x \in \mathcal{A}} \mathbb{1}\{s \leq t\} \tilde{h}(s, x) M^{a,x}(\mathrm{d}s)$$

with $\tilde{h}(s, x)$ P-$\mathcal{F}_t$ predictable with the restriction stated in the above theorem. We make this restriction as any reasonable exchangeability conditions should be placed on the treatment process.

Theorem 2.3: Let $\kappa_t = \int_0^t \sum_{x \in \mathcal{A}} \mathbb{1}\{s \leq t\} \tilde{h}(s, x) M^{a,x}(\mathrm{d}s)$ for some P-$\mathcal{F}_t$-predictable process $\tilde{h}(s, x)$ and that $Q_\kappa\left(\tau^{g^*} = \infty\right) = 1$. Suppose that $\Delta \kappa_t \geq -1$ and $\Delta \kappa_t^* > -1$ if $t < \tau^{g^*}$ and that $\mathcal{E}(\kappa)_t$ is a uniformly integrable P-$\mathcal{F}_t$-martingale. Suppose that $\mathcal{A} = \{a_0, a_1\}$ and that $\pi_t^*(a_1) = 1$. Then any κ with $\tilde{h}(s, a_1)$ a P-$\mathcal{F}_t$-predictable process and

$$\tilde{h}(s, a_0) = \mathbb{1}\{s \leq \tau^{g^*}\} \frac{-1 + \mathbb{1}\{\Lambda^a(\{s\}) > 0\} h(s, a_1) \pi_s(a_1) \Delta \Lambda_s^a}{1 - \mathbb{1}\{\Lambda^a(\{s\}) > 0\}(1 - \pi_s(a_1)) \Delta \Lambda_s^a}$$

will satisfy the condition 2. of Theorem 2.1. Moreover, this solution is unique in the sense that whenever $\int_0^t \tilde{h}(s, a_0) M^{a,a_0}(\mathrm{d}s)$ is of finite variation, this is equal to

$$\int_0^t \mathbb{1}\{s \leq \tau^{g^*}\} \frac{-1 + \mathbb{1}\{\Lambda^a(\{s\}) > 0\} h(s, a_1) \pi_s(a_1) \Delta \Lambda_s^a}{1 - \mathbb{1}\{\Lambda^a(\{s\}) > 0\}(1 - \pi_s(a_1)) \Delta \Lambda_s^a} M^{a,a_0}(\mathrm{d}s).$$

Conclude that, under regularity conditions,

$$\begin{aligned}\mathcal{E}(\kappa)_t = 1 &+ \int_0^t \mathcal{E}(\kappa)_{s-} \mathbb{1}\{s \leq \tau^{g^*}\} \frac{-1 + \mathbb{1}\{\Lambda^a(\{s\}) > 0\} h(s, a_1) \pi_s(a_1) \Delta \Lambda_s^a}{1 - \mathbb{1}\{\Lambda^a(\{s\}) > 0\}(1 - \pi_s(a_1)) \Delta \Lambda_s^a} M^{a,a_0}(\mathrm{d}s) \\ &+ \int_0^t \mathcal{E}(\kappa)_{s-} \tilde{h}(s, a_1) M^{a,a_1}(\mathrm{d}s)\end{aligned}$$

are all the solutions to condition 2. of Theorem 2.1, where $\tilde{h}(s, a_1)$ is any P-$\mathcal{F}_t$-predictable process, ensuring that $\mathcal{E}(\kappa)_t$ is a uniformly integrable P-$\mathcal{F}_t$-martingale.

Proof:

The above theorem gives that we must have

$$\Delta \kappa_{\tau^{g^*}} = \sum_{x \in \mathcal{A}} \tilde{h}\left(\tau^{g^*}, x\right) \Delta M_{\tau^{g^*}}^{a,x} = -1$$

on the event that $\tau^{g^*} < \infty$. Suppose that $\mathcal{A} = \{a_0, a_1\}$ and that $\pi_t^*(a_1) = 1$. In this case, we can write the equation above as

$$h\left(\tau^{g^*}, a_1\right)\left(0 - \pi_{\tau^{g^*}}(a_1) \Delta \Lambda_{\tau^{g^*}}^a\right) + h\left(\tau^{g^*}, a_0\right)\left(1 - (1 - \pi_{\tau^{g^*}}(a_1)) \Delta \Lambda_{\tau^{g^*}}^a\right) = -1$$

or

$$\left(h\left(\tau^{g^*}, a_0\right) - h\left(\tau^{g^*}, a_1\right)\right) \pi_{\tau^{g^*}}(a_1) \Delta \Lambda_{\tau^{g^*}}^a + h\left(\tau^{g^*}, a_0\right)\left(1 - \Delta \Lambda_{\tau^{g^*}}^a\right) = -1$$

We consider various cases:
- Absolutely continuous case: $\Delta \Lambda^a \equiv 0$.
- $\bar{N}^a$ is $\mathcal{F}_t$-predictable.
- Jump times for $\bar{N}^a$ are discrete.
- General case.

2.a) Absolutely continuous case

In this case, conclude that $h\left(\tau^{g^*}, a_0\right) = -1$. However, nothing else can be said about $h\left(\tau^{g^*}, a_1\right)$ as the equation does not place any other restrictions than it being predictable. We can, however, conclude that $\int_0^{t \wedge \tau^{g^*}} h(s, a_0) M^{a,a_0}(\mathrm{d}s) = \int_0^{t \wedge \tau^{g^*}} (-1) M^{a,a_0}(\mathrm{d}s) = -\mathbb{N}^a(t) + \mathbb{L}^a(t)$ whenever that integral happens to be of finite variation. To see this, note that

$$\begin{aligned}\int_0^{t\wedge\tau^{g^*}} h(s,a_0)M^{a,a_0}(\mathrm{d}s) &= \int_0^{t\wedge\tau^{g^*}} h(s,a_0)N^{a,a_0}(\mathrm{d}s) - \int_0^{t\wedge\tau^{g^*}} h(s,a_0)\Lambda^{a,a_0}(\mathrm{d}s) \\ &= \int_0^{t\wedge\tau^{g^*}} (-1)N^{a,a_0}(\mathrm{d}s) - \int_0^{t\wedge\tau^{g^*}} h(s,a_0)\Lambda^{a,a_0}(\mathrm{d}s) \\ &= \int_0^{t\wedge\tau^{g^*}} (-1)M^{a,a_0}(\mathrm{d}s) - \int_0^{t\wedge\tau^{g^*}} (h(s,a_0)+1)\Lambda^{a,a_0}(\mathrm{d}s),\end{aligned}$$

meaning that $\int_0^{t\wedge\tau^{g^*}}(h(s,a_0)+1)\Lambda^{a,a_0}(\mathrm{d}s)$ is of finite variation, a local martingale, predicable and hence constant (and thus zero) by Theorem 15, p. 115 of Protter (2005).

2.b) $\bar{N}^a$ is $\mathcal{F}_t$-predictable

Im this case, $\Delta\Lambda_t^a = \Delta\bar{N}_t^a$ which is 1 at $t = \tau^{g^*}$. Therefore,

$$\left(h\left(\tau^{g^*},a_0\right) - h\left(\tau^{g^*},a_1\right)\right)\pi_{\tau^{g^*}}(a_1) = -1$$

or

$$h\left(\tau^{g^*},a_0\right) = h\left(\tau^{g^*},a_1\right) - \frac{1}{\pi_{\tau^{g^*}}(a_1)}$$

Thus, we have

$$\begin{aligned}K_t^h &= \int_0^{t\wedge\tau^{g^*}} h(s,a_0)M^{a,a_0}(\mathrm{d}s) + \int_0^{t\wedge\tau^{g^*}} h(s,a_1)M^{a,a_1}(\mathrm{d}s) \\ &= \int_0^{t\wedge\tau^{g^*}} h(s,a_0)M^{a,a_0}(\mathrm{d}s) - \int_0^{t\wedge\tau^{g^*}} (h(s,a_1))M^{a,a_0}(\mathrm{d}s) + \int_0^{t\wedge\tau^{g^*}} (h(s,a_1))M^{a}(\mathrm{d}s) \\ &= \int_0^{t\wedge\tau^{g^*}} (h(s,a_0)-h(s,a_1))M^{a,a_0}(\mathrm{d}s) \\ &= \int_0^{t\wedge\tau^{g^*}} \left(-\frac{1}{\pi_s(a_1)}\right)N^{a,a_0}(\mathrm{d}s) - \int_0^{t\wedge\tau^{g^*}} (h(s,a_0)-h(s,a_1))\Lambda^{a,a_0}(\mathrm{d}s) \\ &= \int_0^{t\wedge\tau^{g^*}} \left(-\frac{1}{\pi_s(a_1)}\right)M^{a,a_0}(\mathrm{d}s) - \int_0^{t\wedge\tau^{g^*}} \left((h(s,a_0)-h(s,a_1)) + \frac{1}{\pi_s(a_1)}\right)\Lambda^{a,a_0}(\mathrm{d}s)\end{aligned}$$

Assuming that $\int_0^{t\wedge\tau^{g^*}}(h(s,a_0)-h(s,a_1))M^{a,a_0}(\mathrm{d}s)$ is of finite variation, we have that $\int_0^{t\wedge\tau^{g^*}}((h(s,a_0)-h(s,a_1))M^{a,a_0}(\mathrm{d}s) = \int_0^{t\wedge\tau^{g^*}}\left(-\frac{1}{\pi_s(a_1)}\right)M^{a,a_0}(\mathrm{d}s)$. We conclude that K_t^h if $\int_0^{t\wedge\tau^{g^*}}(h(s,a_0)-h(s,a_1))M^{a,a_0}(\mathrm{d}s)$ is of finite variation does not depend on the choice of h. Therefore, the stochastic exponential $\mathcal{E}\left(K^h\right)_t$ does not depend on the choice of h either, and we may conclude that $\mathcal{E}\left(K^h\right)_t = \mathcal{E}(K)_t$.

2.c) General case

Suppose that $(1-\pi_t(a_1))\Delta\Lambda_t^a < 1$ for all $t > 0$. Otherwise, an argument similar to the one we will give will split into cases.

We have that

$$\begin{aligned}
&\left(h\left(\tau^{g^*},a_0\right)-h\left(\tau^{g^*},a_1\right)\right)\pi_{\tau^{g^*}}(a_1)\Lambda^a\left(\left\{\tau^{g^*}\right\}\right)\mathbb{1}\left\{\Lambda^a\left(\left\{\tau^{g^*}\right\}\right)>0\right\}\\
&\quad+h\left(\tau^{g^*},a_0\right)\left(1-\Lambda^a\left(\left\{\tau^{g^*}\right\}\right)\right)\mathbb{1}\left\{\Lambda^a\left(\left\{\tau^{g^*}\right\}\right)>0\right\}\\
&\quad+h\left(\tau^{g^*},a_0\right)\left(\mathbb{1}\left\{\Lambda^a\left(\left\{\tau^{g^*}\right\}\right)=0\right\}\right)=-1
\end{aligned}$$

By the same argument as in the absolutely continuous case, we have that

$$\begin{aligned}
&\int_0^{t\wedge\tau^{g^*}} h(s,a_0)\mathbb{1}\{\Lambda^a(\{s\})=0\}M^{a,a_0}(\mathrm{d}s)\\
&=-\int_0^{t\wedge\tau^{g^*}} \mathbb{1}\{\Lambda^a(\{s\})=0\}M^{a,a_0}(\mathrm{d}s)\\
&=-\mathbb{N}^a(t)\mathbb{1}\left\{\Lambda^a\left(\left\{\tau^{g^*}\right\}\right)=0\right\}+\mathbb{L}^{a,c}(t),
\end{aligned}$$

where $\mathbb{L}^{a,c}$ is the continuous part of $\mathbb{L}^a$. Next whenever $\Lambda^a\left(\left\{\tau^{g^*}\right\}\right)>0$, we find

$$h\left(\tau^{g^*},a_0\right)=\frac{-1+h\left(\tau^{g^*},a_1\right)\pi_{\tau^{g^*}}(a_1)\Delta\Lambda^a_{\tau^{g^*}}}{1-(1-\pi_{\tau^{g^*}}(a_1))\Delta\Lambda^a_{\tau^{g^*}}}$$

Therefore, it will again be the case that

$$\begin{aligned}
&\int_0^{t\wedge\tau^{g^*}} h(s,a_0)\mathbb{1}\{\Lambda^a(\{s\})>0\}M^{a,a_0}(\mathrm{d}s)\\
&=\int_0^{t\wedge\tau^{g^*}} \frac{-1+h(s,a_1)\pi_s(a_1)\Delta\Lambda^a_s}{1-(1-\pi_s(a_1))\Delta\Lambda^a_s}\mathbb{1}\{\Lambda^a(\{s\})>0\}M^{a,a_0}(\mathrm{d}s)
\end{aligned}$$

Conclude that

$$\begin{aligned}
&\int_0^{t\wedge\tau^{g^*}} h(s,a_0)M^{a,a_0}(\mathrm{d}s)\\
&=\int_0^{t\wedge\tau^{g^*}} \left(\frac{-1+h(s,a_1)\pi_s(a_1)\Delta\Lambda^a_s}{1-(1-\pi_s(a_1))\Delta\Lambda^a_s}\mathbb{1}\{\Lambda^a(\{s\})>0\}-\mathbb{1}\{\Lambda^a(\{s\})=0\}\right)M^{a,a_0}(\mathrm{d}s)\\
&=\int_0^{t\wedge\tau^{g^*}} \frac{-1+\mathbb{1}\{\Lambda^a(\{s\})>0\}h(s,a_1)\pi_s(a_1)\Delta\Lambda^a_s}{1-\mathbb{1}\{\Lambda^a(\{s\})>0\}(1-\pi_s(a_1))\Delta\Lambda^a_s}M^{a,a_0}(\mathrm{d}s),
\end{aligned}$$

and $h(\cdot,a_1)$ freely chosen, predictable satisfying some integrability criteria. Interestingly, this means that the stochastic exponential $\mathcal{E}\left(K^h\right)_t$ will depend on the choice of h in general, but only through $h(s,a_1)$ which can be freely chosen. □

3) Score operator calculations

Theorem 3.1: Let $K^*_t = K_{t \wedge \tau^{g^*}}$.
1. The score operator $S : \mathcal{M}^2(P) \to \mathcal{M}^2(Q)$ is given by

$$\Gamma \mapsto \Gamma - \langle \Gamma, K^* \rangle^P_t$$
$$- \int_0^{t \wedge \tau^{g^*}} \sum_{j=1}^{K} \pi^*_s(\{a_j\}) \frac{\mathrm{d}\langle \Gamma, M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a \rangle^P_t}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} \left(M^{a,a_j}(\mathrm{d}s) - \mathrm{d}\langle M^{a,a_j}, K^* \rangle^P_s\right).$$

2. Let

$D : \mathcal{M}^2(P) \to \mathcal{M}^2(Q)$ be the linear operator defined by

$$\Gamma \mapsto \Gamma - \langle \Gamma, K^* \rangle^P_t.$$

Then its adjoint $D^* : \mathcal{M}^2(Q) \to \mathcal{M}^2(P)$ is given by

$$\Gamma' \mapsto \Gamma'_0 + \int_0^t W_{s-} \,\mathrm{d}(\Gamma' + [\Gamma', K^*])_s$$

The adjoint of the score operator $S^* : \mathcal{M}^2(Q) \to \mathcal{M}^2(P)$ is given by

$$\Gamma' \mapsto D^* \Gamma' - \int_0^t \mathbb{1}\{\tau^{g^*} \leq s\} \sum_j \sum_k \pi^*_s(\{a_j\}) W_{s-} \frac{\mathrm{d}\langle M^{a,a_j}, \Gamma' \rangle^Q_s}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \,\mathrm{d}\left(M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a\right)_s.$$

or $\left(\mathrm{Id} - \sum_j Y^*_j\right)(D^*)$ where $Y_j : \mathcal{M}^2(P) \to \mathcal{M}^2(P)$ is given by

$$Y_j \Gamma = \int_0^{\cdot \wedge \tau^{g^*}} \sum_k \pi^*_s(\{a_j\}) \frac{\mathrm{d}\langle \Gamma, M^{a,a_j} \rangle^P_s - \pi_s(\{a_j\})(P) \,\mathrm{d}\langle \Gamma, M^a \rangle^P_s}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} M^{a,a_j}(\mathrm{d}s)$$

and its adjoint $Y^*_j : \mathcal{M}^2(P) \to \mathcal{M}^2(P)$ is given by

$$Y^*_j \Gamma = \int_0^t \mathbb{1}\{\tau^{g^*} \leq s\} \sum_k \pi^*_s(\{a_j\}) \frac{\mathrm{d}\langle M^{a,a_j}, \Gamma' \rangle^P_s}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \,\mathrm{d}\left(M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a\right)_s$$

3. $\ker(S^*) = \{0\}$ if $\pi_s(\{a_j\})(P) > \eta$ for all $s \in [0, T]$ and j for some $\eta > 0$.
4. $K^* \in \ker(S)$.

1. Let $\bar{K}_t = K^*_t + \mathbb{N}^a_t$

for a given (local) martingale K^*_t with $\Delta K^*_t \geq -1$ and $\Delta K^*_t = -1$ if and only if $t = \tau^{g^*}$ and $\tau^{g^*} < \infty$. Then, it is the case that $\Delta \bar{K}_t = \mathbb{1}\{t \neq \tau^{g^*}\} \Delta K^*_t + \mathbb{1}\{t = \tau^{g^*}\}(0) > -1$ so $\mathcal{E}(\bar{K}) > 0$. First, we see that

$$\begin{aligned}
\mathcal{E}(K^*)_t &= \mathcal{E}(-\mathbb{N}^a)_t \mathcal{E}(K^*)_t \\
&= \mathcal{E}(K^* + \mathbb{N}^a - \mathbb{N}^a)_t \mathcal{E}(-\mathbb{N}^a)_t \\
&= \mathcal{E}(\bar{K})_t \mathcal{E}(-\mathbb{N}^a)_t.
\end{aligned}$$

Let $\bar{W} = \mathcal{E}(\bar{K})$ and $\partial_\varepsilon f(\varepsilon) = \frac{\partial}{\partial \varepsilon} f(\varepsilon)\big|_{\varepsilon=0}$. Then, let $\mathcal{L}$ denote the stochastic logarithm, so that

$$
\begin{aligned}
\frac{1}{\varepsilon}\mathcal{L}\left(\frac{\bar{W}^{\varepsilon}}{\bar{W}^{0}}\right)_t &= \frac{1}{\varepsilon}\mathcal{L}\left(\mathcal{E}(\bar{K}^{\varepsilon})\mathcal{E}\left(-\bar{K}^0 + \sum_{0<s\leq \cdot}\frac{\left(\Delta \bar{K}_s^0\right)^2}{1+\Delta \bar{K}_s^0}\right)\right)_t \\
&= \frac{1}{\varepsilon}\mathcal{L}\left(\mathcal{E}\left(\bar{K}^{\varepsilon} - \bar{K}^0 + \sum_{0<s\leq \cdot}\frac{\left(\Delta \bar{K}_s^0\right)^2}{1+\Delta \bar{K}_s^0} + \left[\bar{K}^{\varepsilon}, -\bar{K}^0 + \sum_{0<s\leq \cdot}\frac{\left(\Delta \bar{K}_s^0\right)^2}{1+\Delta \bar{K}_s^0}\right]\right)\right)_t \\
&= \frac{1}{\varepsilon}\left(\bar{K}_t^{\varepsilon} - \bar{K}_t^0 + \sum_{0<s\leq t}\frac{\left(\Delta \bar{K}_s^0\right)^2}{1+\Delta \bar{K}_s^0} + \left[\bar{K}^{\varepsilon}, -\bar{K}^0 + \sum_{0<s\leq \cdot}\frac{\left(\Delta \bar{K}_s^0\right)^2}{1+\Delta \bar{K}_s^0}\right]_t\right) \\
&= \frac{1}{\varepsilon}\left(\bar{K}_t^{\varepsilon} - \bar{K}_t^0 + \sum_{0<s\leq t}\frac{\left(\Delta \bar{K}_s^0\right)^2}{1+\Delta \bar{K}_s^0} - \left[\bar{K}^{\varepsilon}, \sum_{0<s\leq \cdot}\frac{\Delta \bar{K}_s^0}{1+\Delta \bar{K}_s^0}\right]_t\right) \\
&= \frac{1}{\varepsilon}\left(\bar{K}_t^{\varepsilon} - \bar{K}_t^0 + \sum_{0<s\leq t}\frac{\left(\Delta \bar{K}_s^0\right)^2}{1+\Delta \bar{K}_s^0} - \sum_{0<s\leq t}\Delta \bar{K}_s^{\varepsilon}\frac{\Delta \bar{K}_s^0}{1+\Delta \bar{K}_s^0}\right) \\
&\to \partial_{\varepsilon}\bar{K}_t^{\varepsilon} - \sum_{0<s\leq t}\left(\partial_{\varepsilon}\Delta \bar{K}_s^{\varepsilon}\right)\frac{\Delta \bar{K}_s^0}{1+\Delta \bar{K}_s^0}
\end{aligned}
$$

as $\varepsilon \to 0$, where we use dominated convergence and L'Hopitals rule for the last step. The result is presented in Equation 13.

We will also need to calculate $\partial_{\varepsilon}\left(\pi_s\left(\{a_j\}\right)(P_{\varepsilon})\right)$, which by definition fulfills that

$$
\Lambda^{a,a_j}(\mathrm{d}t)(P_{\varepsilon}) = \pi_t\left(\{a_j\}\right)(P_{\varepsilon})\Lambda^a(\mathrm{d}t)(P_{\varepsilon}),
$$

where $M^a = \sum_{j=1}^K M^{a,a_j}$. Taking the derivative on both sides gives

$$
\langle \Gamma, M^{a,a_j}\rangle_t^P = \left(\partial_{\varepsilon}\left(\pi_t\left(\{a_j\}\right)(P_{\varepsilon})\right)\Lambda^a(\mathrm{d}t)(P) + \pi_t\left(\{a_j\}\right)(P)\,\mathrm{d}\langle \Gamma, M^a\rangle_t^P\right)
$$

so we conclude that

$$
\begin{aligned}
\partial_{\varepsilon}\left(\pi_t\left(\{a_j\}\right)(P_{\varepsilon})\right) &= \frac{\mathrm{d}\langle \Gamma, M^{a,a_j}\rangle_t^P - \pi_t\left(\{a_j\}\right)(P)\,\mathrm{d}\langle \Gamma, M^a\rangle_t^P}{\mathrm{d}\Lambda^a(t)(P)} \\
&= \frac{\mathrm{d}\langle \Gamma, \left(1-\pi_{\cdot}\left(\{a_j\}\right)(P)\right)\bullet M^{a,a_j} - \pi_{\cdot}\left(\{a_j\}\right)(P)\bullet \sum_{i\neq j} M^{a,a_i}\rangle_t^P}{\mathrm{d}\Lambda^a(t)(P)}
\end{aligned}
$$

Here, we have used that $\partial_{\varepsilon}\Lambda_t(P_{\varepsilon}) = \langle \Gamma, M\rangle_t^P$ if $M = N - \Lambda(P)$.

$$
\begin{aligned}
\bar{K}_t &= \int_0^{t\wedge \tau^{g^*}}\sum_{j=1}^K\left(\frac{\pi_s^*\left(\{a_j\}\right)}{\pi_s\left(\{a_j\}\right)} - 1\right)N^{a,a_j}(\mathrm{d}s) + \mathbb{N}_t^a \\
&= \int_0^{t\wedge \tau^{g^*}}\sum_{j=1}^K\left(1-\pi_s^*\left(\{a_j\}\right)\right)(-1)N^{a,a_j}(\mathrm{d}s) + \mathbb{N}_t^a \\
&\quad + \int_0^{t\wedge \tau^{g^*}}\sum_{j=1}^K\pi_s^*\left(\{a_j\}\right)\left(\frac{1}{\pi_s\left(\{a_j\}\right)} - 1\right)N^{a,a_j}(\mathrm{d}s) \\
&= \int_0^{t\wedge \tau^{g^*}}\sum_{j=1}^K\pi_s^*\left(\{a_j\}\right)\left(\frac{1}{\pi_s\left(\{a_j\}\right)} - 1\right)N^{a,a_j}(\mathrm{d}s)
\end{aligned}
\tag{10}
$$

This can also be written as

$$\bar{K}_t = \int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \left(\frac{1}{\pi_s(\{a_j\})} - 1 \right) M^{a,a_j}(\mathrm{d}s) + \mathbb{L}_t^a \tag{11}$$

Calculating the derivative of Equation 11 gives

$$\begin{aligned}
\partial_\varepsilon \bar{K}_t^\varepsilon &= -\int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \frac{1}{\pi_s(\{a_j\})(P)^2} \partial_\varepsilon(\pi_s(\{a_j\})(P_\varepsilon)) M^{a,a_j}(\mathrm{d}s) \\
&\quad - \int_0^{\tau\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \left(\frac{1}{\pi_s(\{a_j\})} - 1 \right) \partial_\varepsilon(\Lambda^{a,a_j}(\mathrm{d}s)(P_\varepsilon)) + \partial_\varepsilon \mathbb{L}_t^a(P_\varepsilon) \\
&= -\int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \frac{1}{\pi_s(\{a_j\})(P)^2} \partial_\varepsilon(\pi_s(\{a_j\})(P_\varepsilon)) M^{a,a_j}(\mathrm{d}s) \\
&\quad - \int_0^{\tau\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \left(\frac{1}{\pi_s(\{a_j\})} - 1 \right) \mathrm{d}\langle \Gamma, M^{a,a_j} \rangle_s^P + \langle \Gamma, \mathbb{L}^a \rangle_t^P \\
&= -\int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \frac{1}{\pi_s(\{a_j\})(P)^2} \partial_\varepsilon(\pi_s(\{a_j\})(P_\varepsilon)) M^{a,a_j}(\mathrm{d}s) - \langle \Gamma, K^* \rangle_t^P,
\end{aligned}$$

again using that $\partial_\varepsilon \Lambda_t(P_\varepsilon) = \langle \Gamma, M \rangle_t^P$ if $M = N - \Lambda(P)$ is a P-martingale. However for Equation 10, we can also get that

$$\partial_\varepsilon \bar{K}_t^\varepsilon = -\int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \frac{1}{\pi_s(\{a_j\})(P)^2} \partial_\varepsilon(\pi_s(\{a_j\})(P_\varepsilon)) N^{a,a_j}(\mathrm{d}s)$$

Also note that

$$\frac{\Delta \bar{K}_s^0}{1 + \Delta \bar{K}_s^0} = \sum_{j=1}^K \pi_s^*(\{a_j\})(1 - \pi_s(\{a_j\})) \Delta N^{a,a_j}(s)$$

by Equation 10. Thus,

$$\begin{aligned}
\sum_{0<s\leq t} (\partial_\varepsilon \Delta \bar{K}_s^\varepsilon) \frac{\Delta \bar{K}_s^0}{1 + \Delta \bar{K}_s^0} &= -\int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \partial_\varepsilon(\pi_s(\{a_j\})(P_\varepsilon)) \frac{1}{(\pi_s(\{a_j\}))^2} (1 - \pi_s(\{a_j\})) N^{a,a_j}(\mathrm{d}s) \\
&= -\int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \frac{1}{\pi_s(\{a_j\})(P)} \partial_\varepsilon(\pi_s(\{a_j\})(P_\varepsilon)) \left(\frac{1}{\pi_s(\{a_j\})} - 1 \right) N^{a,a_j}(\mathrm{d}s) \\
&= -\int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \frac{1}{\pi_s(\{a_j\})(P)} \partial_\varepsilon(\pi_s(\{a_j\})(P_\varepsilon)) \left(\frac{1}{\pi_s(\{a_j\})} - 1 \right) M^{a,a_j}(\mathrm{d}s) \\
&\quad - \int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \frac{1}{\pi_s(\{a_j\})(P)} \partial_\varepsilon(\pi_s(\{a_j\})(P_\varepsilon)) \left(\frac{1}{\pi_s(\{a_j\})} - 1 \right) \Lambda^{a,a_j}(\mathrm{d}s)
\end{aligned}$$

Conclude that, if $\pi_s(\{a_j\})(P) > 0$ for all $s \in [0, T]$ and j,

$$
\begin{aligned}
&\partial_\varepsilon \bar{K}^\varepsilon_t - \sum_{0<s\le t} \left(\partial_\varepsilon \Delta \bar{K}^\varepsilon_s\right) \frac{\Delta \bar{K}^0_s}{1+\Delta \bar{K}^0_s} \\
&= - \int_0^{t\wedge \tau^{g^*}} \sum_{j=1}^K \pi^*_s(\{a_j\}) \frac{1}{\pi_s(\{a_j\})(P)^2} \partial_\varepsilon \left(\pi_s(\{a_j\})(P_\varepsilon)\right) M^{a,a_j}(\mathrm{d}s) - \langle \Gamma, K^* \rangle^P_t \\
&\quad + \int_0^{t\wedge \tau^{g^*}} \sum_{j=1}^K \pi^*_s(\{a_j\}) \frac{1}{\pi_s(\{a_j\})(P)} \partial_\varepsilon \left(\pi_s(\{a_j\})(P_\varepsilon)\right) \left(\frac{1}{\pi_s(\{a_j\})} - 1\right) M^{a,a_j}(\mathrm{d}s) \\
&\quad + \int_0^{t\wedge \tau^{g^*}} \sum_{j=1}^K \pi^*_s(\{a_j\}) \frac{1}{\pi_s(\{a_j\})(P)} \partial_\varepsilon \left(\pi_s(\{a_j\})(P_\varepsilon)\right) \left(\frac{1}{\pi_s(\{a_j\})} - 1\right) \Lambda^{a,a_j}(\mathrm{d}s) \\
&= -\langle \Gamma, K^* \rangle^P_t \\
&\quad - \int_0^{t\wedge \tau^{g^*}} \sum_{j=1}^K \pi^*_s(\{a_j\}) \frac{1}{\pi_s(\{a_j\})(P)} \partial_\varepsilon \left(\pi_s(\{a_j\})(P_\varepsilon)\right) M^{a,a_j}(\mathrm{d}s) \\
&\quad + \int_0^{t\wedge \tau^{g^*}} \sum_{j=1}^K \pi^*_s(\{a_j\}) \frac{1}{\pi_s(\{a_j\})(P)} \partial_\varepsilon \left(\pi_s(\{a_j\})(P_\varepsilon)\right) \left(\frac{1}{\pi_s(\{a_j\})} - 1\right) \Lambda^{a,a_j}(\mathrm{d}s) \\
&= -\langle \Gamma, K^* \rangle^P_t \\
&\quad - \int_0^{t\wedge \tau^{g^*}} \sum_{j=1}^K \pi^*_s(\{a_j\}) \frac{1}{\pi_s(\{a_j\})(P)} \partial_\varepsilon \left(\pi_s(\{a_j\})(P_\varepsilon)\right) \left(M^{a,a_j}(\mathrm{d}s) - \mathrm{d}\langle M^{a,a_j}, K^* \rangle^P_s\right) \\
&= -\langle \Gamma, K^* \rangle^P_t \\
&\quad - \int_0^{t\wedge \tau^{g^*}} \sum_{j=1}^K \pi^*_s(\{a_j\}) \frac{\mathrm{d}\langle \Gamma, M^{a,a_j} \rangle^P_t - \pi_t(\{a_j\})(P)\, \mathrm{d}\langle \Gamma, M^a \rangle^P_t}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} \left(M^{a,a_j}(\mathrm{d}s) - \mathrm{d}\langle M^{a,a_j}, K^* \rangle^P_s\right)
\end{aligned}
\tag{12}
$$

Conclude that the Score operator S is given by

$$
\begin{aligned}
\mathcal{M}^2(P) \ni \Gamma \mapsto \Gamma &- \langle \Gamma, K^* \rangle^P_t \\
&- \int_0^{t\wedge \tau^{g^*}} \sum_{j=1}^K \pi^*_s(\{a_j\}) \frac{\mathrm{d}\langle \Gamma, M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a \rangle^P_t}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} \left(M^{a,a_j}(\mathrm{d}s) - \mathrm{d}\langle M^{a,a_j}, K^* \rangle^P_s\right) \in \mathcal{M}^2(Q)
\end{aligned}
\tag{13}
$$

2. Assume that we have found the adjoint operator of $D : \mathcal{M}^2(P) \to \mathcal{M}^2(Q)$

$$
\mathcal{M}^2(P) \ni \Gamma \mapsto \Gamma - \langle \Gamma, K^* \rangle^P_t \in \mathcal{M}^2(Q)
$$

say $D^* : \mathrm{Range}(D) \subset \mathcal{M}^2(Q) \to \mathcal{M}^2(P)$. Let $H_j : \mathcal{M}^2(P) \to \mathcal{M}^2(Q)$ be given by

$$
H_j \Gamma = \int_0^{t\wedge \tau^{g^*}} \pi^*_s(\{a_j\}) \frac{\mathrm{d}\langle \Gamma, M^{a,a_j} \rangle^P_t - \pi_t(\{a_j\})(P)\, \mathrm{d}\langle \Gamma, M^a \rangle^P_t}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} \left(M^{a,a_j}(\mathrm{d}s) - \mathrm{d}\langle M^{a,a_j}, K^* \rangle^P_s\right)
$$

Then, we have that

$$
\begin{aligned}
\langle H_k\Gamma, \Gamma\rangle_Q &:= \mathbb{E}_Q\left[\langle H_k\Gamma, \Gamma'\rangle^Q\right] \\
&= \mathbb{E}_Q\left[\langle \int_0^{\cdot\wedge\tau^{g^*}} \pi_s^*(\{a_j\}) \frac{\mathrm{d}\langle \Gamma, M^{a,a_j}\rangle_s^P - \pi_s(\{a_j\})(P)\,\mathrm{d}\langle\Gamma, M^a\rangle_s^P}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} DM^{a,a_j}(\mathrm{d}s), \Gamma'\rangle_s^Q\right] \\
&= \mathbb{E}_Q\left[\langle D\int_0^{\cdot\wedge\tau^{g^*}} \pi_s^*(\{a_j\}) \frac{\mathrm{d}\langle \Gamma, M^{a,a_j}\rangle_s^P - \pi_s(\{a_j\})(P)\,\mathrm{d}\langle\Gamma, M^a\rangle_s^P}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} M^{a,a_j}(\mathrm{d}s), \Gamma'\rangle_s^Q\right] \\
&= \mathbb{E}_P\left[\langle \int_0^{\cdot\wedge\tau^{g^*}} \pi_s^*(\{a_j\}) \frac{\mathrm{d}\langle \Gamma, M^{a,a_j}\rangle_s^P - \pi_s(\{a_j\})(P)\,\mathrm{d}\langle\Gamma, M^a\rangle_s^P}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} M^{a,a_j}(\mathrm{d}s), D^*\Gamma'\rangle_s^P\right] \\
&:= \mathbb{E}_P\left[\langle Y_j\Gamma, D^*\Gamma'\rangle_s^P\right] \\
&= \langle \Gamma, Y_j^* D^*\Gamma'\rangle_P,
\end{aligned}
$$

so if we have found D^* and Y_j^*, we have found the adjoint of H_j. Here, we let $Y_j : \mathcal{M}^2(P) \to \mathcal{M}^2(P)$ be given by

$$
Y_j\Gamma = \int_0^{\cdot\wedge\tau^{g^*}} \sum_k \pi_s^*(\{a_j\}) \frac{\mathrm{d}\langle \Gamma, M^{a,a_j}\rangle_s^P - \pi_s(\{a_j\})(P)\,\mathrm{d}\langle\Gamma, M^a\rangle_s^P}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} M^{a,a_j}(\mathrm{d}s)
$$

Then, we may calculate directly that

$$
\begin{aligned}
\langle Y_j\Gamma, \Gamma'\rangle_P &:= \mathbb{E}_P\left[\langle Y_j\Gamma, \Gamma'\rangle^P\right] \\
&= \mathbb{E}_P\left[\langle \int_0^{\cdot\wedge\tau^{g^*}} \sum_k \pi_s^*(\{a_j\}) \frac{\mathrm{d}\langle \Gamma, M^{a,a_j} - \pi_\cdot(\{a_j\})(P)\bullet M^a\rangle_s^P}{\mathrm{d}\Lambda^{a,a_j}(s)(P)}\,\mathrm{d}M^{a,a_j}(s), \Gamma'\rangle^P\right] \\
&= \mathbb{E}_P\left[\langle \Gamma, \mathbb{1}\{\tau^{g^*} \le \cdot\} \sum_k m^*_{\cdot,k,j} \frac{\mathrm{d}\langle M^{a,a_j}, \Gamma'\rangle_\cdot^P}{\mathrm{d}\Lambda^{a,a_j}(\cdot)(P)} \bullet \left(M^{a,a_j} - \pi_\cdot(\{a_j\})(P)\bullet M^a\right)\rangle^P\right] \\
&= \mathbb{E}_P\left[\langle \Gamma, Y_j^*\Gamma'\rangle^P\right] \\
&= \langle \Gamma, Y_j^*\Gamma'\rangle_P,
\end{aligned}
$$

Now note that

$$
\begin{aligned}
D^*\Gamma' &= \Gamma_0' + [\Gamma', W]_t - \int_0^t W_{s-}\,\mathrm{d}\Gamma_s' \\
&= \Gamma_0' + \int_0^t W_{s-}\,\mathrm{d}\left(\Gamma' + [\Gamma', K^*]\right)_s \\
&= \Gamma_t' W_t - \int_0^t \Gamma_{s-}'\,\mathrm{d}W_s \\
&= \Gamma_t' W_t - \int_0^t \Gamma_{s-}' W_{s-}\,\mathrm{d}K_s^*
\end{aligned}
$$

by the arguments in “Projection Notes” and integration by parts for semimartingales. This composition yields,

$$
\int_0^t \mathbb{1}\{\tau^{g^*} \leq s\} \sum_k \pi^*_s(\{a_j\}) \frac{\mathrm{d}\langle M^{a,a_j}, D^*\Gamma' \rangle^P_s}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \,\mathrm{d}\big(M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a\big)_s
$$
$$
= \int_0^t \mathbb{1}\{\tau^{g^*} \leq s\} \sum_k \pi^*_s(\{a_j\}) W_{s-} \frac{\mathrm{d}\langle M^{a,a_j}, \Gamma' + [\Gamma', K^*] \rangle^P_s}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \,\mathrm{d}\big(M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a\big)_s
$$

FACT:

$$
\langle X, Y \rangle^Q = \langle X + [X, K^*], Y \rangle^P = \langle X, Y + [Y, K^*] \rangle^P,
$$

so it follows that

$$
\int_0^t \mathbb{1}\{\tau^{g^*} \leq s\} \sum_k \pi^*_s(\{a_j\}) W_{s-} \frac{\mathrm{d}\langle M^{a,a_j}, \Gamma' + [\Gamma', K^*] \rangle^P_s}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \,\mathrm{d}\big(M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a\big)_s
$$
$$
= \int_0^t \mathbb{1}\{\tau^{g^*} \leq s\} \sum_k \pi^*_s(\{a_j\}) W_{s-} \frac{\mathrm{d}\langle M^{a,a_j}, \Gamma' \rangle^Q_s}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \,\mathrm{d}\big(M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a\big)_s
$$

Note that this operator sends to piecewise constant functions.
3.

Note that the adjoint operator can be written as $\left(\mathrm{Id} - \sum_j Y^*_j\right) D^*$. If $\mathrm{Id} - \sum_j Y^*_j$ is injective, it holds that

$$
\ker\,(S^*) = \ker\,\left(\left(\mathrm{Id} - \sum_j Y^*_j\right) D^*\right) = \ker\,(D^*) = \{0\}
$$

where the last equality follows by "Projection Notes". So this will follow, if we can show that $\mathrm{Id} - \sum_j Y^*_j$ is injective. To this end, we shall show that $\mathrm{Id} - \sum_j Y_j$ has dense range in $\mathcal{M}^2(P)$ (Corollaries to Theorem 4.12 of Rudin, Functional Analysis, 2nd edition).

Take $\Gamma^* \in \mathcal{M}^2(P)$ such that

$$
\Gamma^* = \int_0^\cdot \sum_x \gamma^*_x(s) M^x(\mathrm{d}s)
$$

where $\gamma^*_x(s)$ is bounded and predictable We shall find $\Gamma \in \mathcal{M}^2(P)$ such that

$$
\Gamma^* = \Gamma - \sum_j Y_j \Gamma.
$$

We will find our solution as

$$
\Gamma = \int_0^\cdot \sum_x \gamma_x(s) M^x(\mathrm{d}s),
$$

where $\gamma_x(s)$ is bounded and predictable. We now explain how this results in dense range: The lemma on p. 173 of Protter (2005) shows that any $\Gamma^* \in \mathcal{M}^2(P)$ can be approximated in $\mathcal{M}^2(P)$ by a sequence of such processes.

We will need to find the $\langle \Gamma, M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a \rangle^P_t$ term.

$$
\langle \Gamma, M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a \rangle^P_t = \int_0^t \sum_x h_x(s) \,\mathrm{d}\langle M^x, M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a \rangle^P_s
$$

Now note that

$$
\begin{aligned}
\langle M^x, M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a \rangle_s^P &= \Lambda_s^{*,x,j} - \int_0^s \Delta \langle M^x \rangle_s \, \mathrm{d}\Lambda^{a,a_j}(s) + \int_0^s \pi_s(\{a_j\})(P) \Delta \langle M^x \rangle_s \, \mathrm{d}\Lambda^a(s) \\
&= \Lambda_s^{*,x,a_j} - \int_0^s \Delta \langle M^x \rangle_s \pi_s(\{a_j\})(P) \, \mathrm{d}\Lambda^a(s) + \int_0^s \pi_s(\{a_j\})(P) \Delta \langle M^x \rangle_s \, \mathrm{d}\Lambda^a(s) \\
&= \Lambda_s^{*,x,a_j}
\end{aligned}
$$

where Λ^{*,x,a_j} is the compensator of $\int_0^t \Delta N_s^x \, \mathrm{d}N^{a,a_j}(s) - \int_0^t \pi_s(\{a_j\})(P) \Delta N^x(s)) \, \mathrm{d}N_s^a$. If $x \notin \mathcal{A}$, then $\Lambda^{*,x,a_j} = 0$. If $x = a_i \in \mathcal{A}$, then

$$
\begin{aligned}
&\int_0^t \Delta N_s^x \, \mathrm{d}N^{a,a_j}(s) - \int_0^t \pi_s(\{a_j\})(P) \Delta N^x(s)) \, \mathrm{d}N_s^a \\
&= \mathbb{1}\{j = i\} N^{a,a_j}(t) - \int_0^t \pi_s(\{a_j\})(P) \, \mathrm{d}N^{a_i}(s)
\end{aligned}
$$

This has the compensator

$$
\begin{aligned}
\Lambda_t^{*,x,a_j} &= \mathbb{1}\{j = i\} \Lambda^{a,a_j}(t) - \int_0^t \pi_s(\{a_j\})(P) \, \mathrm{d}\Lambda^{a,a_i}(s) \\
&= \mathbb{1}\{j = i\} \Lambda^{a,a_j}(t) - \int_0^t \pi_s(\{a_i\})(P) \, \mathrm{d}\Lambda^{a,a_j}(s) \\
&= \int_0^t \mathbb{1}\{j = i\} - \pi_s(\{a_i\})(P) \, \mathrm{d}\Lambda^{a,a_j}(s)
\end{aligned}
$$

where the second properties works by properties of Radon-Nikodym derivatives and positivity. Now, we have that

$$
\begin{aligned}
Y_j \Gamma - \Gamma &= \int_0^{\cdot \wedge \tau^{g^*}} \sum_j \sum_k \pi_s^*(\{a_j\}) \frac{\mathrm{d}\langle \Gamma, M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a \rangle_s^P}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \, \mathrm{d}(M^{a,a_j})_s - \int_0^\cdot \sum_x \gamma_x(s) M^x(\mathrm{d}s) \\
&= \int_0^{\cdot \wedge \tau^{g^*}} \sum_j \sum_k \pi_s^*(\{a_j\}) \sum_{x \in \mathcal{A}} \gamma_x(s) (\mathbb{1}\{a_j = x\} - \pi_s(\{x\})(P)) \, \mathrm{d}M_s^{a,a_j} - \int_0^\cdot \sum_x \gamma_x(s) M^x(\mathrm{d}s)
\end{aligned}
$$

Whenever $x \notin \mathcal{A}$, we can choose $\gamma_x(s) = -\gamma_x^*(s)$. Otherwise, we may pick

$$
\begin{aligned}
\gamma_x(s) &= \mathbb{1}\{s \leq \tau^{g^*}\} \sum_k \mathbb{1}\{T_{(k-1)} < s \leq T_{(k)}\} \mathbb{1}\left\{x \neq g_k^*\left(\mathcal{F}_{T_{(k-1)}}, s\right)\right\} (-\gamma_x^*(s)) \\
&\quad + \mathbb{1}\{s \leq \tau^{g^*}\} \sum_k \mathbb{1}\{T_{(k-1)} < s \leq T_{(k)}\} \mathbb{1}\left\{x = g_k^*\left(\mathcal{F}_{T_{(k-1)}}, s\right)\right\} \left(-\frac{\sum_{y \in \mathcal{A}, y \neq x} \pi_s(\{y\})(P) \gamma_y^*(s)}{\pi_s(\{x\})(P)} \right) \\
&\quad + \mathbb{1}\{s > \tau^{g^*}\} (-\gamma_x^*(s)).
\end{aligned}
$$

which is bounded by assumption and predictable.

4.

To this end, we will need to calculate $\langle K^* \rangle_t^P$ and $\langle K^*, M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a \rangle_t^P$.

$$
\begin{aligned}
\langle K^* \rangle_t^P &= \int_0^{t\wedge\tau^{g^*}} \sum_i \sum_j \left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1 \right) \left(\frac{\pi_s^*(\{a_i\})}{\pi_s(\{a_i\})} - 1 \right) \mathrm{d}\langle M^{a,a_i}, M^{a,a_j} \rangle_s^P \\
&= \int_0^{t\wedge\tau^{g^*}} \sum_i \sum_j \left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1 \right) \left(\frac{\pi_s^*(\{a_i\})}{\pi_s(\{a_i\})} - 1 \right) (\mathbb{1}\{i = j\} - \Delta\Lambda^{a,a_i}) \,\mathrm{d}\Lambda^{a,a_j}(s) \\
&= \int_0^{t\wedge\tau^{g^*}} \sum_j \left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1 \right)^2 \mathrm{d}\Lambda^{a,a_j}(s) \\
&\quad - \int_0^{t\wedge\tau^{g^*}} \sum_i \sum_j \left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1 \right) \left(\frac{\pi_s^*(\{a_i\})}{\pi_s(\{a_i\})} - 1 \right) \Delta\Lambda^{a,a_i} \,\mathrm{d}\Lambda^{a,a_j}(s) \\
&= \int_0^{t\wedge\tau^{g^*}} \sum_j \left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1 \right)^2 \mathrm{d}\Lambda^{a,a_j}(s) \\
&\quad - \int_0^{t\wedge\tau^{g^*}} \sum_i \sum_j \left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1 \right) \left(\frac{\pi_s^*(\{a_i\})}{\pi_s(\{a_i\})} - 1 \right) \pi_s(\{a_i\}) \pi_s(\{a_j\}) \Delta\Lambda^a(s) \,\mathrm{d}\Lambda^a(s) \\
&= \int_0^{t\wedge\tau^{g^*}} \sum_j \left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1 \right)^2 \mathrm{d}\Lambda^{a,a_j}(s) \\
&\quad - \int_0^{t\wedge\tau^{g^*}} \sum_i \sum_j (\pi_s^*(\{a_j\}) - \pi_s(\{a_j\}))(\pi_s^*(\{a_i\}) - \pi_s(\{a_i\})) \Delta\Lambda^a(s) \,\mathrm{d}\Lambda^a(s) \\
&= \int_0^{t\wedge\tau^{g^*}} \sum_j \left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1 \right)^2 \mathrm{d}\Lambda^{a,a_j}(s) \\
&= \int_0^{t\wedge\tau^{g^*}} \sum_j \frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} \,\mathrm{d}\Lambda^a(s) - 2 \int_0^{t\wedge\tau^{g^*}} \sum_j (\pi_s^*(\{a_j\})) \,\mathrm{d}\Lambda^a(s) + \int_0^{t\wedge\tau^{g^*}} \sum_j \pi_s(\{a_j\})(P) \,\mathrm{d}\Lambda^a(s) \\
&= \int_0^{t\wedge\tau^{g^*}} \left(\sum_j \frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1 \right) \mathrm{d}\Lambda^a(s).
\end{aligned}
$$

For the next calculation, we have that

$$
\begin{aligned}
\langle K^*, M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a \rangle_t^P &= \int_0^{t\wedge\tau^{g^*}} \sum_i \left(\frac{\pi_s^*(\{a_i\})}{\pi_s(\{a_i\})} - 1 \right) \mathrm{d}\langle M^{a,a_i}, M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a \rangle_s^P \\
&= \int_0^{t\wedge\tau^{g^*}} \sum_i \left(\frac{\pi_s^*(\{a_i\})}{\pi_s(\{a_i\})} - 1 \right) (\mathbb{1}\{i = j\} - \pi_s(\{a_i\})(P)) \,\mathrm{d}\Lambda^{a,a_j}(s).
\end{aligned}
$$

by the previous calculations. Therefore, it holds that

$$
\begin{aligned}
&\int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \frac{\mathrm{d}\langle \Gamma, M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a\rangle_t^P}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} \left(M^{a,a_j}(\mathrm{d}s) - \mathrm{d}\langle M^{a,a_j}, K^*\rangle_s^P\right) \\
&= \int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \sum_i \left(\frac{\pi_s^*(\{a_i\})}{\pi_s(\{a_i\})} - 1\right) (\mathbb{1}\{i=j\} - \pi_s(\{a_i\})(P))(M^{a,a_j}(\mathrm{d}s) - \mathrm{d}\langle M^{a,a_j}, K^*\rangle_s^P) \\
&= \int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \sum_i \left(\frac{\pi_s^*(\{a_i\})}{\pi_s(\{a_i\})} - 1\right) (\mathbb{1}\{i=j\} - \pi_s(\{a_i\})(P)) \\
&\quad \times \left(M^{a,a_j}(\mathrm{d}s) - \left(\sum_v \frac{\pi_s^*(\{a_v\})}{\pi_s(\{a_v\})} - 1\right) \mathrm{d}\langle M^{a,a_j}, M^{a,v}\rangle_s^P\right) \\
&= \int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \sum_i \left(\frac{\pi_s^*(\{a_i\})}{\pi_s(\{a_i\})} - 1\right) (\mathbb{1}\{i=j\} - \pi_s(\{a_i\})(P)) \\
&\quad \times \left(M^{a,a_j}(\mathrm{d}s) - \left(\sum_v \frac{\pi_s^*(\{a_v\})}{\pi_s(\{a_v\})} - 1\right) \pi_s(\{a_v\})\left(\mathbb{1}\{j=v\} - \pi\left(\{a_j\}\Delta\Lambda^a\right) \mathrm{d}\Lambda^a(s)\right)\right) \\
&= \int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \sum_i \left(\frac{\pi_s^*(\{a_i\})}{\pi_s(\{a_i\})} - 1\right) (\mathbb{1}\{i=j\} - \pi_s(\{a_i\})(P)) \\
&\quad \times \left(M^{a,a_j}(\mathrm{d}s) - \left(\sum_v \frac{\pi_s^*(\{a_v\})}{\pi_s(\{a_v\})} - 1\right) \pi_s(\{a_v\}) \mathbb{1}\{j=v\} \,\mathrm{d}\Lambda^a(s)\right) \\
&= \int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \sum_i \left(\frac{\pi_s^*(\{a_i\})}{\pi_s(\{a_i\})} - 1\right) (\mathbb{1}\{i=j\} - \pi_s(\{a_i\})(P)) \\
&\quad \times \left(M^{a,a_j}(\mathrm{d}s) - \left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1\right) \pi_s(\{a_j\}) \,\mathrm{d}\Lambda^a(s)\right) \\
&= \int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \sum_i \left(\frac{\pi_s^*(\{a_i\})}{\pi_s(\{a_i\})} - 1\right) (\mathbb{1}\{i=j\} - \pi_s(\{a_i\})(P)) \\
&\quad \times \left(M^{a,a_j}(\mathrm{d}s) - \left(\pi_s^*(\{a_j\}) - \pi_s(\{a_j\})\right) \mathrm{d}\Lambda^a(s)\right) \\
&= \int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \frac{1}{\pi_s(\{a_j\})} (1 - \pi_s(\{a_i\})(P)) \\
&\quad \times \left(N^{a,a_j}(\mathrm{d}s) - \pi_s^*(\{a_j\}) \,\mathrm{d}\Lambda^a(s)\right) \\
&- \int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \sum_i (\mathbb{1}\{i=j\} - \pi_s(\{a_i\})(P)) \\
&\quad \times \left(N^{a,a_j}(\mathrm{d}s) - \pi_s^*(\{a_j\}) \,\mathrm{d}\Lambda^a(s)\right) \\
&= \int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \pi_s^*(\{a_j\}) \left(\frac{1}{\pi_s(\{a_j\})} - 1\right) (N^{a,a_j}(\mathrm{d}s) - 1) \,\mathrm{d}\Lambda^a(s))
\end{aligned}
$$

This we find the score operator is given by evaluated at K^* by combining the results,

$$
-\int_0^{t\wedge\tau^{g^*}} \sum_{j=1}^K \mathbb{1}\left\{T_{(k-1)} < s \le T_{(k)}\right\} \mathbb{1}\left\{j \neq g_k^*\left(\mathcal{F}_{T_{(k-1)}}, s\right)\right\} N^{a,a_j}(\mathrm{d}s)
$$

which is zero Q-a.s.

3.a) Efficient influence function calculations

Let $I_P \nu\left(\tilde{Y}\right)$ be an influence function for $\nu : \mathcal{P}_{|\sigma\left(\tilde{Y}\right)} \to R$, $\nu(P) = E_P\left[\tilde{Y}\right]$ and let us take $I_P \nu\left(\tilde{Y}\right) = \tilde{Y} - E_P\left[\tilde{Y}\right]$. Let $\zeta_t = \mathbb{E}_P[I_P \nu(Y) \mid \mathcal{F}_t] = \mathbb{E}_P\left[\tilde{Y} \mid \mathcal{F}_t\right] - E_P\left[\tilde{Y}\right]$. Now, we see that

$$
\begin{aligned}
S^* \zeta_T &= W_T(P)\left(Y_T - \mathbb{E}_P\left[\tilde{Y}_T\right]\right) \\
&- \int_0^T \zeta_{s-} W(s-) \, \mathrm{d}K_s^* \\
&- \int_0^T \mathbb{1}\{s \leq \tau^{g^*}\} \sum_j \pi_s^*(\{a_j\}) \frac{\mathrm{d}\langle M^{a,a_j}, \mathbb{E}_P[W_T(P) Y_T \mid \mathcal{F}_\cdot] - E_P\left[\tilde{Y}\right] \rangle_s^Q}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \, \mathrm{d}\left(M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a\right)_s \\
&= W_T(P) Y_T - \left(W_T(P) - \int_0^T W(s-) \, \mathrm{d}K_s^*\right) \mathbb{E}_P\left[\tilde{Y}_T\right] \\
&- \int_0^T \mathbb{E}_Q[Y_T \mid \mathcal{F}_{s-}] W(s-) \, \mathrm{d}K_s^* \\
&- \int_0^T \mathbb{1}\{s \leq \tau^{g^*}\} \sum_j \pi_s^*(\{a_j\}) \frac{\mathrm{d}\langle M^{a,a_j}, \mathbb{E}_P[W_T(P) Y_T \mid \mathcal{F}_\cdot] \rangle_s^Q}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \, \mathrm{d}\left(M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a\right)_s \\
&= W_T(P) Y_T - \mathbb{E}_P\left[\tilde{Y}_T\right] \\
&- \int_0^T \mathbb{E}_Q[Y_T \mid \mathcal{F}_s] W(s-) \, \mathrm{d}K_s^* \\
&+ \left[\mathbb{E}_Q[Y_T \mid \mathcal{F}_\cdot], K^*\right]_T \\
&- \int_0^T \mathbb{1}\{s \leq \tau^{g^*}\} \sum_j \pi_s^*(\{a_j\}) \frac{\mathrm{d}\langle M^{a,a_j}, \mathbb{E}_P[W_T(P) Y_T \mid \mathcal{F}_\cdot] \rangle_s^Q}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \, \mathrm{d}\left(M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a\right)_s \\
&= W_T(P) Y_T - \mathbb{E}_P\left[\tilde{Y}_T\right] \\
&- \int_0^T \mathbb{E}_Q[Y_T \mid \mathcal{F}_s] W(s-) \, \mathrm{d}K_s^* \\
&+ \int_0^T \sum_{j=1}^K \left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1\right) \mathrm{d}\left[\mathbb{E}_P[W_T Y_T \mid \mathcal{F}_\cdot], M^{a,a_j}\right]_s \\
&- \int_0^T \mathbb{1}\{s \leq \tau^{g^*}\} \sum_j \pi_s^*(\{a_j\}) \frac{\mathrm{d}\langle M^{a,a_j}, \mathbb{E}_P[W_T(P) Y_T \mid \mathcal{F}_\cdot] \rangle_s^P}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \, \mathrm{d}\left(M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a\right)_s \\
&- \int_0^T \mathbb{1}\{s \leq \tau^{g^*}\} \sum_j \pi_s^*(\{a_j\}) \frac{\mathrm{d}\langle M^{a,a_j}, [\mathbb{E}_P[W_T(P) Y_T \mid \mathcal{F}_\cdot], K^*] \rangle_s^P}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \, \mathrm{d}\left(M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a\right)_s
\end{aligned}
$$

The integrators can be replaced with their corresponding martingales.

Ohlendorff et al. (2025) suggested the EIF:

$$
\begin{aligned}
& W_T(P) Y_T - E_P\left[\tilde{Y}_T\right] \\
& - \int_0^T W(s-) \mathbb{E}_Q[Y_T \mid \mathcal{F}_s] \sum_j \frac{\pi^*_s(\{a_j\})}{\pi_s(\{a_j\})} N^{a,a_j}(\mathrm{d}s) + \int_0^T W(s-) \mathbb{E}_Q[Y_T \mid \mathcal{F}^g_s] N^a(\mathrm{d}s) \\
& = W_T(P) Y_T - E_P\left[\tilde{Y}_T\right] \\
& - \int_0^T W(s-) \mathbb{E}_Q[Y_T \mid \mathcal{F}_s] \sum_j \frac{\pi^*_s(\{a_j\})}{\pi_s(\{a_j\})} N^{a,a_j}(\mathrm{d}s) + \int_0^T W(s-) \mathbb{E}_Q[Y_T \mid \mathcal{F}^g_s] N^a(\mathrm{d}s) \\
& - \int_0^T W(s-) \mathbb{E}_Q[Y_T \mid \mathcal{F}_s] N^\ell(\mathrm{d}s) + \int_0^T W(s-) \mathbb{E}_Q[Y_T \mid \mathcal{F}_s] N^\ell(\mathrm{d}s) \\
& = -E_P\left[\tilde{Y}_T\right] \\
& - \int_0^T W(s-) \mathbb{E}_Q[Y_T \mid \mathcal{F}_s] \sum_j \frac{\pi^*_s(\{a_j\})}{\pi_s(\{a_j\})} N^{a,a_j}(\mathrm{d}s) + \int_0^T W(s-) \mathbb{E}_Q[Y_T \mid \mathcal{F}^g_s] N^a(\mathrm{d}s) \\
& - \int_0^T W(s-) \mathbb{E}_Q[Y_T \mid \mathcal{F}_s] N^\ell(\mathrm{d}s) + \int_0^T W(s-) \mathbb{E}_Q[Y_T \mid \mathcal{F}_s] N^\ell(\mathrm{d}s) \\
& + \int_0^T W(s-) N^y(\mathrm{d}s) \\
& = \mathbb{E}_Q[Y_t \mid \mathcal{F}_0] - E_P\left[\tilde{Y}_T\right] \\
& - \int_0^T W(s-) \mathbb{E}_Q[Y_T \mid \mathcal{F}_s] \sum_j \frac{\pi^*_s(\{a_j\})}{\pi_s(\{a_j\})} N^{a,a_j}(\mathrm{d}s) + \int_0^T W(s-) \mathbb{E}_Q[Y_T \mid \mathcal{F}^g_s] N^a(\mathrm{d}s) \\
& - \int_0^T W(s-) \mathbb{E}_Q[Y_T \mid \mathcal{F}_s] N^\ell(\mathrm{d}s) + \int_0^T W(s-) \mathbb{E}_Q[Y_T \mid \mathcal{F}_s] N^\ell(\mathrm{d}s) \\
& + \int_0^T W(s-) N^y(\mathrm{d}s) - \mathbb{E}_Q[Y_t \mid \mathcal{F}_0]
\end{aligned}
$$

The latter function is a rewriting of the one in Ohlendorff et al. (2025), while the first is suggested by projecting onto the propensity tangent space. Here, $\mathcal{F}^g_s$ denotes the natural filtration, generated by observed data with the treatment decisions specified by the treatment regime g^*.

We can write

$$
\begin{aligned}
& - \int_0^T W(s-) \mathbb{E}_Q[Y_T \mid \mathcal{F}_s] \sum_j \frac{\pi^*_s(\{a_j\})}{\pi_s(\{a_j\})} N^{a,a_j}(\mathrm{d}s) + \int_0^T W(s-) \mathbb{E}_Q[Y_T \mid \mathcal{F}^g_s] N^a(\mathrm{d}s) \\
& = - \int_0^T W(s-) \mathbb{E}_Q[Y_T \mid \mathcal{F}_s] \sum_j \left(\frac{\pi^*_s(\{a_j\})}{\pi_s(\{a_j\})} - 1 \right) N^{a,a_j}(\mathrm{d}s) + \int_0^T W(s-) \mathbb{E}_Q[Y_T \mid \mathcal{F}^g_s] N^a(\mathrm{d}s) \\
& - \int_0^T W(s-) \mathbb{E}_Q[Y_T \mid \mathcal{F}_s] N^a(\mathrm{d}s)
\end{aligned}
$$

3.b) Comparisons of the positivity assumptions in Ryalen (2024)

One may ask oneself if positivity holds in Ryalen (2024); under what assumptions does positivity in Theorem 1.1 hold? In general, however, it would appear that the two positivity conditions are different and neither implies the other.

Can we find a process φ such that $\mathcal{E}(K) = \frac{\mathcal{E}(-\mathbb{N}^a)}{\mathcal{E}(-\mathbb{L}^a)}\mathcal{E}(\varphi)$?

Theorem 3.2.1: φ is given by

$$\varphi_t = K_t - \mathbb{L}_t^a + \mathbb{N}_t^a - [K, \mathbb{L}^a]_t,$$

where $[\cdot, \cdot]$ denotes the quadratic covariation process (Protter (2005)), where

$$[K, \mathbb{L}^a]_t = \int_0^{t \wedge \tau^{g^*}} \sum_v \mathbb{1}\left\{T_{(v-1)} < s \leq T_{(v)}\right\} \sum_{i \neq g_v^*\left(\mathcal{F}_{T_{(v-1)}}, T_{(v)}\right)} \pi_s(\{a_j\}) \Delta\Lambda^a(s)$$
$$\times \sum_{j=1}^k \left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1\right) N^{a,a_j}(\mathrm{d}s)$$

In the absolutely continuous case, $[K, \mathbb{L}^a]_t = 0$ as $\Delta\Lambda_t^a = 0$ for all $t > 0$. If, further, $\pi_t^*(\{a_j\}) = 1$ for some j, then

$$\varphi_t = \int_0^{t \wedge \tau^{g^*}} \left(\frac{1}{\pi_s} - 1\right) M^{a,a_j}(\mathrm{d}s).$$

Proof: To this end, let $v := \mathbb{1}\left\{W(t) > 0, \widetilde{W}(t) > 0\right\} = \mathbb{1}\{\tau^{g^*} > t\} = \mathcal{E}(-\mathbb{N}^a)$ and calculate

$$\begin{aligned}
\mathcal{E}(\varphi)v &= \frac{\mathcal{E}(K)\mathcal{E}(-\mathbb{L}^a)}{\mathcal{E}(-\mathbb{N}^a)}v \\
&= \mathcal{E}(K)\mathcal{E}(-\mathbb{L}^a)v \\
&= \mathcal{E}(K - \mathbb{L}^a - [K, \mathbb{L}^a])v \\
&= \mathcal{E}(K - \mathbb{L}^a + \mathbb{N}^a - [K, \mathbb{L}^a])v,
\end{aligned}$$

where the last equality follows since $\mathbb{N}^a v \equiv 0$. Note that

$$\begin{aligned}
[K, \mathbb{L}^a]_t &= \int_0^t \Delta\mathbb{L}_s^a \sum_{j=1}^k \left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1\right) N^{a,a_j}(\mathrm{d}s) \\
&\stackrel{*}{=} \int_0^{t \wedge \tau^{g^*}} \sum_v \mathbb{1}\left\{T_{(v-1)} < s \leq T_{(v)}\right\} \sum_{i \neq g_v^*\left(\mathcal{F}_{T_{(v-1)}}, T_{(v)}\right)} \pi_s(\{a_j\}) \Delta\Lambda^a(s) \\
&\quad \times \sum_{j=1}^k \left(\frac{\pi_s^*(\{a_j\})}{\pi_s(\{a_j\})} - 1\right) N^{a,a_j}(\mathrm{d}s)
\end{aligned}$$

In the case where $\Delta\Lambda_s^a \equiv 0$ and $\pi_s^*(\{a_j\}) = 1$ for some j, then

$$v(K_t - \mathbb{L}_t^a + \mathbb{N}_t^a) = v \int_0^{t \wedge \tau^{g^*}} \left(\frac{1}{\pi_s} - 1\right) M^{a,a_j}(\mathrm{d}s)$$

and $[K, \mathbb{L}^a]_t = 0$. □

A simple consequence of this is the following. Assume that $Q_{\text{ryalen}} \ll P$ with $Q_{\text{ryalen}} = \frac{\mathcal{E}(-\mathbb{N}^a)}{\mathcal{E}(-\mathbb{L}^a)} \bullet P$. and, say, $\mathcal{E}(\varphi)$ is a uniformly integrable Q_{ryalen}-$\mathcal{F}_t$-martingale, i.e., that $Q \ll Q_{\text{ryalen}}$, then $Q_{\text{ryalen}} \ll P$ implies that $Q \ll P$. This happens for example if $\mathcal{E}(\varphi)$ is uniformly bounded by a constant.

Bibliography

Andersen, P. K., Borgan, Ø., Gill, R. D., & Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer US. https://doi.org/10.1007/978-1-4612-4348-9

Gill, R. D., & Robins, J. M. (2001). Causal Inference for Complex Longitudinal Data: The Continuous Case. *The Annals of Statistics*, *29*(6), 1785–1811. https://doi.org/10.1214/aos/1015345962

Last, G., & Brandt, A. (1995). *Marked Point Processes on the Real Line: The Dynamical Approach*. Springer. https://link.springer.com/book/9780387945477

Ohlendorff, J. S., Munch, A., & Gerds, T. A. (2025). *A Sequential Regression Approach for Efficient Continuous-Time Causal Inference*.

Protter, P. E. (2005). *Stochastic differential equations*. Springer.

Ryalen, P. (2024). *On the role of martingales in continuous-time causal inference*.

Rytgaard, H. C., Gerds, T. A., & van der Laan, M. J. (2022). Continuous-Time Targeted Minimum Loss-Based Estimation of Intervention-Specific Mean Outcomes. *The Annals of Stat Istics*, *50*(5), 2469–2491. https://doi.org/10.1214/21-AOS2114

van der Vaart, A. (2004). On Robins' formula. *Statistics & Decisions*, *22*(3), 171–200. https://doi.org/doi:10.1524/stnd.22.3.171.57065