now includes the old ρ, ρ_1, and G.
$\mathbf{IC(Y \mid Q, G, D_h(\cdot \mid \mu, \rho))}$: an observed data estimating function for μ with nuisance parameters $Q(F_X, G), G$, and ρ, which is obtained by applying the mapping $D \rightarrow IC(Y \mid Q, G, D)$ to the particular full data estimating function D_h. If $h = (h_1, \ldots, h_k) \in \mathcal{H}^{Fk}$, then $IC(Y \mid Q, G, D_h(\cdot \mid \mu, \rho))$ denotes

$$(IC(Y \mid Q, G, D_{h_1}(\cdot \mid \mu, \rho)), \ldots, IC(Y \mid Q, G, D_{h_k}(\cdot \mid \mu, \rho))).$$

$\mathbf{S^*_{eff}(Y \mid F_X, G)}$: the canonical gradient (also called the efficient influence curve) of the pathwise derivative of the parameter μ in the observed data model $\mathcal{M}$.
$\mathbf{IC(Y \mid F_X, G, D_{h_{opt}}(\cdot \mid \mu(F_X), \rho(F_X, G)))} = S^*_{eff}(Y \mid F_X, G)$: that is, h_{opt} indexes the choice of full data estimating function that results in the optimal observed data estimating function for μ in the *observed data model* $\mathcal{M}$, $\mathcal{M}(\mathcal{G})$, and $\mathcal{M}(CAR)$. Here $h_{opt} = h_{opt}(F_X, G)$ depends on F_X and G.
$\mathbf{h_{ind,F_X}} : L^2_0(F_X) \rightarrow \mathcal{H}^F$: We call it the *index mapping* since it maps a full data function into an index h defining the projection onto $T^{F,\perp}_{nuis}(F_X)$. It is defined by

$$D_{h_{ind,F_X}(D)}(X \mid \mu(F_X), \rho(F_X)) = \Pi(D \mid T^{F,\perp}_{nuis}(F_X)).$$

$\mathbf{A_{F_X}} : L^2_0(F_X) \rightarrow L^2_0(P_{F_X,G}) : A_{F_X}(h)(Y) = E_{F_X}(h(X) \mid Y)$: the nonparametric score operator that maps a score of a one-dimensional fluctuation F_ϵ at F_X into the score of the corresponding one-dimensional fluctuation $P_{F_\epsilon,G}$ at $P_{F_X,G}$.
$\mathbf{A_G} : L^2_0(P_{F_X,G}) \rightarrow L^2_0(F_X)$: $A_G(V)(X) = E_G(V(Y) \mid X)$: the adjoint of the nonparametric score operator A_{F_X}.
$\mathbf{I_{F_X,G}} = A_G^\top A_{F_X}: L^2_0(F_X) \rightarrow L^2_0(F_X)$: $\mathbf{I}_{F_X,G}(h)(X) = E_G(E_{F_X}(h(X) \mid Y) \mid X)$: the nonparametric information operator. If we write $\mathbf{I}^{-1}_{F_X,G}(h)$, then it is implicitly assumed that $I_{F_X,G}$ is 1-1 and h lies in the range of $I_{F_X,G}$.
$\mathbf{IC(Y \mid F_X, G, D)} \equiv A_{F_X}\mathbf{I}^-_{F_X,G}(D)$ is an optimal mapping (assuming that the generalized inverse is defined) from full data estimating functions into the observed data estimating function. For any $IC_0(Y \mid F_X, G, D)$ satisfying $E(IC_0(Y \mid F_X, G, D) \mid X) = D(X)$ F_X-a.e., the optimal mapping can be more generally defined by $IC(Y \mid F_X, G, D) = IC_0(Y \mid F_X, G, D) - \Pi(IC_0(Y \mid F_X, G, D) \mid T_{CAR}(P_{F_X,G}))$.
$\mathbf{I^*_{F_X,G}} = \Pi(\mathbf{I}_{F_X,G} \mid T^F(F_X))$: the information operator. If we write $I^{*-1}_{F_X,G}(h)$, then it is implicitly assumed that $\mathbf{I}^*_{F_X,G}$ is 1-1 and h lies in the range of $I^*_{F_X,G}$.
The projection operator can be expressed as a sum of a projection on a finite space and the projection on $T^{F,\perp}_{nuis}$ since $T^F(F_X) = \langle S_{eff}(\cdot \mid F_X)\rangle \oplus T^F_{nuis}(F_X)$.
$S^*_{eff}(Y \mid F_X, G) = A_{F_X}\mathbf{I}^{*-}_{F_X,G}(S^{*F}_{eff}(\cdot \mid F_X))$: that is, the efficient influence curve can be expressed in terms of the inverse of the information operator, assuming that the inverse is defined.

$\mathbf{R(B)}$: the range of a previously defined linear operator $\mathbf{B}$.
$\mathbf{N(B)}$: the null space of a previously defined linear operator $\mathbf{B}$.
$\overline{\mathbf{H}}$ for a set of elements in a Hilbert space $L^2_0(P_{F_X,G})$ is defined as the closure of its linear span.
$\mathbf{Pf} \equiv \int f(y)dP(y)$.
$\mathcal{L}(\mathcal{X})$: all real-valued functions of X that are uniformly bounded on a set that contains the true X with probability one.
$\mathcal{H}$: some index set indexing observed data estimating functions.
$\mathbf{c}(\mu) = d/d\mu EIC(Y \mid Q, G, D_h(\cdot \mid \mu, \rho))$ $(h = (h_1, \ldots, h_k))$: the derivative matrix of the expected value of the observed data estimating function.