A fundamental condition in our asymptotics Theorem 2.4 for this estimator μ_n^0 of μ is that $D_h(X \mid \mu(F_X), \rho(F_X)) \in \mathcal{D}(\rho_1(F_X), G)$. This assumption can be a model assumption (e.g., one assumes that $P(X observed \mid X) > \delta > 0$, F_X-a.e., for some $\delta > 0$), but, if it is not a reasonable assumption for the particular application of interest, then it can be weakened by making this membership requirement an additional parameter of the full data estimating function in the manner described in Subsection 2.3.1 of Chapter 2, assuming that the set $\mathcal{D}(\rho_1, G)$ is not empty. Thus, one 1) identifies a subset $\mathcal{H}^F(\mu, \rho, \rho_1, G) \subset \mathcal{H}^F$ so that for all $h \in \mathcal{H}^F(\mu, \rho, \rho_1, G)$ $E_G(IC_0(Y \mid G, D_h(\cdot \mid \mu, \rho)) \mid X) = D_h(X \mid \mu, \rho)$ F_X-a.e. (for all possible μ, ρ, G) and 2) reparametrizes this restricted class of full data structure estimating functions $\{D_h : h \in \mathcal{H}^F(\mu, \rho, \rho_1, G)\}$ as $\{D_h^r : h \in \mathcal{H}^F\}$ by incorporating the extra nuisance parameter ρ_1, G (needed to map any $h \in \mathcal{H}^F$ into $\mathcal{H}^F(\mu, \rho, \rho_1, G)$) in the nuisance parameter of D_h^r. This reparametrization is specified in (2.14). Subsequently, we denote this reparametrized class of estimating functions with $\{D_h(\cdot \mid \mu, \rho) : h \in \mathcal{H}^F\}$ again, where ρ now includes the old ρ, ρ_1, and G. We showed this reparametrization in action in a right-censored data example covered in Examples 2.3, 2.4, and 2.5 in Chapter 2. We refer the reader to this worked-out example.

In the next subsection, we propose IPCW estimators of the regression parameters in a multiplicative intensity model and derive the confidence intervals by establishing a fundamental lemma establishing the projection on the tangent space of the censoring mechanism G under the Cox proportional hazards model $\mathcal{G}$. Subsequently, we complete this estimation problem by proposing locally efficient estimators according to our general methodology, as covered in the next section. We also show the extension of the estimators to the proportional rate models.

3.3.2 *Estimation of a marginal multiplicative intensity model*

Let $X(t) = (N(t), V_1(t), V_2(t))$ be a time-dependent process containing a counting process $N(t)$ and time-dependent covariates $V(t) = (V_1(t), V_2(t))$, where $V_1(t)$ represents covariates for which one wants to adjust in the multiplicative intensity model below. Let $Z(t) = (N(t), V_1(t))$. Suppose that for the full data of interest we have $\bar{X}(T)$ for some time variable (random or fixed) T. If T is random, then it is assumed that $R(t) = I(T \le t)$ is a component of $Z(t)$. As the full data model, we consider a multiplicative intensity model

$$E(dN(t) \mid \bar{Z}(t-)) = Y(t)\lambda_0(t)\exp(\beta W(t)), \tag{3.8}$$

where $Y(t)$ and $W(t)$ are functions of $\bar{Z}(t-)$, and $Y(t)$ is the indicator that $N(t)$ is at risk of jumping at time $t-$.

Suppose that for the observed data we have $Y = (C, \bar{X}(C))$ (recall that

fined as $(C, \bar{X}(C))$ by defining $C = \infty$ if $T < C$). Our parameter of interest is $\mu = \beta$.

This estimation problem covers a large class of important problems. For example, consider a trial in which subjects are randomized to a treatment arm Tr (e.g., Tr is a 1-0 variable), and suppose that the goal of this trial is to estimate the causal effect of treatment on the survival time T. In this case, let $N(t) = I(T \le t)$ and $Z(t) = (N(t), Tr)$ so that

$$E(dN(t) \mid \bar{Z}(t-)) = I(T \ge t)\lambda(t \mid Tr)dt,$$

where $\lambda(t \mid Tr)dt = P(T \in dt \mid T \ge t, Tr)$ is the hazard of failure within treatment arm Tr. Assuming a multiplicative intensity model, in this case known as the Cox proportional hazards model (Andersen and Gill, 1982; Gill, 1984; Ritov and Wellner, 1988), corresponds with

$$\lambda(t \mid Tr) = \lambda_0(t)\exp(\beta Tr).$$

The (causal) parameter of interest is now the regression coefficient β in front of Tr. An ad hoc method for estimation of β would be to fit a Cox proportional hazards model for right-censored data on T ignoring the covariates beyond Tr. However, if C is not independent of T, given Tr, then this estimator will be inconsistent. For example, in this clinical trial, people might drop out of the treatment arm because of possible side effects or other severe complications measured by $X(t)$. In addition, this ad hoc method will be very inefficient, even when C is known to be independent of T, given Tr. In other words, there is a real need for a general methodology for estimation of a Cox proportional hazards model that does not adjust for all measured covariates of interest. Another example would be obtained by letting $N(t)$ be a counting process that repeatedly jumps. For example, $N(t)$ might jump each time a particular type of patient is admitted into the hospital. We refer to Pavlic, van der Laan, and Buttler (2001) for an application of the methods described in this subsection to estimate the intensity and rates of the lung exacerbations in cystic fibrosis patients.

Firstly, we need to know the orthogonal complement of the nuisance tangent space in the full data model in order to determine the class of full data estimating functions. Subsequently, we will map these into observed data estimating functions with our general mapping $IC_0(\cdot \mid G, D) - \Pi(IC_0(\cdot \mid G, D) \mid T_{CAR})$ for which we have a closed-form representation (Chapter 1, Theorem 1.1, and the next Section 3.4). Finally, we propose estimators of the nuisance parameters of this observed data estimating function for β.

In Lemma 2.2 in Chapter 2, we showed that all full data estimating functions are given by

$$\left\{D_h(\cdot \mid \mu, \Lambda_0) = \int \{h(t, \bar{Z}(t-)) - g_\mu(h)\} dM_{\mu,\Lambda_0}(t) : h\right\}, \tag{3.9}$$

where $dM(t) = dN(t) - E(dN(t) \mid \bar{Z}(t-))$ and

$g_\mu(h)(t) = E(h(t, \bar{Z}(t-))Y(t)\exp(\beta W(t)))/E(Y(t)\exp(\beta W(t)))$.