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;x 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 $\mathbb{R}^d$. Such a random measure gives rise to a filtration $\left(\mathcal{F}_t\right)_{t \geq 0}$, where

$$\mathcal{F}_t := \sigma\left(N^y(\mathrm{d}s), N^a(\mathrm{d}s \times \{x\}), N^\ell(\mathrm{d}s \times \{y\}) \mid s \in (0, t], x \in \mathcal{A}, y \in \mathcal{L}\right).$$

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 functions $g_k^* : \mathbb{H}_{k-1} \times [0, T] \to \mathcal{A}$ such that

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

i.e., g_k^* is $\mathcal{F}_{T_{(k-1)}} \otimes \mathcal{B}([0, T])$-measurable 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 $N^a(\mathrm{d}t \times \mathcal{A})$. 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

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

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

Definition 1.1: 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. 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 $\mathscr{L}$ denote the P-$\mathcal{F}_t$-canonical compensator of N^{g^*} and let Λ denote the P-$\mathcal{F}_t$-compensator of N.

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 $x \in \{y, \ell\}$.

The main outcome of interest is the counterfactual outcome process $\tilde{Y} := \tilde{N}^y$; and we wish to identify $\mathbb{E}_P\left[\tilde{Y}_t\right]$.

Let $\left(T_{(k)}, \Delta_{(k)}, A\left(T_{(k)}\right)\right)$ denote the ordered event times, event types, and treatment decisions at event k. Note that Equation 1 is the same likelihood ratio as in Rytgaard et al. (2022). As a shorthand, we let $N^{a,x}(\mathrm{d}t) := N^a(\mathrm{d}t \times \{x\})$ for $x \in \mathcal{A}$.

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 $\mathcal{H}_t := \mathcal{F}_t \vee \sigma\left(\tilde{Y}\right)$. The Radon-Nikodym derivative of the P-$\mathcal{F}_t$ compensator with respect to the total P-$\mathcal{H}_t$ compensator is the same as the Radon-Nikodym derivative of the P-$\mathcal{F}_t$ compensator with respect to the total P-$\mathcal{F}_t$ compensator.
- **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

$$\mathbb{E}_P[Y_t W(t)] \overset{(**)}{=} \mathbb{E}_P\left[\tilde{Y}_t W(t)\right] = \mathbb{E}_P\left[\tilde{Y}_t \mathbb{E}_P[W(t) \mid \mathcal{H}_0]\right] = \mathbb{E}_P\left[\tilde{Y}_t W(0)\right] = \mathbb{E}_P\left[\tilde{Y}_t\right]$$

where in $(**)$ we used consistency by noting that $W(t) \neq 0$ if and only if $\tau^{g^*} > t$. □

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 $\mathcal{H}_t := \mathcal{F}_t \vee \sigma\left(\tilde{Y}\right)$. 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.

1.a) Comparison with the exchangeability condition of Ryalen (2024)

Theorem 1.1.1: Let $\mathbb{N}_t^a = \mathbb{1}\{\tau^{g^*} \leq t\}$. The exchangeability condition of Theorem 1.1 implies the one of Ryalen (2024), e.g., $\mathbb{L}_t$ is both the P-$\mathcal{F}_t$ compensator and the P-$\mathcal{H}_t$ compensator of $\mathbb{N}_t^a$.

Proof: First, let

$$N^{a,*}(\mathrm{d}t) = \sum_k \sum_j \mathbb{1}\left\{T_{(k-1)} < t \leq T_{(k)}\right\} \mathbb{1}\left\{j \neq g_k^*\left(\mathcal{F}_{T_{(k-1)}}, t\right)\right\} N^{a,a_j}(\mathrm{d}t),$$

which is the counting measure for the number of deviations from the treatment regime g^* and consider

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

for some K. Note that $\left(\mathbb{1}\{\tau^{g^*} \leq t\} N^{a,*}(\mathrm{d}t)\right)^p = \mathbb{1}\{\tau^{g^*} \leq t\} \Lambda^{a,*}(\mathrm{d}t)$, where p denotes the predictable projection (Last & Brandt (1995)), by definition of the predictable projection. On the other hand,

$$\begin{aligned}
&\left(\sum_{k=1}^{K} \sum_{j} \mathbb{1}\left\{T_{(k-1)} < t \leq T_{(k)}\right\} \mathbb{1}\left\{j \neq g_k^*\left(\mathcal{F}_{T_{(k-1)}}, t\right)\right\} N^{a,a_j}(\mathrm{d}t)\right)^p \\
&= \sum_{k=1}^{K} \sum_{j} \left(\mathbb{1}\left\{T_{(k-1)} < t \leq T_{(k)}\right\} \mathbb{1}\left\{j \neq g_k^*\left(\mathcal{F}_{T_{(k-1)}}, t\right)\right\} N^{a,a_j}(\mathrm{d}t)\right)^p \\
&\overset{(*)}{=} \sum_{k=1}^{K} \sum_{j} \mathbb{1}\left\{T_{(k-1)} < t \leq T_{(k)}\right\} \mathbb{1}\left\{j \neq g_k^*\left(\mathcal{F}_{T_{(k-1)}}, t\right)\right\} (N^{a,a_j}(\mathrm{d}t))^p \\
&= \sum_{k=1}^{K} \sum_{j} \mathbb{1}\left\{T_{(k-1)} < t \leq T_{(k)}\right\} \mathbb{1}\left\{j \neq g_k^*\left(\mathcal{F}_{T_{(k-1)}}, t\right)\right\} \Lambda^{a,a_j}(\mathrm{d}t),
\end{aligned} \tag{3}$$

where we use in $(*)$ that $\mathbb{1}\left\{T_{(k-1)} < t \leq T_{(k)}\right\} \mathbb{1}\left\{j \neq g_k^*\left(\mathcal{F}_{T_{(k-1)}}, t\right)\right\}$ is predictable (Theorem 2.2.22 of Last & Brandt (1995)). Letting $K \to \infty$ and using the assumption of no explosion shows that the P-$\mathcal{F}_t$-compensator of $N^{a,*}(\mathrm{d}t)$ is given by the limit as $K \to \infty$ of the right-hand side of Equation 3. Since $\mathbb{N}_t^a = N^{a,*}\left(\left(0, t \wedge \tau^{g^*}\right]\right)$, the corollary on p. 10 of Protter (2005), implies that the P-$\mathcal{F}_t$-compensator of $\mathbb{N}_t^a$ is

$$\mathbb{L}_t^a = \sum_{k} \sum_{j} \mathbb{1}\left\{T_{(k-1)} < t \leq T_{(k)}\right\} \mathbb{1}\left\{j \neq g_k^*\left(\mathcal{F}_{T_{(k-1)}}, t\right)\right\} \Lambda^{a,a_j}\left(t \wedge \tau^{g^*}\right).$$

by using the limit of the right-hand side of Equation 3. If exchangeability holds (given in Theorem 1.1), then the same argument applies with $\mathcal{H}_t$ instead of $\mathcal{F}_t$. This is the desired result. □

1.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 1.2.1: φ is given by

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

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 $\Delta\Lambda^a_t = 0$ 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

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

In the case where $\Delta\Lambda^a_s \equiv 0$ and $\pi^*_s(\{a_j\}) = 1$ for some j, then

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

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.

2) Comparison with the 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 slightly weaker than 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 2.1 (g-formula): Let $\Lambda_P^a(\mathrm{d}t)$ denote the total P-$\mathcal{F}_t$-compensator of N^a. Furthermore, let Λ_P^x denote the P-$\mathcal{F}_t$-compensator of N^x for $x \in \{y, \ell\}$. 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

$$\int_0^t \frac{1}{W_{s-}} \, \mathrm{d}\langle W, X\rangle_s^P = \langle K, X\rangle_t^P \tag{4}$$

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

$$
\begin{aligned}
\langle K, M^{a,x}\rangle_t^P &= \int_0^t \sum_j \left(\frac{\pi_s^*(a_j)}{\pi_s(a_j)} - 1\right) \mathrm{d}\langle M_P^{a,a_j}, M_P^{a,x}\rangle_s^P \\
&= \int_0^t \left(\frac{\pi_s^*(x)}{\pi_s(x)} - 1\right) \mathrm{d}\langle M_P^{a,x}\rangle_s^P \\
&\quad + \sum_{j \neq x} \int_0^t \left(\frac{\pi_s^*(a_j)}{\pi_s(a_j)} - 1\right) \mathrm{d}\langle M_P^{a,a_j}, M_P^{a,x}\rangle_s^P \\
&= \int_0^t \left(\frac{\pi_s^*(x)}{\pi_s(x)} - 1\right) \pi_s(x)\Lambda_P^a(\mathrm{d}s) \\
&\quad - \int_0^t \left(\frac{\pi_s^*(x)}{\pi_s(x)} - 1\right) \Delta(\pi(x)\Lambda_P^a)_s \pi_s(x)\Lambda_P^a(\mathrm{d}s) \qquad (5) \\
&\quad - \sum_{j \neq x} \int_0^t \left(\frac{\pi_s^*(a_j)}{\pi_s(a_j)} - 1\right) \Delta(\pi(x)\Lambda_P^a)_s \pi_s(a_j)\Lambda_P^a(\mathrm{d}s) \\
&= \int_0^t (\pi_s^*(x) - \pi_s(x))\Lambda_P^a(\mathrm{d}s) \\
&\quad - \sum_j \int_0^t (\pi_s^*(a_j) - \pi_s(a_j))\Delta(\pi(x)\Lambda_P^a)_s \Lambda_P^a(\mathrm{d}s) \\
&= \int_0^t (\pi_s^*(x) - \pi_s(x))\Lambda_P^a(\mathrm{d}s).
\end{aligned}
$$

Girsanov's theorem (Theorem 41, p. 136 of Protter (2005)) together with Equation 4 and Equation 5 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. □

2.a) Comparison of the exchangeability condition of Rytgaard et al. (2022)

Additionally, consistency is not explicitly stated, so we cannot compare this condition. The stated positivity condition is the same as the one in Theorem 1.1, so left is to compare the exchangeability conditions.

We can also ask ourselves: Is the exchangeability criterion in Theorem 1.1 close in interpretation to the statement of Rytgaard et al. (2022)? In Rytgaard et al. (2022), the statement is:

$$
\left(\tilde{Y}_t\right)_{t \in [0,T]} \perp A(T_k^a) \mid \mathcal{F}_{T_k^a -}, \qquad (')
$$

for all k, where T^a_k are the ordered treatment event times, where $\mathcal{F}_T^-$ is defined on p. 62 of Last & Brandt (1995). This σ-algebra contains all the information that occurs strictly before time T.

In this case, we can express one condition (in addition to another) for our exchangeability as something that is very similar (Theorem 2.1.1),

$$\left(\tilde{Y}_t\right)_{t \in [0,T]} \perp A\left(T_{(k)}\right) \mid \Delta N^a_{T_{(k)}} = 1, T_{(k)}, \mathcal{F}_{T_{(k-1)}}, \qquad (*)$$

for all k. A slightly weaker statement than (*) is that

- The Radon-Nikodym derivative of $\Lambda^{a,a_j}(\mathrm{d}t)$ with respect to $\Lambda^a(\mathrm{d}t)$ is the same for $\mathcal{F}_t$ and $\mathcal{H}_t$. (*!)

The statements $(')$ and $(*)$ appear similar, but are generally not the same, since S and ΔN^a_S are not generally $\mathcal{F}_T^-$ measurable for a stopping time S.

Let $S = T^a_k$ for some k. If S and ΔN^a_S are $\mathcal{F}_S^-$ measurable, then whenever $(')$ holds, we have

$$\left(\tilde{Y}_t\right)_{t \in [0,T]} \perp A(T^a_k) \mid T^a_k, \Delta N^a_{T^a_k}, \mathcal{F}_{T^a_k}^-$$

which implies (*).

If $N^a(\cdot \times \{0,1\})$ is a predictable counting process with respect to $\mathcal{F}_t$, we will see that this indeed is the case. If this holds, S is predictable, and so $S \in \sigma(\mathcal{F}_S^-)$ (Theorem 2.2.19 of Last & Brandt (1995)). If N^a_t-predictable, then $N^a_t \in \sigma(\mathcal{F}_{t-}, t)$ (Theorem 2.2.6 of Last & Brandt (1995)); therefore $\Delta N^a_S \in \sigma(\mathcal{F}_S^-, S)$.

To have exchangeability in the sense of Theorem 2.1.1, we also need that the compensator for $N^a = N^a(\cdot \times \{0,1\})$ is the same under $\mathcal{F}_t$ and $\mathcal{H}_t$, i.e., that

- $\Lambda^a(\mathrm{d}t)$ is the P-$\mathcal{F}_t$-compensator and the P-$\mathcal{H}_t$-compensator of $N^a(\mathrm{d}t \times \{0,1\})$. (**)

This is obviously the case if $N^a(\cdot \times \mathcal{A})$ is $\mathcal{F}_t$-predictable.

We then have the following result.

Theorem 2.1.1: The conditions (*!) and (**) hold if and only if the exchangeability condition of Theorem 1.1 holds.

Proof: Obvious.

□

3) Which other solutions can we think of besides the one given in Theorem 1.1 and in Ryalen (2024)?

Theorem 3.1: Let K_t^* be a (local) martingale with $\Delta K_t^* \geq -1$ and $\Delta K_t^* > -1$ if $t < \tau^{g^*}$. Then,

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

if and only if

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

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

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

This holds if and only if

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

if and only if

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

and this is

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

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

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

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

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

□

Now, we consider only K^*'s of the form

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

with $\tilde{h}(s, x)$ P-$\mathcal{F}_t$ predictable with the restriction stated in the above theorem. The above theorem gives that we must have

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

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

$$h(\tau^{g^*}, a_1)(0 - \pi_{\tau^{g^*}}(a_1)\Delta\Lambda^a_{\tau^{g^*}}) + h(\tau^{g^*}, a_0)(1 - (1 - \pi_{\tau^{g^*}}(a_1))\Delta\Lambda^a_{\tau^{g^*}}) = -1$$

or

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

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.

3.a) Absolutely continuous case

In this case, conclude that $h(\tau^{g^*}, a_0) = -1$. However, nothing else can be said about $h(\tau^{g^*}, a_1)$ 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).

3.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,

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

or

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

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 $\mathcal{E}(K^h)_t$ does not depend on the choice of h either, and we may conclude that $\mathcal{E}(K^h)_t = \mathcal{E}(K)_t$.

3.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

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

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

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

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

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

Therefore, it will again be the case that

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

Conclude that

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

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

4) Score operator calculations

Let $\bar{K}_t = K_t^* + \mathbb{N}_t^a$ 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

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

as $\varepsilon \to 0$, where we use dominated convergence and L'Hopitals rule for the last step. The result is presented in Equation 9.

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

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

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

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

so we conclude that

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

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

Let $m^*_{s,k,j}$ predictable be given by $m^*_{s,k,j} = \mathbb{1}\left\{T_{(k-1)} < s \leq T_{(k)}\right\}\mathbb{1}\left\{j = g^*_k\left(\mathcal{F}_{T_{(k-1)}}, s\right)\right\}$. With $K^*_t = K_{t \wedge \tau^{g*}}$, we can take

$$\begin{aligned}\bar{K}_t &= \int_0^{t\wedge\tau^{g^*}} \sum_k \sum_{j=1}^K \mathbb{1}\left\{T_{(k-1)} < s \leq T_{(k)}\right\}\left(\frac{\pi^*_s(\{a_j\})}{\pi_s(\{a_j\})} - 1\right) N^{a,a_j}(\mathrm{d}s) + \mathbb{N}^a_t \\ &= \int_0^{t\wedge\tau^{g^*}} \sum_k \sum_{j=1}^K \mathbb{1}\left\{T_{(k-1)} < s \leq T_{(k)}\right\}\mathbb{1}\left\{j \neq g^*_k\left(\mathcal{F}_{T_{(k-1)}}, s\right)\right\}(-1) N^{a,a_j}(\mathrm{d}s) + \mathbb{N}^a_t \\ &\quad + \int_0^{t\wedge\tau^{g^*}} \sum_k \sum_{j=1}^K m^*_{s,k,j}\left(\frac{1}{\pi_s(\{a_j\})} - 1\right) N^{a,a_j}(\mathrm{d}s) \\ &= \int_0^{t\wedge\tau^{g^*}} \sum_k \sum_{j=1}^K m^*_{s,k,j}\left(\frac{1}{\pi_s(\{a_j\})} - 1\right) N^{a,a_j}(\mathrm{d}s)\end{aligned} \tag{6}$$

This can also be written as

$$\bar{K}_t = \int_0^{t\wedge\tau^{g^*}} \sum_k \sum_{j=1}^K m^*_{s,k,j}\left(\frac{1}{\pi_s(\{a_j\})} - 1\right) M^{a,a_j}(\mathrm{d}s) + \mathbb{L}^a_t \tag{7}$$

Calculating the derivative of Equation 7 gives

$$
\begin{aligned}
\partial_\varepsilon \bar{K}^\varepsilon_t &= -\int_0^{t\wedge\tau^{g^*}} \sum_k \sum_{j=1}^K m^*_{s,k,j} \frac{1}{\pi_s(\{a_j\})(P)^2} \partial_\varepsilon\big(\pi_s(\{a_j\})(P_\varepsilon)\big) M^{a,a_j}(\mathrm{d}s) \\
&\quad - \int_0^{\tau\wedge\tau^{g^*}} \sum_k \sum_{j=1}^K m^*_{s,k,j} \left(\frac{1}{\pi_s(\{a_j\})} - 1\right) \partial_\varepsilon(\Lambda^{a,a_j}(\mathrm{d}s)(P_\varepsilon)) + \partial_\varepsilon \mathbb{L}^a_t(P_\varepsilon) \\
&= -\int_0^{t\wedge\tau^{g^*}} \sum_k \sum_{j=1}^K m^*_{s,k,j} \frac{1}{\pi_s(\{a_j\})(P)^2} \partial_\varepsilon\big(\pi_s(\{a_j\})(P_\varepsilon)\big) M^{a,a_j}(\mathrm{d}s) \\
&\quad - \int_0^{\tau\wedge\tau^{g^*}} \sum_k \sum_{j=1}^K m^*_{s,k,j} \left(\frac{1}{\pi_s(\{a_j\})} - 1\right) \mathrm{d}\langle \Gamma, M^{a,a_j}\rangle^P_s + \langle \Gamma, \mathbb{L}^a\rangle^P_t \\
&= -\int_0^{t\wedge\tau^{g^*}} \sum_k \sum_{j=1}^K m^*_{s,k,j} \frac{1}{\pi_s(\{a_j\})(P)^2} \partial_\varepsilon\big(\pi_s(\{a_j\})(P_\varepsilon)\big) M^{a,a_j}(\mathrm{d}s) - \langle \Gamma, K^*\rangle^P_t,
\end{aligned}
$$

again using that $\partial_\varepsilon \Lambda_t(P_\varepsilon) = \langle \Gamma, M\rangle^P_t$ if $M = N - \Lambda(P)$ is a P-martingale. On the other hand also calculating it for Equation 6 gives

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

Also note that

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

Thus,

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

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

$$
\begin{aligned}
&\partial_\varepsilon \bar{K}_t^\varepsilon - \sum_{0<s\le t} \left(\partial_\varepsilon \Delta \bar{K}_s^\varepsilon\right) \frac{\Delta \bar{K}_s^0}{1+\Delta \bar{K}_s^0} \\
&= -\int_0^{t\wedge\tau^{g^*}} \sum_k \sum_{j=1}^K m_{s,k,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_t^P \\
&\quad + \int_0^{t\wedge\tau^{g^*}} \sum_k \sum_{j=1}^K m_{s,k,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_k \sum_{j=1}^K m_{s,k,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_t^P \\
&\quad - \int_0^{t\wedge\tau^{g^*}} \sum_k \sum_{j=1}^K m_{s,k,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_k \sum_{j=1}^K m_{s,k,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_t^P \qquad (8) \\
&\quad - \int_0^{t\wedge\tau^{g^*}} \sum_k \sum_{j=1}^K m_{s,k,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) - \sum_v \sum_i m_{s,v}^* \left(\frac{1}{\pi_s(\{a_i\})} - 1\right) \Lambda^{a,a_i}(\mathrm{d}s)\right) \\
&= -\langle \Gamma, K^* \rangle_t^P \\
&\quad - \int_0^{t\wedge\tau^{g^*}} \sum_k \sum_{j=1}^K m_{s,k,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) - \sum_v \sum_i \Bigg(m_{s,v}^* \left(\frac{1}{\pi_s(\{a_i\})} - 1\right) \\
&\quad - \left(\frac{0}{\pi_s(\{a_j\})} - 1\right) \mathbb{1}\left\{T_{(v-1)} < s \le T_{(v)}\right\} \mathbb{1}\left\{j \neq g_v^*\left(\mathcal{F}_{T_{(v-1)}}, s\right)\right\} \Lambda^{a,a_i}(\mathrm{d}s)\Bigg) \\
&= -\langle \Gamma, K^* \rangle_t^P \\
&\quad - \int_0^{t\wedge\tau^{g^*}} \sum_k \sum_{j=1}^K m_{s,k,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) - \mathrm{d}\langle M^{a,a_j}, K^* \rangle_s^P) \\
&= -\langle \Gamma, K^* \rangle_t^P \\
&\quad - \int_0^{t\wedge\tau^{g^*}} \sum_k \sum_{j=1}^K m_{s,k,j}^* \frac{\mathrm{d}\langle \Gamma, M^{a,a_j} \rangle_t^P - \pi_t(\{a_j\})(P)\, \mathrm{d}\langle \Gamma, M^a \rangle_t^P}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} (M^{a,a_j}(\mathrm{d}s) - \mathrm{d}\langle M^{a,a_j}, K^* \rangle_s^P)
\end{aligned}
$$

Note that $\frac{\mathrm{d}\langle \Gamma, M^{a,a_j} \rangle_t^P - \pi_t(\{a_j\})(P)\, \mathrm{d}\langle \Gamma, M^a \rangle_t^P}{\mathrm{d}\Lambda^{a,a_j}(t)(P)}$ can be be chosen predictable so that the corresponding term is a (local) martingale and that the last two terms in Equation 8 can be written as $\langle \Gamma, Z \rangle_t^P$ for some (local) martingale Z not specified (here). Conclude that the Score operator S is given by

$$
\begin{aligned}
&\mathcal{L}^2(P) \ni \Gamma \mapsto \Gamma - \langle \Gamma, K^* \rangle_t^P \\
&\qquad - \int_0^{t\wedge\tau^{g^*}} \sum_k \sum_{j=1}^K m_{s,k,j}^* \frac{\mathrm{d}\langle \Gamma, M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a \rangle_t^P}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} (M^{a,a_j}(\mathrm{d}s) - \mathrm{d}\langle M^{a,a_j}, K^* \rangle_s^P) \in \mathcal{L}^2(Q) \qquad (9)
\end{aligned}
$$

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

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

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

$$H_j\Gamma = \int_0^{t\wedge \tau^{g^*}} \sum_k m^*_{s,k,j} \frac{\mathrm{d}\langle \Gamma, M^{a,a_j}\rangle^P_t - \pi_t\big(\{a_j\}\big)(P)\,\mathrm{d}\langle \Gamma, M^a\rangle^P_t}{\mathrm{d}\Lambda^{a,a_j}(t)(P)}\big(M^{a,a_j}(\mathrm{d}s) - \mathrm{d}\langle M^{a,a_j}, K^*\rangle^P_s\big)$$

Then, we have that

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

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

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

Then, we may calculate directly that

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

Now note that

$$
\begin{aligned}
D^*\Gamma' &= \Gamma'_0 + [\Gamma', W]_t - \int_0^t W_{s-} \,\mathrm{d}\Gamma'_s \\
&= \Gamma'_0 + \int_0^t W_{s-} \,\mathrm{d}(\Gamma' + [\Gamma', K^*])_s \\
&= \Gamma'_t W_t - \int_0^t \Gamma'_{s-} \,\mathrm{d}W_s \\
&= \Gamma'_t W_t - \int_0^t \Gamma'_{s-} W_{s-} \,\mathrm{d}K^*_s
\end{aligned}
$$

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

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

FACT:

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

so it follows that

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

Note that this operator sends to piecewise constant functions. 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. We show that $\sum_j Y^*_j$ is a contraction; this then follows by a fixed-point theorem.

$$
\begin{aligned}
\langle\sum_j Y_j^*\Gamma - \sum_j Y_j^*\Gamma'\rangle_P &= \langle\sum_j Y_j^*(\Gamma-\Gamma')\rangle_P \\
&= \mathbb{E}_P\left[\langle\sum_j Y_j^*(\Gamma-\Gamma')\rangle_t^P\right] \\
&= \mathbb{E}_P\left[\int_0^t \mathbb{1}\{\tau^{g^*}\le s\}\sum_j\sum_k m_{s,k,j}^* \frac{\mathrm{d}\langle M^{a,a_j},\Gamma-\Gamma'\rangle_s^P}{\mathrm{d}\Lambda^{a,a_j}(s)(P)}\,\mathrm{d}\langle M^{a,a_j}-\pi_\cdot(\{a_j\})(P)\bullet M^a\rangle_s^P\right] \\
&= \mathbb{E}_P\left[\int_0^t \mathbb{1}\{\tau^{g^*}\le s\}\sum_j\sum_k m_{s,k,j}^* \frac{\mathrm{d}\langle M^{a,a_j}-\pi_\cdot(\{a_j\})(P)\bullet M^a\rangle_s^P}{\mathrm{d}\Lambda^{a,a_j}(s)(P)}\,\mathrm{d}\langle M^{a,a_j},\Gamma-\Gamma'\rangle_s^P\right] \\
&\le \mathbb{E}_P\left[\left(\left[\int_0^t \mathbb{1}\{\tau^{g^*}\le s\}\sum_j\sum_k m_{s,k,j}^* \left(\frac{\mathrm{d}\langle M^{a,a_j}-\pi_\cdot(\{a_j\})(P)\bullet M^a\rangle_s^P}{\mathrm{d}\Lambda^{a,a_j}(s)(P)}\right)^2 \mathrm{d}\langle M^{a,a_j}\rangle_s^P\right]\right)^{\frac{1}{2}}\right. \\
&\quad\left.\times\left(\left[\int_0^t \mathbb{1}\{\tau^{g^*}\le s\}\sum_j\sum_k m_{s,k,j}^* \left(\frac{\mathrm{d}\langle M^{a,a_j}-\pi_\cdot(\{a_j\})(P)\bullet M^a\rangle_s^P}{\mathrm{d}\Lambda^{a,a_j}(s)(P)}\right)^2 \mathrm{d}\langle \Gamma-\Gamma'\rangle_s^P\right]\right)^{\frac{1}{2}}\right]
\end{aligned}
$$

Kunita Watanabe inequality or Cauchy-Schwarz like inequality gives??? Second integrand can b ebounded by something ≤ 1? What to do about square root.

Can this contraction be zero ???? in the absolutely continuous case. No, that's not how bilinearity works. First, let Γ be MG by MG representation theorem so that

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

for some predictable $h_x(s)$. Then,

$$
\begin{aligned}
\langle\Gamma, M^{a,a_j}-\pi_\cdot(\{a_j\})(P)\bullet M^a\rangle_t^P &= \int_0^t \sum_x h_x(s)\,\mathrm{d}\langle M^x, M^{a,a_j}-\pi_\cdot(\{a_j\})(P)\bullet M^a\rangle_s^P \\
&= \int_0^t \sum_x h_x(s)\,\mathrm{d}\langle M^x, M^{a,a_j}\rangle_s^P - \int_0^t \pi_s(\{a_j\})(P)\left(\sum_x h_x(s)\right)\mathrm{d}\langle M^x, M^a\rangle_s^P
\end{aligned}
$$

Have:

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

where $\Lambda^{*,x}$ 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 $\Lambda^{*,x}=0$. 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

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

Now calculate

$$\begin{aligned}
SS^*\Gamma &= \left(D - \sum_j DY_j\right)\left(D^*\Gamma - \sum_j Y_j^* D^*\Gamma\right) \\
&= DD^*\Gamma - D\sum_j Y_j^* D^*\Gamma - \sum_j DY_j D^*\Gamma + \left(\sum_j DY_j\right)\sum_i Y_i^* D^*\Gamma \\
&= \Gamma_0 + \int_0^\cdot \bar{W}_s \,\mathrm{d}\Gamma_s - D\sum_j Y_j^* D^*\Gamma - \left(\sum_j DY_j\right) D^*\Gamma + \left(\sum_j DY_j\right)\left(\sum_i Y_i^* D^*\Gamma\right)
\end{aligned}$$

Note that

$$\begin{aligned}
\left(\sum_j DY_j\right)\left(\sum_i Y_i^* D^*\Gamma\right) &= \sum_j \sum_i DY_j Y_i^* D^*\Gamma \\
&= \sum_j \sum_i \int_0^{\cdot \wedge \tau^{g^*}} \sum_k m_{s,k,j}^* \frac{\mathrm{d}\langle Y_i^* D^*\Gamma, M^{a,a_j} - \pi_s\big(\{a_j\}\big)(P) \bullet M^a\rangle_s^P}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} \\
&\times \mathrm{d}\big(M^{a,a_j} - \mathrm{d}\langle M^{a,a_j}, K^*\rangle_s^P\big) \\
&= \sum_j \sum_i \int_0^{\cdot \wedge \tau^{g^*}} \sum_k m_{s,k,j}^* \\
&\times \frac{\mathbb{1}\{\tau^{g^*} \le s\} \sum_k m_{s,k,i}^* W_{s-} \frac{\mathrm{d}\langle M^{a,a_j}, \Gamma\rangle_s^Q}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \,\mathrm{d}\langle M^{a,a_j} - \pi_\cdot\big(\{a_j\}\big)(P) \bullet M^a, M^{a,a_i} - \pi_\cdot(\{a_i\})(P) \bullet M^a\rangle_s}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} \\
&= \sum_j \int_0^{\cdot \wedge \tau^{g^*}} W_{s-} \sum_k m_{s,k,j}^* \frac{\mathrm{d}\langle M^{a,a_j}, \Gamma\rangle_s^Q}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \frac{\mathrm{d}\langle M^{a,a_j} - \pi_\cdot\big(\{a_j\}\big)(P) \bullet M^a\rangle_s}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} \\
&\times \mathrm{d}\big(M^{a,a_j} - \mathrm{d}\langle M^{a,a_j}, K^*\rangle_s^P\big)
\end{aligned}$$

and

$$\begin{aligned}
D\sum_j Y_j^* D^*\Gamma &= D\sum_j \int_0^{\cdot \wedge \tau^{g^*}} \mathbb{1}\{\tau^{g^*} \le s\} \sum_k m_{s,k,j}^* W_{s-} \frac{\mathrm{d}\langle M^{a,a_j}, \Gamma\rangle_s^Q}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \,\mathrm{d}\big(M^{a,a_j} - \pi_\cdot\big(\{a_j\}\big)(P) \bullet M^a\big)_s \\
&= \sum_j \int_0^{\cdot \wedge \tau^{g^*}} \mathbb{1}\{\tau^{g^*} \le s\} \sum_k m_{s,k,j}^* W_{s-} \frac{\mathrm{d}\langle M^{a,a_j}, \Gamma\rangle_s^Q}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \,\mathrm{d}\big(D\big(M^{a,a_j} - \pi_\cdot\big(\{a_j\}\big)(P) \bullet M^a\big)\big)_s \\
&= \sum_j \int_0^{\cdot \wedge \tau^{g^*}} \mathbb{1}\{\tau^{g^*} \le s\} \sum_k m_{s,k,j}^* W_{s-} \\
&\times \frac{\mathrm{d}\langle M^{a,a_j}, \Gamma\rangle_s^Q}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \,\mathrm{d}\big(M^{a,a_j} - \pi_\cdot\big(\{a_j\}\big)(P) \bullet M^a - \langle M^{a,a_j} - \pi_\cdot\big(\{a_j\}\big)(P) \bullet M^a, K^*\rangle^P\big)_s
\end{aligned}$$

Finally, note that

$$\sum_j DY_j D^* \Gamma = \int_0^{t \wedge \tau^{g^*}} W_{s-} \sum_k \sum_{j=1}^K m^*_{s,k,j} \frac{\mathrm{d}\langle \Gamma + [\Gamma, K^*], M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a \rangle_t^P}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} \left(M^{a,a_j}(\mathrm{d}s) - \mathrm{d}\langle M^{a,a_j}, K^* \rangle_s^P\right)$$

$$\int_0^{t \wedge \tau^{g^*}} W_{s-} \sum_k \sum_{j=1}^K m^*_{s,k,j} \frac{\mathrm{d}\langle \Gamma, M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a \rangle_t^Q}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} \left(M^{a,a_j}(\mathrm{d}s) - \mathrm{d}\langle M^{a,a_j}, K^* \rangle_s^P\right)$$

Thus, we find

$$\begin{aligned}
S^* S \Gamma = & \sum_j \int_0^{\cdot \wedge \tau^{g^*}} W_{s-} \sum_k m^*_{s,k,j} \\
& \times \left(\frac{\mathrm{d}\langle M^{a,a_j}, \Gamma \rangle_s^Q}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \left(\frac{\mathrm{d}\langle M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a \rangle_s}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} - 1 \right) - \frac{\mathrm{d}\langle \Gamma, M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a \rangle_t^Q}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} \right) \\
& \times \mathrm{d}\left(M^{a,a_j} - \mathrm{d}\langle M^{a,a_j}, K^* \rangle_s^P\right) \\
& + \sum_j \int_0^{\cdot \wedge \tau^{g^*}} \mathbb{1}\{\tau^{g^*} \leq s\} \sum_k m^*_{s,k,j} W_{s-} \pi_s(\{a_j\})(P) \frac{\mathrm{d}\langle M^{a,a_j}, \Gamma \rangle_s^Q}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \, \mathrm{d}\left(M^a - \langle M^a, K^* \rangle^P\right)_s \\
& + \Gamma_0 + \int_0^\cdot \bar{W}_s \, \mathrm{d}\Gamma_s \\
& = \sum_j \int_0^{\cdot \wedge \tau^{g^*}} W_{s-} \sum_k m^*_{s,k,j} \\
& \times \left(\frac{\mathrm{d}\langle M^{a,a_j}, \Gamma \rangle_s^Q}{\mathrm{d}\Lambda^{a,a_j}(s)(P)} \left(\frac{\mathrm{d}\langle M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a \rangle_s^P}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} - 1 \right) - \frac{\mathrm{d}\langle \Gamma, M^{a,a_j} - \pi_\cdot(\{a_j\})(P) \bullet M^a \rangle_t^Q}{\mathrm{d}\Lambda^{a,a_j}(t)(P)} \right) \\
& \times \mathrm{d}\left(M^{a,a_j} - \mathrm{d}\langle M^{a,a_j}, K^* \rangle_s^P\right) \\
& + \sum_j \int_0^{\cdot \wedge \tau^{g^*}} \mathbb{1}\{\tau^{g^*} \leq s\} \sum_k m^*_{s,k,j} W_{s-} \frac{\mathrm{d}\langle M^{a,a_j}, \Gamma \rangle_s^Q}{\mathrm{d}\Lambda^{a}(s)(P)} \, \mathrm{d}\left(M^a - \langle M^a, K^* \rangle^P\right)_s \\
& + \Gamma_0 + \int_0^\cdot \bar{W}_s \, \mathrm{d}\Gamma_s
\end{aligned}$$

5) Sequential representation of exchangeabillity

Theorem 5.1: 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

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

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,

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

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\left[W_{T_{(m)}} \mathbb{1}\left\{T_{(m)} < t\right\} N^{a,a_j}\left(T_{(m+1)} \wedge t\right)\right. \\
&\quad \left. \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_{(m+1)} = a\right]\right] \\
&= \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)\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_{(m+1)} = a\right] \\
&\quad \left. \times \mathbb{E}_P\left[\tilde{Y}_t \mid \mathcal{F}_{T_{(m)}}, T_{(m+1)} \leq \tau, \Delta_{(m+1)} = a\right]\right] \\
&= \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_{(m+1)} = a\right]\right. \\
&\quad \left. \times \mathbb{E}_P\left[\tilde{Y}_t \mid \mathcal{F}_{T_{(m)}}, T_{(m+1)} \leq \tau, \Delta_{(m+1)} = a\right]\right] \\
&= \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_{(m+1)} = a\right]\right] \\
&= 0.
\end{aligned}
$$

□

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

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

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