Comparison of meta-learners for estimating multi-valued treatment heterogeneous effects

Naoufal Acharki [1][2] **Ramiro Lugo** [2] **Antoine Bertoncello** [2] **Josselin Garnier** [1]

Abstract

Conditional Average Treatment Effects (CATE) estimation is one of the main challenges in causal inference with observational data. In addition to Machine Learning based-models, nonparametric estimators called *meta-learners* have been developed to estimate the CATE with the main advantage of not restraining the estimation to a specific supervised learning method. This task becomes, however, more complicated when the treatment is not binary as some limitations of the naive extensions emerge. This paper looks into meta-learners for estimating the heterogeneous effects of multi-valued treatments. We consider different meta-learners, and we carry out a theoretical analysis of their error upper bounds as functions of important parameters such as the number of treatment levels, showing that the naive extensions do not always provide satisfactory results. We introduce and discuss meta-learners that perform well as the number of treatments increases. We empirically confirm the strengths and weaknesses of those methods with synthetic and semi-synthetic datasets.

1. Introduction

With the rapid development of Machine Learning (ML) and its efficiency in predicting outcomes, the question of counterfactual prediction *"what would happen if ?"* arises. Engineers and specialists want to know how the outcome would be affected after an intervention on a parameter. It will help them personalize the parameter at efficient levels and optimize the outcome. Recently, much effort has been devoted to supervised ML to find the optimal intervention strategy. Yet, the results are not always convincing. These models cannot distinguish between correlations and causal relationships in the data (Pearl, 2019).

Based on the Potential Outcomes theory (Neyman, 1923; Rubin, 1974), epidemiologists and statisticians have developed a set of tools that reduce the inference of causal effects to statistical inference under certain assumptions about, for example, the data-generating process. They have been successfully applied in many fields such as medicine (Alaa & van der Schaar, 2017), economics (Knaus et al., 2020b), public policy (Imai & Strauss, 2011) and marketing (Diemert et al., 2018) to infer causal effects. However, most of these studies are limited to a binary treatment setting whereas many causal questions in real-world cases are not binary. It would be helpful to give an in-depth analysis of the impact of the treatment across its possible levels instead of just considering a binary scenario where the treatment is either assigned or not. In addition, the heterogeneity of effects may provide valuable information regarding this treatment's effectiveness and help users personalize their intervention policies and strategies.

The problem of Causal Inference beyond binary treatment settings is gaining attention from the Causal Inference community. There are two major challenges: Firstly, the lack of the ground truth due to the fundamental problem of causal inference (Holland, 1986) makes Heterogeneous Treatment Effects estimation more challenging (Alaa & van der Schaar, 2018) as standard metrics can not be used to assess performances. Secondly, binarizing the multi-valued treatment setting leads to a violation of the Stable Unit Treatment Value Assumption (SUTVA) as it violates the principle of "no hidden variations of treatment". It may yield a bias known as position bias in recommendation systems (Chen et al., 2023; Wu et al., 2022). It happens when some units tend to have/select high treatment values and, therefore, there is a hidden variation of the treatment that is not taken into consideration when one attempt to binarize the treatment. From a statistical point of view, this bias was established by Heiler & Knaus (2022) who show that the binarization of multi-valued treatments does not disassociate the heterogeneity of the treatment from the heterogeneity of the effects of each value.

The extension of the binary setting does not seem trivial as several versions are possible and turn out to have different

[1]CMAP, Ecole polytechnique, Institut Polytechnique de Paris, Palaiseau, France [2]TotalEnergies One Tech, Palaiseau, France. Correspondence to: Naoufal Acharki <naoufal.acharki@gmail.com>.

behaviours. Moreover, – to the best of our knowledge – there is no result so far on the impact of the number of possible treatments on the performances of heterogeneous treatment effect (nonparametric or ML-based) estimators.

In this paper, we study the problem of nonparametric estimation of Heterogeneous Treatment Effects, also known as Conditional Average Treatment Effects (CATEs), for multi-valued treatments. We consider nonparametric estimators, also referred to as *meta-learners* (Künzel et al., 2019) or *model-agnostic algorithms* (Curth & van der Schaar, 2021a). We put our main focus on discussing the theoretical properties of meta-learners for estimating CATEs. Finally, along the lines of Curth & van der Schaar (2021a), we consider it significant to draw strengths and weaknesses theoretically and compare scenarios in which some methods would perform better than others.

Contributions. The paper considers meta-learners for multi-valued treatments. First, we generalize *meta-learners* to the multi-treatment setting for CATE estimation. We overview, in particular, Debiased Machine Learning (DML) estimators in observational studies and we establish a new version for the X-learner based on regression adjustment for multi-valued treatments. We also highlight the multi-treatments R-learner's main drawbacks. Second, and this is the major contribution of the paper, we conduct a theoretical analysis of the proposed meta-learners for multiple treatments based on an asymptotic bias-variance analysis (see Silverman (1986) for an example of this analysis for kernel density estimation). We analyze the biases and upper bounds on errors of the M-, DR-, X-learners as well as the T- and naive X-learners. Thanks to this analysis, we can identify the effect of the number of possible treatment levels, in addition to other parameters such as the propensity score lower bound and the outcome model estimation error. This approach is different from what has been conducted in binary treatment with the minimax approach (Künzel et al., 2019; Curth & van der Schaar, 2021a; Kennedy, 2020) as it allows a direct analysis of the roles of the important parameters (e.g. the impact of the number of possible treatments K) instead of relying on the smoothness of treatment effects. Following this analysis, we present some key points about the expected performances of each meta-learner then we present a summarizing table of our findings. We note also that our analysis sheds new light on binary meta-learners' performances as it also clarifies the influence of the sampling probability for both T- and naive X-learners.

2. Related work

Meta-learners for CATEs estimation. The recent interest in the CATE's estimation has motivated the Causal Inference community to develop numerous algorithms and methods (see Caron et al. (2022b) for a review). This includes a wide variety of statistical and ensemble methods (Hill, 2011; Athey & Imbens, 2016; Alaa & van der Schaar, 2017; Wager & Athey, 2018; Powers et al., 2018; Hahn et al., 2020; Caron et al., 2022a) as well as neural networks (Johansson et al., 2016; Shalit et al., 2017; Yoon et al., 2018; Shi et al., 2019) (see Dorie et al. (2019) for a review of -hybrid- ML models for causal inference). In contrast, some methods, known as *meta-learners*, are nonparametric and do not require a specific ML method. The theory of *meta-learners* was initially introduced and discussed by Künzel et al. (2019) for the CATE estimation in the binary setting with three meta-learners: the S-learner, the T-learner (which use either a *Single* or *Two* models) and the X-learner. Later, Kennedy (2020) proposes the DR-learner (Doubly-Robust) to overcome the problem of model misspecifications when estimating nuisance functions (e.g. the propensity score and outcome models). Nie & Wager (2020) present the R-learner that estimates the CATE by minimizing an orthogonalized loss function. Curth & van der Schaar (2021a) consider the PW-learner (Propensity Weighting, also known as the M-learner) and the RA-learner (Regression-Adjustment), which is an improved version of the X-learner. They show that, under some conditions, the RA-, PW-, and DR-learners can attain the oracle convergence rate.

Multiple and continuous treatments. Recently, there has also been an increased interest in causal inference with multi-valued and continuous treatments. The theoretical work of Imbens (2000); Lechner (2001); Frölich (2002); Imai & Dyk (2004) extended the potential outcome framework and the propensity score to the non-binary treatment setting, including also continuous treatments. The average dose-response estimation was considered and successfully applied in many domains in medicine and economics (Dominici et al., 2002; Flores, 2007; Kallus, 2017; Saini et al., 2019; Lin et al., 2019; Hu et al., 2020; Knaus, 2022). Additionally, Colangelo & Lee (2020) apply doubly debiased machine learning methods to dose-response modelling with continuous treatment. The CATE's estimation, however, remains less prominently studied in the literature. Hill (2011) (briefly) and Hu et al. (2020) consider Bayesian additive regression trees for the estimation of counterfactual response and causal effects. An extension of GRF to multi-valued treatments is developed by Tibshirani et al. (2020) (which can be seen as an M-learner with multi-valued treatments). Schwab et al. (2020); Harada & Kashima (2021); Nie et al. (2021) and Kaddour et al. (2021) applied neural networks and representations learning to estimate counterfactual response curves for multiple continuous treatments (more precisely for graph-structured treatments). Zhao & Harinen (2019) naively extended binary meta-learners (X- and R-learners) without any theoretical analysis of their behaviour.

The remainder of the paper is structured as follows. Section

[3](#) presents the CATE estimation for multi-treatments. In Section [4](#) we introduce CATE nonparametric estimators (meta-learners) and we discuss their consistency. In Section [5](#), we establish the theoretical analysis of meta-learners' error bounds and provide some discussions. We present numerical experiments and results in Section [6](#). Finally, we present our conclusion in Section [7](#).

3. Problem setting

To address the problem of causal inference under multiple treatments, we follow the Rubin-Neyman model as extended by Imbens (2000); Lechner (2001); Imai & Dyk (2004) and we consider the following statistical problem.

We suppose the existence of $Y(t)$, the real-valued counterfactual outcome that would have been observed under treatment level $t \in \mathcal{T} = \{t_0, t_1, \ldots, t_K\}$. We consider $(\boldsymbol{X}, T, Y(t)_{t \in \mathcal{T}}) \sim \mathbb{P}$ where $\boldsymbol{X} = (X^{(1)}, \ldots, X^{(d)}) \in \mathcal{D} \subseteq \mathbb{R}^d$ denotes a random vector of covariates and T denotes the treatment assignment random variable. We suppose finally that we observe data that has the form of an independent and identically distributed sample of n units $\mathbf{D}_{\mathrm{obs}} = (D_{\mathrm{obs},i})_{i=1}^n$ where $D_{\mathrm{obs},i} = (\boldsymbol{X}_i, T_i, Y_{\mathrm{obs},i})$ is distributed as $(\boldsymbol{X}, T, Y_{\mathrm{obs}})$ and $Y_{\mathrm{obs}} = Y(T)$ (consistency assumption). We define the Generalized Propensity Score (GPS) $r(t, \boldsymbol{x}) := \mathbb{P}(T = t | \boldsymbol{X} = \boldsymbol{x})$ (Imbens, 2000) as the generalization of the classical Propensity Score with the same balancing propriety (Rosenbaum & Rubin, 1983) to remove *selection bias* in observational studies.

We aim to infer the effect of the treatment T on the outcome Y. More precisely, we want to estimate the CATE between two levels of treatment t_k and t_0, defined as

$$\tau_k(\boldsymbol{x}) = \mathbb{E}[Y(t_k) - Y(t_0) | \boldsymbol{X} = \boldsymbol{x}], \qquad (1)$$

which can be interpreted as the expected treatment effect between levels t_0 (defined as the baseline treatment value) and t_k given covariates $\boldsymbol{X} = \boldsymbol{x}$. Note that other definitions and alternatives of the CATE are possible (Kaddour et al., 2021).

Unfortunately, it is impossible to infer this effect directly. We observe only one potential outcome corresponding to the treatment T (i.e. the real outcome) for every unit. All other potential outcomes are missing (inherently unobservable). Consequently, to identify the causal effects from the observed sample data, we shall consider the following assumptions (Assumption 3.1 is unfortunately untestable).

Assumption 3.1 (Unconfoundedness)**.** The treatment mechanism is unconfounded given the observed covariates $Y(t) \perp\!\!\!\perp \mathbf{1}\{T = t\} \mid \boldsymbol{X}$ for all $t \in \mathcal{T}$.

Assumption 3.2 (Overlap)**.** The probability of receiving the treatment given the observed covariates is positive i.e. there

exists $r_{\min} > 0$ such that $r_{\min} \leq \mathbb{P}(T = t | \boldsymbol{X} = \boldsymbol{x})$ for all $t \in \mathcal{T}$ and $\boldsymbol{x} \in \mathcal{D}$.

With the previous assumptions, the expected potential outcome satisfies $\mathbb{E}(Y(t) \mid \boldsymbol{X} = \boldsymbol{x}) = \mathbb{E}(Y_{\mathrm{obs}} \mid T = t, \boldsymbol{X} = \boldsymbol{x})$ and the CATE can be estimated.

The problem of the CATE estimation can be seen as a nonparametric estimation problem. We tackle it by generalizing the notion of meta-learners to derive consistent estimators. This task can be achieved either by modelling the CATE directly in one or two steps: by decomposing it into regularized regression problems or by addressing a minimization problem with respect to an appropriate loss function. Moreover, all considered meta-learners below, except the R-learner, have the advantage of supporting any supervised regression method (e.g. random forest, gradient boosting methods, neural networks).

4. Generalizing meta-learners to multi-valued treatments

In the following, we adopt a similar taxonomy of CATEs estimators as Curth & van der Schaar (2021a) and Knaus et al. (2020a). Namely, direct plug-in (one-step) meta-learners, pseudo-outcome (two-step) meta-learners and Neyman-Orthogonality-based learners (R-learner).

4.1. Direct plug-in meta-learners

This subsection presents direct plug-in meta-learners, also known as *one-step* learners. They estimate the CATE in (1) by targeting it directly from $\mathbf{D}_{\mathrm{obs}}$. They are the naive extension of the T- and S-learners in the binary case.

T-learner with multiple treatments. T-learner is a naive approach to estimating CATEs. It consists on estimating the *two* conditional response surfaces $\mu_t(\boldsymbol{x}) = \mathbb{E}(Y(t) \mid \boldsymbol{X} = \boldsymbol{x})$ using $\mathbf{S}_t = \{i, T_i = t\}$ for $t \in \{t_k, t_0\}$, then estimates the CATE as $\widehat{\tau}_k^{(\mathrm{T})}(\boldsymbol{x}) := \widehat{\mu}_{t_k}(\boldsymbol{x}) - \widehat{\mu}_{t_0}(\boldsymbol{x})$.

The T-learning approach does not account for the interaction between T and Y and creates different models for different treatments. Still, it may suffer from selection bias (Curth & van der Schaar, 2021b), i.e. when the outcome models μ_t are estimated with respect to the wrong distribution when sampling $(D_{\mathrm{obs},i})_{i \in \mathbf{S}_t}$. Therefore, $\widehat{\mu}_t$ should be estimated by minimizing the expected squared error on the nominal *weighted* distribution using Importance Sampling (Hassanpour & Greiner, 2019); see Appendix A for details.

S-learner with multiple treatments. The S-learner in multi-valued treatments uses the identification of the CATE and considers a *single* model $\mu(t, \boldsymbol{x}) = \mathbb{E}(Y_{\mathrm{obs}} \mid T = t, \boldsymbol{X} = \boldsymbol{x})$. μ is estimated using the whole dataset

$\mathbf{D}_{\mathrm{obs}}$. The CATE can be computed therefore as $\widehat{\tau}_k^{(\mathrm{S})}(\boldsymbol{x}) := \widehat{\mu}(\boldsymbol{x}, t_k) - \widehat{\mu}(\boldsymbol{x}, t_0)$.

Including the treatment T as a feature and sharing information between covariates $\boldsymbol{X}$ and T may provide better predictions. However, this advantage is conditioned by the ability of the base learner to capture and distinguish contributions of both $\boldsymbol{X}$ and T on Y_{obs} as we will see in Section 5. Note that the S-learner may also suffer from confounding and regularization biases (Chernozhukov et al., 2018; Hahn et al., 2020) when estimating $\widehat{\mu}$.

4.2. Pseudo-outcome meta-learners

To overcome *selection bias*, a usual alternative is to consider specific representations of the observed outcome Y_{obs}, called *pseudo-outcome* Z_k. They incorporate *nuisance components* that generally include valuable information such as the dependence between covariates $\boldsymbol{X}$ and T (i.e. the GPS r) and the occurrence of a particular treatment assignment. Under the *well-specification* of nuisance components, regressing Z_k on $\boldsymbol{X}$ produces a *consistent* estimator i.e. $\mathbb{E}(Z_k \mid \boldsymbol{X} = \boldsymbol{x}) = \tau_k(\boldsymbol{x})$ while keeping the same sample size as $\mathbf{D}_{\mathrm{obs}}$. In the following, we say that an estimator ($\widehat{\mu}$ or $\widehat{r}$) is *well-specified* if it is based on a well-specified statistical model, that is, the class of distributions assumed for modelling contains the unknown probability distribution from which the sample used for estimation is drawn.

M-learner with multiple treatments. The *M-learner* (Horvitz & Thompson, 1952), where M refers to the *modified* learned pseudo-outcome in the algorithm, is inspired from the Inverse Propensity Weighting (IPW). It is defined in the multi-valued setting, for $k = 1, \dots, K$, as the regression of Z_k^M such that

$$Z_k^M = \frac{\mathbf{1}\{T = t_k\}}{\widehat{r}(t_k, \boldsymbol{X})} Y_{\mathrm{obs}} - \frac{\mathbf{1}\{T = t_0\}}{\widehat{r}(t_0, \boldsymbol{X})} Y_{\mathrm{obs}},$$

where $\widehat{r}$ is an estimator of the GPS r.

DR-learner with multiple treatments. The *Doubly Robust* (DR) method (Robins et al., 1994; Kennedy et al., 2017; Kennedy, 2020) is helpful in overcoming the problem of the model's misspecification by estimating two components, the outcome model $\mu_{\cdot}$ and the GPS r, instead of relying on the correctness of one (and the only) parameter. If $\widehat{\mu}$ and $\widehat{r}$ denote arbitrary estimators of the outcome μ and the GPS r (we assume that $\widehat{r}$ satisfies Assumption 3.2), then the DR-learner regresses $Z_{\widehat{\mu}, \widehat{r}, k}^{DR}$ such that:

$$Z_{\widehat{\mu}, \widehat{r}, k}^{DR} = \frac{Y_{\mathrm{obs}} - \widehat{\mu}_T(\boldsymbol{X})}{\widehat{r}(t_k, \boldsymbol{X})} \mathbf{1}\{T = t_k\} + \widehat{\mu}_{t_k}(\boldsymbol{X})$$
$$- \frac{Y_{\mathrm{obs}} - \widehat{\mu}_T(\boldsymbol{X})}{\widehat{r}(t_0, \boldsymbol{X})} \mathbf{1}\{T = t_0\} - \widehat{\mu}_{t_0}(\boldsymbol{X}).$$

X-learner with multiple treatments. The X-learner (Künzel et al., 2019), also known as *Regression-Adjustment* (RA)-learning in a version developed by Curth & van der Schaar (2021a), has been proposed as an alternative to T-learning in the case where one treatment group is overrepresented. The idea consists of a *cross* procedure of estimation between observations Y_{obs} and outcome models when one of the treatments occurs. For $k = 1, \dots, K$, we define the *Regression-Adjustment* pseudo-outcome Z_k^X as

$$Z_k^X = \mathbf{1}\{T = t_k\}(Y_{\mathrm{obs}} - \widehat{\mu}_{t_0}(\boldsymbol{X})) + \sum_{l \neq k} \mathbf{1}\{T = t_l\} \times$$
$$(\widehat{\mu}_{t_k}(\boldsymbol{X}) - Y_{\mathrm{obs}}) + \sum_{l \neq k} \mathbf{1}\{T = t_l\}(\widehat{\mu}_{t_l}(\boldsymbol{X}) - \widehat{\mu}_{t_0}(\boldsymbol{X})).$$

For comparison purposes, we consider also the naive extension of the binary X-learner (Künzel et al., 2019) to multi-treatments as proposed by Zhao & Harinen (2019). This extension considers two random variables $D^{(k)} := Y(t_k) - \widehat{\mu}_{t_0}(\boldsymbol{X})$ and $D^{(0)} := \widehat{\mu}_{t_k}(\boldsymbol{X}) - Y(t_0)$ where $\widehat{\mu}_{t_k}$ and $\widehat{\mu}_{t_0}$ are trained on the samples $\mathbf{S}_{t_k}$ and $\mathbf{S}_{t_0}$. Then it regresses $(D_i^{(k)})_{i \in \mathbf{S}_{t_k}}$ and $(D_i^{(0)})_{i \in \mathbf{S}_{t_0}}$ on $\boldsymbol{X}$ to obtain $\widehat{\tau}^{(k)}$ and $\widehat{\tau}^{(0)}$, and estimates the CATE as:

$$\widehat{\tau}_k^{(\mathrm{X,nv})}(\boldsymbol{x}) := \frac{\widehat{r}(t_k, \boldsymbol{x})}{\widehat{r}(t_k, \boldsymbol{x}) + \widehat{r}(t_0, \boldsymbol{x})} \widehat{\tau}^{(k)}(\boldsymbol{x}) +$$
$$\frac{\widehat{r}(t_0, \boldsymbol{x})}{\widehat{r}(t_k, \boldsymbol{x}) + \widehat{r}(t_0, \boldsymbol{x})} \widehat{\tau}^{(0)}(\boldsymbol{x}).$$

We note that the consistency of the M- and DR-learners is already established in the literature (Knaus, 2022). We show it also for the extended X-learner in Appendix A. We note also that there are three main approaches possible to learn nuisance components (r and μ) and then estimate the τ_k, namely, Full-Sample, Sample-Split and Cross-Fit methods (Okasa, 2022). This paper does not discuss estimation procedures and adopts the Full-Sample strategy.

4.3. Neyman-Orthogonality based learner: R-learner

The R-learner is based mainly on the Robinson (1988) decomposition and was proposed by Nie & Wager (2020) to provide a flexible CATE estimator avoiding regularization bias. It states that the potential outcome error $\epsilon = Y_{\mathrm{obs}} - \mu_T(\boldsymbol{X})$ satisfies $\mathbb{E}(\epsilon \mid T, \boldsymbol{X}) = 0$ and

$$\epsilon = Y_{\mathrm{obs}} - m(\boldsymbol{X}) - \sum_{k=1}^{K} \left(\mathbf{1}\{T = t_k\} - r(t_k, \boldsymbol{X}) \right) \tau_k(\boldsymbol{X}),$$

where $m(\boldsymbol{x}) = \mathbb{E}(Y_{\mathrm{obs}} \mid \boldsymbol{X} = \boldsymbol{x})$ is the observed outcome model. Therefore, considering the mean squared error of ϵ (the generalized R-loss function) and minimizing it estimate

K models $\{\widehat{\tau}_k^{(\mathrm{R})}\}_{k=1}^K$ simultaneously such that

$$\{\widehat{\tau}_k^{(\mathrm{R})}\}_{k=1}^K := \operatorname{argmin}_{\{\overline{\tau}_k\} \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n \Big[\big(Y_{\mathrm{obs},i} - \widehat{m}(\boldsymbol{X}_i)\big)$$

$$- \sum_{k=1}^K \big(\mathbf{1}\{T_i = t_k\} - \widehat{r}(t_k, \boldsymbol{X}_i)\big)\overline{\tau}_k(\boldsymbol{X}_i) \Big]^2,$$

where $\widehat{m}$ (respectively, $\widehat{r}$) is an estimator of m (respectively, r) and $\mathcal{F}$ is the space of candidate models $[\{\overline{\tau}_k\}_{k=1}^K]$.

The generalized R-learner suffers from two major limitations: First, it cannot be written as *weighted* supervised learning problem with a specific pseudo-outcome. Only parametric families $\mathcal{F}$ can be considered in the multi-treatment regime. The second drawback is the non-identifiability of the generalized R-loss described previously in (4.3) without a regularization. Indeed, this problem does not have a unique solution (see Appendix A for details) and leads thus to poor estimation performance. This point is also shown recently by Zhang et al. (2022) for continuous treatments and our numerical results in Appendix D also confirm that the R-learner fails to estimate CATEs $(\tau_k)_{k=1}^K$.

5. Theoretical analysis of the error upper bound

In this section, we analyze the error's upper bounds for different meta-learners. The theoretical analysis will be carried out under the following framework:

Assumption 5.1. We assume that $(T, \boldsymbol{X})$ satisfies the overlap assumption 3.2 and that, for $t \in \mathcal{T}$, the outcome $Y(t)$ is generated from a function $f : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ such that

$$Y(t) = f(t, \boldsymbol{X}) + \varepsilon(t), \tag{2}$$

where $\varepsilon(t)$ are i.i.d. Gaussian $\mathcal{N}(0, \sigma^2)$ and independent of $(T, \boldsymbol{X})$.

Assumption 5.2. We assume the existence of $\boldsymbol{\beta}_t \in \mathbb{R}^p$ such that, for all $t \in \mathcal{T}$ and $\boldsymbol{x} \in \mathcal{D}$

$$f(t, \boldsymbol{x}) = \sum_{j=0}^{p-1} \beta_{t,j} f_j(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})^\top \boldsymbol{\beta}_t,$$

where f_j are some bounded predefined basis functions.

The assumption of a product effect is reasonable. One can show the universality of this representation in the Reproducing Kernel Hilbert Space (RKHS) (Proposition 1 of Kaddour et al. (2021)) if we allow the dimension p to be large enough.

Under these two assumptions, the CATE τ_k can be written as:

$$\tau_k(\boldsymbol{x}) = \sum_{j=0}^{p-1} \beta_{k,j}^* f_j(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})^\top \boldsymbol{\beta}_k^*,$$

where $\boldsymbol{\beta}_k^* = (\beta_{k,j}^*)_{j=0}^{p-1} = \boldsymbol{\beta}_{t_k} - \boldsymbol{\beta}_{t_0} \in \mathbb{R}^p$.

From a theoretical point of view, the S-learner corresponds to the naive Ordinary Least Square (OLS) $\widehat{\boldsymbol{\beta}}_k$ of $\boldsymbol{\beta}_k^*$. The statistical task of the CATE's estimation holds immediately. However, we cannot properly analyse the base-learner's ability to learn $\boldsymbol{\beta}_k^*$ under confounding effects.

Theorem 5.3. *Under Assumptions (5.1-5.2), the OLS estimators $\widehat{\boldsymbol{\beta}}_k^*$ of the T-learner and the naive X-learner are unbiased and have an asymptotic covariance matrix $\mathbb{V}(\widehat{\boldsymbol{\beta}}_k^*) = \mathbf{C}/n$, whose terms $\mathbf{C}_{ij}$ are bounded by:*

$$\mathcal{E}^T = \mathcal{E}^{X,nv} = \mathcal{O}\left(\frac{1}{\rho(t_k)} + \frac{1}{\rho(t_0)}\right),$$

where $\mathbb{P}(T = t) = \rho(t) > 0$ for all $t \in \mathcal{T}$.

We consider now pseudo-outcome meta-learners (M-, DR- and X-learners). When investigating the pseudo-outcomes Z_k, one can see that, for $k = 1, \dots, K$, these pseudo-outcomes are linear with respect to Y_{obs} i.e.

$$Z_k = A_{t_k}(T, \boldsymbol{X})Y_{\mathrm{obs}} + B_{t_k}(T, \boldsymbol{X}),$$

where $A_{t_k}(T, \boldsymbol{X})$ and $B_{t_k}(T, \boldsymbol{X})$ are given for each pseudo-outcome meta-learner.

Theorem 5.4. *Under Assumptions (5.1-5.2), the OLS estimator $\widehat{\boldsymbol{\beta}}_k^*$ is unbiased if the nuisance parameters ($\widehat{\mu}$ and $\widehat{r}$) are well-specified, and has an asymptotic covariance matrix $\mathbb{V}(\widehat{\boldsymbol{\beta}}_k^*) = \mathbf{C}/n$, whose terms, for all $\epsilon > 0$, are bounded by:*

$$\mathcal{E}^M = \mathcal{O}\left(\frac{1}{r_{\min}^{1+\epsilon}}\right) \text{ for the M-learner,}$$

$$\mathcal{E}^{DR} = \mathcal{O}\left(\frac{\mathrm{err}(\widehat{\mu}_{t_k}) + \mathrm{err}(\widehat{\mu}_{t_0})}{r_{\min}^{1+\epsilon}}\right) \text{ for the DR-learner,}$$

$$\mathcal{E}^X = \mathcal{O}\left(K^2 \sum_{l \neq k} \mathrm{err}(\widehat{\mu}_{t_l})\right) \text{ for the X-learner,}$$

where $\mathrm{err}(\widehat{\mu}_t) = \mathbb{E}_{\boldsymbol{X}}\big[\big(f(t, \boldsymbol{X}) - \widehat{\mu}_t(\boldsymbol{X})\big)^2\big]$ is the expected mean squared error of $\widehat{\mu}_t$.

Sketch of the proofs of Theorems 5.3 and 5.4. Both proofs are similar and are structured in 3 steps: 1) Express the OLS estimator $\widehat{\boldsymbol{\beta}}_k^*$ as a function of the true $\boldsymbol{\beta}_k^*$; 2) Apply the multivariate Central Limit Theorem and the Slutsky theorem (or the Delta method in the general case with a biased $\widehat{\boldsymbol{\beta}}_k^*$); 3) Bound the asymptotic covariance matrix terms. The full proofs can be found in Appendix B.

Remark 5.5. For $K = 2$, one recovers the error's upper bounds of pseudo-outcome meta-learners and the important parameters as shown by Curth & van der Schaar (2021a). However, the influence of the sampling probability ρ for the T- and naive X-learners is original.

Finally, regarding the binarized (Kaddour et al., 2021) or generalized R-learners, they cannot be expressed as a *model-agnostic* regression problem but rather as a minimization problem of a loss function (generalized R-loss or binarized R-loss). Therefore, the actual theoretical analysis of the error's upper bound cannot be conducted similarly.

Using Theorems 5.3 and 5.4, we establish the following discussions (see Table 1 for a summary).

5.1. General comments and insights

About Theorem 5.3. The significant implication of Theorem 5.3 is the ability to anticipate the performances of the T-learner with respect to the treatment distribution when the probability $\rho(t)$ of sampling the treatment value t is small. The performances of the T- and naive X-learners become poor when the number K of treatments increases because these learners imply learning in small samples $\mathbf{S}_t$. Moreover, the naive X-learner also does not seem to really differ from the T-learner in any meaningful way (see Appendix B.2). We note that this result is original and cannot be obtained by the minimax approach.

About Theorem 5.4. From this theorem, one can establish the relationship between the number of observations n and the number of treatment values K for a given error bound. The same theorem is also useful for an RCT setting ($r_{\min}$ is known) and when the error of the outcome model is small for all treatment levels except for one treatment value t.

Moreover, although pseudo-outcome meta-learners have the advantage of learning on the full sample, they, unfortunately, may lead to high error and poor performance in how nuisance components intervene. On the one hand, the GPS is in the M- and DR-learners denominators. The error bound is likely to be high when there is a lack of overlap or when the number of treatments K increases (we have necessarily $r_{\min} \leq 1/K$). On the other hand, the upper bounds of the X- and DR-learners depend on the quality of the estimated $\widehat{\mu}$. One can expect that the more precise the outcome models are, the lower the error is.

5.2. Specific comments for each meta-learner

M-learner. Without surprise, the M-learner is very sensitive to the estimated GPS $\widehat{r}$ and suffers from high variance. This is even more critical as the number K of treatments increases.

DR-learner. The error's estimation term $\mathrm{err}(\widehat{\mu}_t)$ of $\widehat{\mu}_t$ and $\widehat{\mu}_{t_0}$ in the numerator dramatically improves the DR-learner's performances.

X-learner. The X-learner incorporates only $\widehat{\mu}_{\cdot}$ and does not imply the GPS r. Thus, the X-learner is likely to have the smallest error compared to other meta-learners when the

overlap assumption is not sufficiently respected. However, the consistency of $(\widehat{\mu}_t)_{t \neq t_k}$ is required to estimate CATEs correctly.

M-learner vs DR-learner. If the outcome models $\widehat{\mu}_t$ and $\widehat{\mu}_{t_0}$ are well-specified, the error's upper bound is expected to be smaller for the DR-learner than for the M-learner. However, if they are misspecified (but the propensity score is well-specified), then there is no guarantee that the DR-learner would perform better than M-learner. It may perform even worse, as we will see in Appendix D (Table 8).

M-learner vs X-learner. The X-learner is likely to have a lower error upper bound if the expected squared error $\mathrm{err}(\widehat{\mu}_t)$ is small and if some conditions on K and $r_{\min}$ hold.

X-learner vs DR-learner. Analytically, it is difficult to anticipate which meta-learner would perform better. This depends mainly on $\mathrm{err}(\widehat{\mu}_{\cdot})$, K and $r_{\min}$, whom, in some cases, make the X-learner have less error than the DR-learner and the opposite in other cases. Still, our numerical results in Appendix D show that the X-learner outperforms the DR-learner in most cases.

X-learner vs Naive X-learner. We cannot theoretically compare these meta-learners without knowledge about the distribution of T. However, we can see numerically that the X-learner clearly outperforms the naive X-learner.

We conclude this subsection with a small discussion about the generalized and the binarized R-learners (see Appendix C). Indeed, the binary R-loss function may be solved separately in a low sample (instead of $\mathbf{D}_{\mathrm{obs}}$) and one can obtain a unique solution, unlike the generalized R-loss. However, the optimization procedure is two-stage iterative and computationally heavy. We do not consider it in Section 6.

Table 1. Summary table of multi-treatments meta-learners.

Meta-learner	Advantages	Disadvantages
T-learner (nv X-learner)	Simple approach	Selection bias Low samples
S-learner	Simple approach	Confounding effects Regularization bias
M-learner	Consistency	High variance
DR-learner	Consistency Doubly Robust	Possibly high variance
X-learner	Consistency Low variance	Non-intuitive
R-learner	Interaction effects	Non-identifiability
Bin R-learner	Identifiability	Computational cost

In Table 1, *Bin R-learner* refers to the binarized R-learner (Kaddour et al., 2021) and *nv X-learner* refers to the naive extension of the X-learner described in subsection 4.2. *Possibly high variance* refers to the case where the variance can be significantly high due to the lack of overlap caused by the inverse propensity weighting in some DGPs. *Selection bias* refers to the bias that occurs when sampling $\mathbf{S}_t$ and comparing units directly, as we describe in Proposition 4.1 in Appendix A. The confounding effects represent the statistical and intrinsic dependence between the treatment T and the covariates $\mathbf{X}$, which prevent some base-learner (e.g. random forest) from distinguishing and disassociating the treatment T and the covariates $\mathbf{X}$. Finally, *low samples* refer to cases where the samples $\mathbf{S}_t = \{i, T_i = t\}$ become small for some treatment levels and the model has then few observations to learn from.

5.3. Practical recommendations for selecting a meta-learner

In this subsection, based on previous insights and our numerical findings (see Section 6 for more detail), we provide some instructions and recommendations for selecting meta-learners given a dataset $\mathbf{D}_{\text{obs}}$:

- Pseudo-outcomes meta-learners and the S-learner are preferred in low sample regimes.

- It is recommended to examine the distribution of T to understand at which values of t the T-learner may fail to learn CATEs.

- The S-learner remains a simple and reasonable choice, especially when $K \geq 10$.

- The X-learner is useful to learn the CATE τ_t when there is not enough information about a specific treatment value t.

- If one is interested in quantifying uncertainties, then the DR-learner is recommended as it allows having a consistent estimation of the CATE.

- After estimating the GPS r, one needs to check if there is a possible lack of overlap. If a lack of overlap is detected, one should avoid considering the M- and DR-learners.

6. Numerical experiments

In this section, we assess the performances of the different meta-learners, additional numerical results are shown in Appendix D.

In synthetic or semi-synthetic examples where the CATEs are known, the error in estimation is given by **mPEHE** (respectively, **sdPEHE**), the mean (respectively, the standard deviation) of the Precision in Estimation of Heterogeneous

Effect (PEHE) (Hill, 2011; Shalit et al., 2017) defined as the mean squared error in the estimation of the treatment effect $\widehat{\tau}_k$, over all possible treatment levels t_k for $k = 1, \ldots, K$:

$$\mathbf{mPEHE} = \sqrt{\frac{1}{K}\sum_{k=1}^{K} PEHE(\widehat{\tau}_k)^2},$$

where $PEHE(\widehat{\tau}_k)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\widehat{\tau}_k(\mathbf{X}_i) - \tau_k(\mathbf{X}_i)\right)^2$, and,

$$\mathbf{sdPEHE} = \sqrt[4]{\frac{1}{K-1}\sum_{k=1}^{K}\left(PEHE(\widehat{\tau}_k)^2 - \mathbf{mPEHE}^2\right)^2}$$

Those metrics will be used to compare meta-learners under different scenarios (sample size n, number of possible treatments K, the correctness of nuisance parameters). We do not consider here model-fitting of base-learners. All hyperparameters (e.g. the number of trees, depth etc.) are fixed to their default values during all experiments. In addition, we do not consider Neural Networks because it would require choosing between at least five possible architectures (Curth & van der Schaar, 2021a) to define tasks of learning nuisance components and estimating CATEs τ_k while the main focus of the paper is on the choice of the meta-learner.

6.1. Synthetic datasets: analytical functions in randomized and non-randomized studies

In this subsection, we begin by empirically evaluating the performances of meta-learners when the treatment T is taking $K + 1 = 10$ possible values in $[0, 1]$ on a linear outcome:

$$Y(t) \text{ of the form (2) with}$$
$$f(t, X) = (1 + t)X, \text{ and } X \sim \mathcal{U}[0, 1],$$

in Randomized Controlled Trials (RCT) setting where the T and X are independent. Second, we evaluate meta-learners on the hazard rate outcome:

$$Y(t) \text{ of the form (2) with}$$
$$f(t, \mathbf{X}) = t + \|\mathbf{X}\| \exp\left(-t\|\mathbf{X}\|\right) \text{ and } \mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_5)$$

in a non-randomized setting as will be described below.

To simulate observational data, instead of removing some rows, we create a selection bias in the data by selecting preferentially only observations with specific characteristics (see subsection D.1 in Appendix D). This strategy comes in line with the findings and recommendations of Curth et al. (2021) about creating a biased sub-sample and evaluating CATEs' estimators.

The GPS is estimated using the XGBoost model, and the outcome models μ_t are either estimated by the T- or S-learning approaches. In Tables 2 and 3 and Appendix D,

Table 2. **mPEHE** and **sdPEHE** for XGBoost and RandomForest; linear model in a RCT setting with $n = 2000$ units.

Meta-learner	XGBoost	RandomForest
T-Learner	0.065 (0.019)	0.041 (0.016)
S-Learner	**0.033 (0.018)**	0.032 (0.028)
*Nv*X-Learner	0.060 (0.019)	**0.037 (0.016)**
M-Learner	1.25 (0.610)	1.22 (0.621)
DR-Learner	0.068 (0.019) – 0.063 (0.020)	0.068 (0.018) – 0.068 (0.018)
X-Learner	0.063 (0.020) – **0.033 (0.017)**	0.045 (0.016) – 0.061 (0.040)
RLin-Learner	0.135 (0.130)	0.137 (0.128)

For the DR- and X-learners: μ_t are estimated by T- (left value) or S- (right value). The bold font indicates the best meta-learner (row) per base-learner (column).

Table 3. **mPEHE** and **sdPEHE** for XGBoost and RandomForest. Hazard rate model in an observational setting with $n = 10000$ units.

Meta-learner	XGboost	RandomForest
T-Learner	0.183 (0.039)	0.286 (0.155)
*Reg*T-Learner	0.176 (0.044)	0.286 (0.155)
S-Learner	0.176 (0.056)	0.306 (0.153)
*Nv*X-Learner	0.190 (0.096)	0.336 (0.200)
M-Learner	1.61 (0.505)	1.58 (0.472)
DR-Learner	0.168 (0.045) - 0.178 (0.048)	0.304 (0.158) – 0.322 (0.162)
X-Learner	**0.167 (0.053)** – 0.172 (0.057)	0.302 (0.169) – 0.332 (0.167)
RLin-Learner	0.231 (0.081)	**0.186 (0.123)**

For the DR- and X-learners: μ_t are estimated by T- (left value) or S- (right value). The bold font indicates the best meta-learner (row) per base-learner (column).

RLin-learner denotes the generalized R-learner with linear regression models in (4.3) with $p = 2$. For each meta-learner (row) and base-learner (column), we indicate the **mPEHE** followed by the **sdPEHE** between brackets.

In Tables 2 and 3, we find that, as expected, the M-learner predicts poorly. The T- and naive X-learners give better predictions for Random Forest, whereas the S-learner gives better results for XGBoost. Regularizing T-learner (*Reg*T-Learner) against selection bias increases its performance. The X- and DR-learners improve the predictions of the S-learner for XGBoost, but this improvement is not always observable for Random Forests. Unfortunately, the actual results (and also additional numerical experiments in Appendix D) confirm the claim: The RLin-learner generally fails to identify CATEs correctly.

Despite these satisfying results, we highlight the problem of over-fitted gradient boosting models and Random Forest by comparing them with the linear model in Appendix D. This problem should be taken further while estimating CATEs. We think that using out-sample prediction supervised models might solve this problem.

We consider now the effect of increasing K on the hazard rate function with XGBoost. The results are shown in Figure 1 in Appendix D.3. On the one hand, the performances of the T- and the naive X-learners become compromised. The regularized T-learner suffers also from the same issue (with a linear effect with respect to K), which can be also quantified on the DR- and X-learners when applying the regularized T-learning. For $K \geq 20$, the T- and the naive X-learners perform better than the previous meta-learners. Regarding the M-learner, it has poor performances in all cases as can be expected with Theorem 5.4. On the other hand, the S-learner stabilises once K is large enough and, consequently, the DR- and X-learners also stabilize when applying the S-learning for large K. Therefore, we recommend the S-learner's estimated potential outcome model when $K \geq 10$ for pseudo-outcome meta-learners. To conclude, two-step meta-learners are robust. In particular, the X-learner improves the quality of plug-in meta-learners; when it does not, the differences are very small.

6.2. Semi-synthetic dataset: estimating heterogeneous treatment effects on a non-randomized dataset.

In this subsection, we consider a multistage fracturing Enhanced Geothermal System (EGS) (Olasolo et al., 2016). We assume that the heat extraction performance satisfies the physical model: $Q_{well}(\ell_L) = Q_{fracture} \times \ell_L / d \times \eta_d$, where

Table 4. **mPEHE** and **sdPEHE** for XGBoost and RandomForest. Heat Extraction model in an observational setting.

Meta-learner	XGBoost	RandomForest
T-learner	0.172 (0.052)	0.157 (0.067)
*Reg*T-Learner	0.156 (0.042)	0.154 (0.067)
S-learner	0.101 (0.040)	0.218 (0.129)
*Nv*X-Learner	0.102 (0.042)	**0.143 (0.067)**
M-learner	1.04 (0.423)	0.898 (0.417)
DR-learner	0.148 (0.042) – 0.097 (0.029)	0.164 (0.068) – 0.203 (0.108)
X-learner	0.142 (0.041) – **0.094 (0.034)**	0.173 (0.077) – 0.211 (0.120)
RLin-learner	0.357 (0.274)	0.362 (0.278)

For the DR- and X-learners: μ_t are estimated by T- (left value) or S- (right value). The bold font indicates the best meta-learner (row) per base-learner (column).

$Q_{fracture}$ is the *unknown* heat extraction performance from a single fracture, that can be generated using a numerical model with eight input parameters including reservoir characteristics and fracture design. ℓ_L is the Lateral Length of the well, d is the average spacing between two fractures and η_d is the stage efficiency penalizing the individual contribution when fractures are close to each other. We refer to Appendix E for a detailed description of the model and the semi-synthetic dataset.

We consider the Lateral Length as treatment T with $K+1 = 13$ possible values and the covariates $\boldsymbol{X} \in \mathbb{R}^{11}$ are the remaining variables. We also consider a logarithmic transformation of the heat performance for a meaningful **mPEHE**, and we normalize the treatment T. Following the *preferential selection*, we sample $n = 10000$ units such that wells with high lateral length are likely to have larger fractures and vice versa. The GPS is estimated using gradient boosting models. Table 4 resumes the **mPEHE** and **sdPEHE** in brackets for different meta-learners. Most findings of subsection 6.1 remain valid: XGBoost model is generally a better choice than Random Forests (except for T-learning); the X-learner, followed by the DR-learner, outperforms all other learners.

7. Conclusion

We have investigated heterogeneous treatment effects estimation with multi-valued treatments. In addition to standard plug-in meta-learners, we have considered representations to build pseudo-outcome meta-learners, and we have proposed the generalized Robinson decomposition to build the R-learner. Using the bias-variance analysis, we have conducted an in-depth analysis of the error's upper bounds of pseudo-outcomes meta-learners. Thanks to this analysis, we could discuss the advantages and limitations of each pseudo-outcome meta-learner. In particular, we have identified the impacts of the number of treatment levels and the lower bound $r_{\min}$ on the M-, DR and X-learners. Through

synthetic and semi-synthetic industrial datasets, we have illustrated the performances of different meta-learners in a non-randomized case where some covariates are confounded with the treatment. We have demonstrated the ability of the X-learner to reconstruct the ground truth model. We have also highlighted how the choice of base-learner can affect the quality of CATEs estimation.

8. Software and Data

The code and the semi-synthetic dataset in subsection 6.2 are available at https://github.com/nacharki/multipleT-MetaLearners.

9. Acknowledgements

References

Alaa, A. and van der Schaar, M. Limits of estimating heterogeneous treatment effects: Guidelines for practical algorithm design. In Dy, J. and Krause, A. (eds.), *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *PMLR*, pp. 129–138. PMLR, 10–15 Jul 2018.

Alaa, A. M. and van der Schaar, M. Bayesian inference of individualized treatment effects using multi-task gaussian processes. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, pp. 3427–3435. Curran Associates Inc., 2017. ISBN 9781510860964.

Athey, S. and Imbens, G. Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, 113(27):7353–7360, 2016. doi: 10.1073/pnas.1510489113.

Caron, A., Baio, G., and Manolopoulou, I. Shrinkage bayesian causal forests for heterogeneous treatment effects estimation. *Journal of Computational and Graphical Statistics*, pp. 1–13, 2022a.

Caron, A., Baio, G., and Manolopoulou, I. Estimating individual treatment effects using non-parametric regression models: A review. *Journal of the Royal Statistical Society Series A*, 185(3):1115–1149, July 2022b. doi: 10.1111/rssa.12824.

Chen, J., Dong, H., Wang, X., Feng, F., Wang, M., and He, X. Bias and debias in recommender system: A survey and future directions. *ACM Transactions on Information Systems*, 41(3):1–39, 2023.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68, 01 2018. ISSN 1368-4221. doi: 10.1111/ectj.12097.

Colangelo, K. and Lee, Y.-Y. Double debiased machine learning nonparametric inference with continuous treatments. *arXiv preprint arXiv:2004.03036*, 2020.

Curth, A. and van der Schaar, M. Nonparametric estimation of heterogeneous treatment effects: From theory to learning algorithms. In Banerjee, A. and Fukumizu, K. (eds.), *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of *PMLR*, pp. 1810–1818. PMLR, 13–15 Apr 2021a.

Curth, A. and van der Schaar, M. On inductive biases for heterogeneous treatment effect estimation. In Beygelzimer, A., Dauphin, Y., Liang, P., and Vaughan, J. W. (eds.), *Advances in Neural Information Processing Systems*. PMLR, 2021b.

Curth, A., Svensson, D., Weatherall, J., and van der Schaar, M. Really doing great at estimating CATE? a critical look at ML benchmarking practices in treatment effect estimation. In *35th Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*, 2021.

Diemert, E., Betlei, A., Renaudin, C., and Amini, M.-R. A large scale benchmark for uplift modeling. In *Proceedings of the AdKDD and TargetAd Workshop, KDD, London, United Kingdom, August, 20, 2018*. ACM, 2018.

Dominici, F., Daniels, M., Zeger, S. L., and Samet, J. M. Air pollution and mortality: estimating regional and national dose-response relationships. *Journal of the American Statistical Association*, 97(457):100–111, 2002.

Dorie, V., Hill, J., Shalit, U., Scott, M., and Cervone, D. Automated versus do-it-yourself methods for causal inference: Lessons learned from a data analysis competition. *Statistical Science*, 34(1):43–68, 2019.

Flores, C. A. Estimation of Dose-Response Functions and Optimal Doses with a Continuous Treatment. Technical Report 0707, University of Miami, Department of Economics, November 2007.

Frölich, M. Programme evaluation with multiple treatments. *Wiley-Blackwell: Journal of Economic Surveys*, 2002.

Hahn, P., Murray, J., and Carvalho, C. Bayesian regression tree models for causal inference: Regularization, confounding, and heterogeneous effects (with discussion). *Bayesian Analysis*, 15(3):965–1056, 2020. ISSN 1936-0975. doi: 10.1214/19-BA1195.

Harada, S. and Kashima, H. Graphite: Estimating individual effects of graph-structured treatments. In *Proceedings of the 30th ACM International Conference on Information and Knowledge Management*, pp. 659–668. Association for Computing Machinery, 2021. ISBN 9781450384469.

Hassanpour, N. and Greiner, R. Counterfactual regression with importance sampling weights. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence, IJCAI-19*, pp. 5880–5887. International Joint Conferences on Artificial Intelligence Organization, 7 2019. doi: 10.24963/ijcai.2019/815.

Heiler, P. and Knaus, M. C. Effect or Treatment Heterogeneity? Policy Evaluation with Aggregated and Disaggregated Treatments. Iza discussion papers, Institute of Labor Economics (IZA), September 2022.

Hill, J. L. Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics*, 20(1):217–240, 2011. doi: 10.1198/jcgs.2010.08162.

Holland, P. W. Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–960, 1986.

Horvitz, D. G. and Thompson, D. J. A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47(260):663–685, 1952. ISSN 01621459.

Hu, L., Gu, C., Lopez, M., Ji, J., and Wisnivesky, J. Estimation of causal effects of multiple treatments in observational studies with a binary outcome. *Statistical methods in medical research*, 29(11):3218–3234, 2020.

Imai, K. and Dyk, D. A. V. Causal inference with general treatment regimes: Generalizing the propensity score. *Journal of the American Statistical Association*, 99(467): 854–866, 2004.

Imai, K. and Strauss, A. Estimation of heterogeneous treatment effects from randomized experiments, with application to the optimal planning of the get-out-the-vote campaign. *Political Analysis*, 19(1):1–19, 2011. doi: 10.1093/pan/mpq035.

Imbens, G. W. The role of the propensity score in estimating dose-response functions. *Biometrika*, 87(3):706–710, 2000.

Johansson, F., Shalit, U., and Sontag, D. Learning representations for counterfactual inference. In Balcan, M. F. and Weinberger, K. Q. (eds.), *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *PMLR*, pp. 3020–3029. PMLR, 20–22 Jun 2016.

Kaddour, J., Zhu, Y., Liu, Q., Kusner, M. J., and Silva, R. Causal effect inference for structured treatments. In Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P., and Vaughan, J. W. (eds.), *Advances in Neural Information Processing Systems*, volume 34, pp. 24841–24854. Curran Associates, Inc., 2021.

Kallus, N. Recursive partitioning for personalization using observational data. In *International conference on machine learning*, pp. 1789–1798. PMLR, 2017.

Kennedy, E. H. Optimal doubly robust estimation of heterogeneous causal effects, 2020.

Kennedy, E. H., Ma, Z., McHugh, M. D., and Small, D. S. Non-parametric methods for doubly robust estimation of continuous treatment effects. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79 (4):1229–1245, 2017.

Knaus, M. C. Double machine learning-based programme evaluation under unconfoundedness. *The Econometrics Journal*, 25(3):602–627, 2022.

Knaus, M. C., Lechner, M., and Strittmatter, A. Machine learning estimation of heterogeneous causal effects: Empirical Monte Carlo evidence. *The Econometrics Journal*, 24(1):134–161, 06 2020a.

Knaus, M. C., Lechner, M., and Strittmatter, A. Heterogeneous employment effects of job search programmes: A machine learning approach. *Journal of Human Resources*, pp. 0718–9615R1, Mar 2020b. ISSN 1548-8004. doi: 10.3368/jhr.57.2.0718-9615r1.

Künzel, S. R., Sekhon, J. S., Bickel, P. J., and Yu, B. Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences*, 116(10):4156–4165, Feb 2019. ISSN 1091-6490. doi: 10.1073/pnas.1804597116.

Lechner, M. Identification and estimation of causal effects of multiple treatments under the conditional independence assumption. In Lechner, M. and Pfeiffer, F. (eds.), *Econometric Evaluation of Labour Market Policies*, pp. 43–58. Physica-Verlag HD, 2001.

Lin, L., Zhu, Y., and Chen, L. Causal inference for multilevel treatments with machine-learned propensity scores. *Health Services and Outcomes Research Methodology*, 19(2):106–126, 2019.

Neyman, J. On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9. *Statistical Science*, 5:465 – 472, 1923. Translated and published in English in 1990.

Nie, L., Ye, M., qiang liu, and Nicolae, D. {VCN}et and functional targeted regularization for learning causal effects of continuous treatments. In *International Conference on Learning Representations*, 2021.

Nie, X. and Wager, S. Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika*, 09 2020. doi: 10.1093/biomet/asaa076.

Okasa, G. Meta-learners for estimation of causal effects: Finite sample cross-fit performance. *arXiv preprint arXiv:2201.12692*, 2022.

Olasolo, P., Juárez, M., Morales, M., Liarte, I., et al. Enhanced geothermal systems (egs): A review. *Renewable and Sustainable Energy Reviews*, 56:133–144, 2016.

Pearl, J. The seven tools of causal inference, with reflections on machine learning. *Commun. ACM*, 62(3):54–60, feb 2019. ISSN 0001-0782. doi: 10.1145/3241036.

Powers, S., Qian, J., Jung, K., Schuler, A., Shah, N. H., Hastie, T., and Tibshirani, R. Some methods for heterogeneous treatment effect estimation in high dimensions. *Statistics in medicine*, 37(11):1767–1787, 2018.

Robins, J. M., Rotnitzky, A., and Zhao, L. P. Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association*, 89(427):846–866, 1994.

Robinson, P. M. Root-n-consistent semiparametric regression. *Econometrica*, 56(4):931–954, 1988.

Rosenbaum, P. R. and Rubin, D. B. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 04 1983.

Rubin, D. Estimating causal effects if treatment in randomized and nonrandomized studies. *J. Educ. Psychol.*, 66, 01 1974.

Saini, S. K., Dhamnani, S., Aakash, Ibrahim, A. A., and Chavan, P. Multiple treatment effect estimation using deep generative model with task embedding. In *The World Wide Web Conference*, WWW '19, pp. 1601–1611. Association for Computing Machinery, 2019. ISBN 9781450366748. doi: 10.1145/3308558.3313744.

Schwab, P., Linhardt, L., Bauer, S., Buhmann, J., and Karlen, W. Learning counterfactual representations for estimating individual dose-response curves. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34: 5612–5619, 04 2020. doi: 10.1609/aaai.v34i04.6014.

Shalit, U., Johansson, F. D., and Sontag, D. Estimating individual treatment effect: Generalization bounds and algorithms. In *Proceedings of the 34th International Conference on Machine Learning - Volume 70*, ICML'17, pp. 3076–3085. JMLR.org, 2017.

Shi, C., Blei, D., and Veitch, V. Adapting neural networks for the estimation of treatment effects. In *Advances in neural information processing systems*, volume 32. PMLR, 2019.

Silverman, B. W. *Density estimation for statistics and data analysis*, volume 26. CRC press, 1986.

Sudret, B. Global sensitivity analysis using polynomial chaos expansions. *Reliability engineering & system safety*, 93(7):964–979, 2008.

Tibshirani, J., Athey, S., and Wager, S. *grf: Generalized Random Forests*, 2020. URL https://CRAN.R-project.org/package=grf. R package version 1.1.0.

Wager, S. and Athey, S. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242, 2018.

Wilks, S. Multidimensional statistical scatter. *Collected Papers, Contributions to Mathematical Statistics*, 01 1967.

Wilks, S. S. Certain generalizations in the analysis of variance. *Biometrika*, 24(3/4):471–494, 1932. ISSN 00063444.

Wu, P., Li, H., Deng, Y., Hu, W., Dai, Q., Dong, Z., Sun, J., Zhang, R., and Zhou, X.-H. On the opportunity of causal learning in recommendation systems: Foundation, estimation, prediction and challenges. In Raedt, L. D. (ed.), *Proceedings of the 31st International Joint Conference on Artificial Intelligence, IJCAI-22*, pp. 5646–5653.

International Joint Conferences on Artificial Intelligence Organization, 7 2022. doi: 10.24963/ijcai.2022/787.

Yoon, J., Jordon, J., and van der Schaar, M. GANITE: estimation of individualized treatment effects using generative adversarial nets. In *6th International Conference on Learning Representations, ICLR 2018*, 2018.

Zhang, Y., Kong, D., and Yang, S. Towards R-learner of conditional average treatment effects with a continuous treatment: T-identification, estimation, and inference. *arXiv preprint arXiv:2208.00872*, 2022.

Zhao, Z. and Harinen, T. Uplift modeling for multiple treatments with cost optimization. In *2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 422–431. IEEE, 2019.

The Appendix of the paper is divided into the following sections:

- Appendix A contains the proofs of the main propositions in Sections 3: Regularization of the T-learner, the consistency of the naive and extended X-learners, and the generalization of the Robinson decomposition. It also includes a solution for the generalized R-learner with linear regression models.

- Appendix B is divided into two parts: The first part B.1 establishes the bias-variance analysis of pseudo-outcome meta-learners (proof of Theorem 5.4). It provides the main framework for the proof based on the Delta Method and Slutsky's theorem, which will then be applied to the M-, DR-, and X-learners to show their error's upper bound. The second part B.2 establishes the bias-variance of the T- and the naive X-learners (proof of Theorem 5.3). This appendix follows the same logic as Appendix B.1's main proof.

- Appendix C discusses further the comparison of the generalized R-learner and the binarized R-learner.

- Appendix D concerns Section 6.1 with more insights about the numerical experiments and results.

- Appendix E describes the Semi-synthetic dataset used in subsection 6.2

A. Proofs of propositions of Section 4

Proposition A.1 (Regularizing the T-learner against selection bias). *For a treatment level $t \in \mathcal{T}$, the expected squared error of the estimator $\widehat{\mu}_t$ on the outcome surface μ_t satisfies:*

$$
\begin{aligned}
\mathbb{E}_{\boldsymbol{X} \sim \mathbb{P}(\cdot)}&\left[(\widehat{\mu}_t(\boldsymbol{X}) - \mu_t(\boldsymbol{X}))^2\right] = \\
\mathbb{E}_{\boldsymbol{X} \sim \mathbb{P}(\cdot \mid T=t)}&\left[\frac{\mathbb{P}(T=t)}{r(t, \boldsymbol{X})}\left(\widehat{\mu}_t(\boldsymbol{X}) - \mu_t(\boldsymbol{X})\right)^2\right].
\end{aligned}
\tag{3}
$$

where $\mathbb{P}(\cdot)$ is the marginal distribution of $\boldsymbol{X}$ and $\mathbb{P}(\cdot \mid T = t)$ is the conditional distribution of $\boldsymbol{X}$ given $T = t$.

A.1. Proof of Proposition A.1

This proof is similar to the proof of equation (5) in supplementary of Curth & van der Schaar (2021a). For simplicity, we assume that the distribution of X and the conditional distribution of $\boldsymbol{X}$ given $T = t$ are absolutely continuous with respect to the Lebesgue measure over $\mathbb{R}^d$. Let $p_{\boldsymbol{X}}(\boldsymbol{x})$ denote the probability distribution function of $\boldsymbol{X}$, let $p(\boldsymbol{x} \mid T = t)$ denote the probability distribution function of $\boldsymbol{X}$ given $T = t$ and let $R_t = \int (\widehat{\mu}_t(\boldsymbol{x}) - \mu_t(\boldsymbol{x}))^2 p(\boldsymbol{x} \mid T = t) d\boldsymbol{x}$.

$$\mathbb{E}_{\boldsymbol{X} \sim \mathbb{P}(\cdot)}\left[(\widehat{\mu}_t(\boldsymbol{X}) - \mu_t(\boldsymbol{X}))^2\right] = \int (\widehat{\mu}_t(\boldsymbol{x}) - \mu_t(\boldsymbol{x}))^2 p(\boldsymbol{x}) d\boldsymbol{x}$$

$$= \mathbb{P}(T = t) \int (\widehat{\mu}_t(\boldsymbol{x}) - \mu_t(\boldsymbol{x}))^2 p(\boldsymbol{x} \mid T = t) d\boldsymbol{x} + \sum_{t' \neq t} \mathbb{P}(T = t') \int (\widehat{\mu}_t(\boldsymbol{x}) - \mu_t(\boldsymbol{x}))^2 p(\boldsymbol{x} \mid T = t') d\boldsymbol{x}$$

$$= \mathbb{P}(T = t) R_t + \sum_{t' \neq t} \mathbb{P}(T = t') \int (\widehat{\mu}_t(\boldsymbol{x}) - \mu_t(\boldsymbol{x}))^2 \frac{p(\boldsymbol{x} \mid T = t')}{p(\boldsymbol{x} \mid T = t)} p(\boldsymbol{x} \mid T = t) d\boldsymbol{x}$$

$$= \mathbb{P}(T = t) R_t + \sum_{t' \neq t} \mathbb{P}(T = t') \int (\widehat{\mu}_t(\boldsymbol{x}) - \mu_t(\boldsymbol{x}))^2 \frac{\frac{\mathbb{P}(T = t' \mid \boldsymbol{x}) p(\boldsymbol{x})}{\mathbb{P}(T = t')}}{\frac{\mathbb{P}(T = t \mid \boldsymbol{x}) p(\boldsymbol{x})}{\mathbb{P}(T = t)}} p(\boldsymbol{x} \mid T = t) d\boldsymbol{x} \qquad \text{(Bayes rule)}$$

$$= \mathbb{P}(T = t) R_t + \mathbb{P}(T = t) \sum_{t' \neq t} \int (\widehat{\mu}_t(\boldsymbol{x}) - \mu_t(\boldsymbol{x}))^2 \frac{\mathbb{P}(T = t' \mid \boldsymbol{x})}{\mathbb{P}(T = t \mid \boldsymbol{x})} p(\boldsymbol{x} \mid T = t) d\boldsymbol{x} \qquad (4)$$

$$= \mathbb{P}(T = t) R_t + \mathbb{P}(T = t) \int (\widehat{\mu}_t(\boldsymbol{x}) - \mu_t(\boldsymbol{x}))^2 \frac{\sum_{t' \neq t} \mathbb{P}(T = t' \mid \boldsymbol{x})}{\mathbb{P}(T = t \mid \boldsymbol{x})} p(\boldsymbol{x} \mid T = t) d\boldsymbol{x}$$

$$= \mathbb{P}(T = t) R_t + \mathbb{P}(T = t) \int \frac{1 - r(t, \boldsymbol{x})}{r(t, \boldsymbol{x})} (\widehat{\mu}_t(\boldsymbol{x}) - \mu_t(\boldsymbol{x}))^2 p(\boldsymbol{x} \mid T = t) d\boldsymbol{x}$$

$$= \mathbb{P}(T = t) \int \left(1 + \frac{1 - r(t, \boldsymbol{x})}{r(t, \boldsymbol{x})}\right) (\widehat{\mu}_t(\boldsymbol{x}) - \mu_t(\boldsymbol{x}))^2 p(\boldsymbol{x} \mid T = t) d\boldsymbol{x}$$

$$= \mathbb{E}_{\boldsymbol{X} \sim p(\cdot \mid T = t)}\left[\frac{\mathbb{P}(T = t)}{r(t, \boldsymbol{X})} (\widehat{\mu}_t(\boldsymbol{X}) - \mu_t(\boldsymbol{X}))^2\right].$$

A.2. Consistency of the X-learner

By direct calculations, we show that

$$\mathbb{E}(Z_k^X \mid \boldsymbol{X} = \boldsymbol{x}) = \mathbb{E}\left[\mathbf{1}\{T = t_k\} Y(t_k) \mid \boldsymbol{X} = \boldsymbol{x}\right] - r(t_k, \boldsymbol{x}) \mu_{t_0}(\boldsymbol{x}) + \sum_{l \neq k} r(t_l, \boldsymbol{x})\Big(\mu_{t_k}(\boldsymbol{x}) \qquad (5)$$

$$- \mathbb{E}\left[\mathbf{1}\{T = t_l\} Y(t_l) \mid \boldsymbol{X} = \boldsymbol{x}\right]\Big) + \sum_{l \neq k} r(t_l, \boldsymbol{x})(\mu_{t_l}(\boldsymbol{x}) - \mu_{t_0}(\boldsymbol{x})) \qquad (6)$$

$$= r(t_k, \boldsymbol{x}) \mu_{t_k}(\boldsymbol{x}) - r(t, \boldsymbol{x}) \mu_{t_0}(\boldsymbol{x}) + \sum_{l \neq k} \big(r(t_l, \boldsymbol{x}) \mu_t(\boldsymbol{x}) - r(t_l, \boldsymbol{x}) \mu_{t'}(\boldsymbol{x})\big) \qquad (7)$$

$$+ \sum_{l \neq k} r(t_l, \boldsymbol{x})(\mu_{t_l}(\boldsymbol{x}) - \mu_{t_0}(\boldsymbol{x})) \qquad \text{(by Assumption 3.1)} \qquad (8)$$

$$= r(t_k, \boldsymbol{x}) \mu_{t_k}(\boldsymbol{x}) - r(t_k, \boldsymbol{x}) \mu_{t_0}(\boldsymbol{x}) + \sum_{l \neq k} r(t_l, \boldsymbol{x}) \mu_{t_k}(\boldsymbol{x}) - \sum_{l \neq k} r(t_l, \boldsymbol{x}) \mu_{t_0}(\boldsymbol{x}) \qquad (9)$$

$$= (\mu_{t_k}(\boldsymbol{x}) - \mu_{t_0}(\boldsymbol{x}))\Big(r(t_k, \boldsymbol{x}) + \sum_{l \neq k} r(t_l, \boldsymbol{x})\Big) \qquad (10)$$

$$= \mu_{t_k}(\boldsymbol{x}) - \mu_{t_0}(\boldsymbol{x}) = \tau_k(\boldsymbol{x}). \qquad (11)$$

A.3. Consistency of the naive extension of the X-learner

Let us consider the two random variables $D^{(k)} := Y(t_k) - \mu_{t_0}(\boldsymbol{X})$ and $D^{(0)} := \mu_{t_k}(\boldsymbol{X}) - Y(t_0)$. We have

$$\tau^{(k)}(\boldsymbol{x}) = \mathbb{E}(D^{(k)} \mid \boldsymbol{X} = \boldsymbol{x}) = \mathbb{E}(Y(t_k) - \mu_{t_0}(\boldsymbol{X}) \mid \boldsymbol{X} = \boldsymbol{x})$$
$$= \mathbb{E}[Y(t_k) \mid \boldsymbol{X} = \boldsymbol{x}] - \mu_{t_0}(\boldsymbol{x}) \qquad (12)$$
$$= \mu_{t_k}(\boldsymbol{x}) - \mu_{t_0}(\boldsymbol{x}) = \tau_k(\boldsymbol{x}),$$

and

$$\tau^{(0)}(\boldsymbol{x}) = \mathbb{E}(D^{(0)} \mid \boldsymbol{X} = \boldsymbol{x}) = \mathbb{E}(\mu_{t_k}(\boldsymbol{X}) - Y(t_k) \mid \boldsymbol{X} = \boldsymbol{x})$$
$$= \mu_{t_k}(\boldsymbol{x}) - \mathbb{E}[\mathbb{E}(Y(t_k) \mid \boldsymbol{X} = \boldsymbol{x}] \qquad (13)$$
$$= \mu_{t_k}(\boldsymbol{x}) - \mu_{t_0}(\boldsymbol{x}) = \tau_k(\boldsymbol{x}).$$

Therefore,

$$\tau_k^{(\text{X,nv})}(\boldsymbol{x}) = \frac{r(t_k, \boldsymbol{x})}{r(t_k, \boldsymbol{x}) + r(t_0, \boldsymbol{x})} \tau^{(k)}(\boldsymbol{x}) + \frac{r(t_0, \boldsymbol{x})}{r(t_k, \boldsymbol{x}) + r(t_0, \boldsymbol{x})} \tau^{(0)}(\boldsymbol{x}) = \tau_k(\boldsymbol{x}). \tag{14}$$

A.4. Generalizing the Robinson decomposition

We show first the Neyman-Orthogonality propriety, i.e. $\mathbb{E}\left(\epsilon \mid T, \boldsymbol{X}\right) = 0$. Indeed, for $t \in \mathcal{T}$ and $\boldsymbol{x} \in \mathcal{D}$, we have

$$\begin{aligned}
\mathbb{E}\big[\epsilon \mid T = t, \boldsymbol{X} = \boldsymbol{x}\big] &= \mathbb{E}\big[Y_{\text{obs}} - \mu_T(\boldsymbol{X}) \mid T = t, \boldsymbol{X} = \boldsymbol{x}\big] \\
&= \mathbb{E}\big[Y(t) - \mu_T(\boldsymbol{X}) \mid T = t, \boldsymbol{X} = \boldsymbol{x}\big] \\
&= \mu_t(\boldsymbol{x}) - \mu_t(\boldsymbol{x}) = 0.
\end{aligned} \tag{15}$$

Thus, the observed outcome model satisfies:

$$\begin{aligned}
\mathbb{E}(Y_{\text{obs}} \mid \boldsymbol{X} = \boldsymbol{x}) &= \mathbb{E}\big[\epsilon + \sum_{k=0}^{K} \mathbf{1}\{T = t_k\} \mu_{t_k}(\boldsymbol{X}) \mid \boldsymbol{X} = \boldsymbol{x}\big] \\
&= \mathbb{E}\big[\mathbb{E}[\epsilon \mid T, \boldsymbol{X}] \mid \boldsymbol{X} = \boldsymbol{x}\big] + \sum_{k=0}^{K} \mathbb{E}\big[\mathbf{1}\{T = t_k\} \mid \boldsymbol{X} = \boldsymbol{x}\big] \mu_{t_k}(\boldsymbol{x}) \\
&= \sum_{k=0}^{K} \mu_{t_k}(\boldsymbol{x}) r(t_k, \boldsymbol{x}) = \mu_{t_0}(\boldsymbol{x}) r(t_0, \boldsymbol{x}) + \sum_{k=1}^{K} \mu_{t_k}(\boldsymbol{x}) r(t_k, \boldsymbol{x}) \\
&= \mu_{t_0}(\boldsymbol{x})\big[1 - \sum_{k=1}^{K} r(t_k, \boldsymbol{x})\big] + \sum_{k=1}^{K} \mu_{t_k}(\boldsymbol{x}) r(t_k, \boldsymbol{x}) \\
&= \mu_{t_0}(\boldsymbol{x}) + \sum_{k=1}^{K} r(t, \boldsymbol{x})\left[\mu_{t_k}(\boldsymbol{x}) - \mu_{t_0}(\boldsymbol{x})\right] \\
&= \mu_{t_0}(\boldsymbol{x}) + \sum_{k=1}^{K} r(t_k, \boldsymbol{x}) \tau_k(\boldsymbol{x}) = m(\boldsymbol{x}).
\end{aligned} \tag{16}$$

By gathering both quantities :

$$\begin{aligned}
Y_{\text{obs}} - m(\boldsymbol{X}) &= \sum_{k=0}^{K} \mathbf{1}\{T = t_k\} \mu_{t_k}(\boldsymbol{X}) - \mu_{t_0}(\boldsymbol{X}) - \sum_{k=1}^{K} r(t_k, \boldsymbol{X}) \tau_k(\boldsymbol{X}) + \epsilon \\
&= \mathbf{1}\{T = t_0\} \mu_{t_0}(\boldsymbol{X}) + \sum_{k=1}^{K} \mathbf{1}\{T = t_k\} \mu_{t_k}(\boldsymbol{X}) - \mu_{t_0}(\boldsymbol{X}) - \sum_{k=1}^{K} r(t_k, \boldsymbol{X}) \tau_k(\boldsymbol{X}) + \epsilon \\
&= \big(\mathbf{1}\{T = t_0\} - 1\big) \mu_{t_0}(\boldsymbol{X}) + \sum_{k=1}^{K} \big(\mathbf{1}\{T = t_k\} \mu_{t_k}(\boldsymbol{X}) - r(t_k, \boldsymbol{X}) \tau_k(\boldsymbol{X})\big) + \epsilon \\
&= \sum_{k=1}^{K} \big(\mathbf{1}\{T = t_k\} \mu_{t_k}(\boldsymbol{X}) - r(t_k, \boldsymbol{X}) \tau_k(\boldsymbol{X})\big) - \sum_{k=1}^{K} \mathbf{1}\{T = t_k\} \mu_{t_0}(\boldsymbol{X}) + \epsilon \\
&= \sum_{k=1}^{K} \big(\mathbf{1}\{T = t_k\} \mu_{t_k}(\boldsymbol{X}) - \mathbf{1}\{T = t_k\} \mu_{t_0}(\boldsymbol{X}) - r(t_k, \boldsymbol{X}) \tau_k(\boldsymbol{X})\big) + \epsilon \\
&= \sum_{k=1}^{K} \big[\mathbf{1}\{T = t_k\} - r(t_k, \boldsymbol{X})\big] \tau_k(\boldsymbol{X}) + \epsilon.
\end{aligned} \tag{17}$$

Therefore, we obtain the generalized Robinson decomposition for the multi-treatment regime.

A.5. Solving the generalized R-learner for linear models

For $k = 1, \ldots, K$, we assume that $\overline{\tau}_k$ belongs to the family of linear regression models such that:

$$\mathcal{F} = \left\{ \left\{ \overline{\tau}_k(\boldsymbol{x}) := \beta_{k,0} + \sum_{j=1}^{p-1} \beta_{k,j} f_j(\boldsymbol{x}) \right\}_{k=1}^{K} \Big/ \boldsymbol{\beta}_k = (\beta_{k,0}, \ldots, \beta_{k,p-1})^\top \in \mathbb{R}^p \right\}. \tag{18}$$

f_j are predefined functions (e.g. polynomial functions). It is also possible to use a matrix notation and write $\overline{\tau}_k(\mathbf{X}) = \mathbf{H}\boldsymbol{\beta}_k$ where $\mathbf{H} = (f_j(\boldsymbol{X}_i)) \in \mathbb{R}^{n \times p}$ assumed to be full rank matrix $rank(\mathbf{H}) = p \leq n$.

Let $\overline{Y} = (\overline{Y}_i)_{i=1}^n$ and $\overline{T}_k = (\overline{T}_{i,k})_{i=1}^n$ such that $\overline{Y}_i = Y_{\mathrm{obs},i} - \widehat{m}(\boldsymbol{X}_i)$ and $\overline{T}_{i,k} = \mathbf{1}\{t_i = t_k\} - \widehat{r}(t_k, \boldsymbol{X}_i)$. Let $\epsilon = (\epsilon_i)_{i=1}^n$ denote the vector of errors obtained for the generalized Robinson (1988) decomposition in Proposition 3.3.

We show immediately that $\mathcal{L}$, the generalized R-loss function associated with the mean squared error of ϵ in (2) in the paper, is quadratic with respect to $\boldsymbol{\beta}$. Indeed,

$$\begin{aligned}
\mathcal{L}(\{\overline{\tau}_k\}_{t \neq t^{(0)}}) &= \frac{1}{n}\epsilon^\top \epsilon = \frac{1}{n}\left(\overline{Y} - \sum_{k=1}^K \overline{T}_k \odot (\mathbf{H}\boldsymbol{\beta}_k) \right)^\top \left(\overline{Y} - \sum_{k=1}^K \overline{T}_k \odot (\mathbf{H}\boldsymbol{\beta}_k) \right) \\
&= \frac{1}{n}\left[\overline{Y}^\top \overline{Y} - 2\sum_{k=1}^K \overline{Y}^\top \left(\overline{T}_k \odot (\mathbf{H}\boldsymbol{\beta}_k) \right) + \sum_{k,k'=1}^K \left(\overline{T}_k \odot (\mathbf{H}\boldsymbol{\beta}_k) \right)^\top \left(\overline{T}_{k'} \odot (\mathbf{H}\boldsymbol{\beta}_{k'}) \right) \right] \\
&= \frac{1}{n}\left(\overline{Y}^\top \overline{Y} - 2\sum_{k=1}^K \overline{Y}^\top \mathbf{D}_{\overline{T}_k} \mathbf{H}\boldsymbol{\beta}_k + \sum_{k,k'=1}^K \boldsymbol{\beta}_k^\top \mathbf{H}^\top \mathbf{D}_{\overline{T}_k} \mathbf{D}_{\overline{T}_{k'}} \mathbf{H}\boldsymbol{\beta}_{k'} \right),
\end{aligned} \tag{19}$$

where $\odot$ is the Hadamard product (element-wise product). The last line holds because $\overline{T}_k \odot (\mathbf{H}\boldsymbol{\beta}_k) = \mathbf{D}_{\overline{T}_k}\mathbf{H}\boldsymbol{\beta}_k$ with $\mathbf{D}_{\overline{T}_k}$ is the diagonal matrix of the vector $\overline{T}_k = (\overline{T}_{i,k})_{i=1}^n$

By differentiating $\partial\mathcal{L}/\partial\boldsymbol{\beta}_k = 0$ for $k = 1, \ldots, K$:

$$\begin{cases}
-\boldsymbol{a}_1 + \mathbf{B}_1\widehat{\boldsymbol{\beta}}_1 + \sum_{k=2}^K \mathbf{C}_{1k}\widehat{\boldsymbol{\beta}}_k = 0 \\
\qquad \vdots \qquad \vdots \qquad \vdots \quad = 0 \\
-\boldsymbol{a}_K + \sum_{k=1}^K \mathbf{C}_{Kk}\widehat{\boldsymbol{\beta}}_k + \mathbf{B}_K\widehat{\boldsymbol{\beta}}_K = 0
\end{cases} \tag{20}$$

$$\iff \begin{bmatrix} \mathbf{B}_1 & \mathbf{C}_{12} & \cdots & \mathbf{C}_{1K} \\ \mathbf{C}_{21} & \mathbf{B}_2 & \cdots & \mathbf{C}_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_{K1} & \mathbf{C}_{K2} & \cdots & \mathbf{B}_K \end{bmatrix} \begin{bmatrix} \widehat{\boldsymbol{\beta}}_1 \\ \widehat{\boldsymbol{\beta}}_2 \\ \vdots \\ \widehat{\boldsymbol{\beta}}_K \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_K \end{bmatrix}, \tag{21}$$

where

$$\boldsymbol{a}_j = \frac{1}{n}\mathbf{H}^\top \mathbf{D}_{\overline{T}_j}\overline{Y} \in \mathbb{R}^p, \tag{22}$$

$$\mathbf{B}_j = \frac{1}{n}\mathbf{H}^\top \mathbf{D}_{\overline{T}_j}^2 \mathbf{H} \in \mathbb{R}^{p \times p}, \tag{23}$$

$$\mathbf{C}_{ij} = \frac{1}{n}\mathbf{H}^\top \mathbf{D}_{\overline{T}_i} \mathbf{D}_{\overline{T}_j} \mathbf{H} \in \mathbb{R}^{p \times p}. \tag{24}$$

Let $\boldsymbol{\beta} = \left(\boldsymbol{\beta}_1^\top, \ldots, \boldsymbol{\beta}_K^\top \right)^\top \in \mathbb{R}^{K \times p}$ and consider the block matrix $\mathbf{A}$ defined as

$$\mathbf{A} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{C}_{12} & \cdots & \mathbf{C}_{1K} \\ \mathbf{C}_{21} & \mathbf{B}_2 & \cdots & \mathbf{C}_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_{K1} & \mathbf{C}_{K2} & \cdots & \mathbf{B}_K \end{bmatrix}. \tag{25}$$

16

The matrix $\mathbf{A}$ is real symmetric and satisfies:

$$
\begin{aligned}
\boldsymbol{\beta}^\top \mathbf{A} \boldsymbol{\beta} &= \sum_{1 \le k,l \le K} \boldsymbol{\beta}_k^\top \mathbf{H}^\top \mathbf{D}_{\overline{T}_k} \mathbf{D}_{\overline{T}_l} \mathbf{H} \boldsymbol{\beta}_l \\
&= \left\| \sum_{k=1}^{K} \mathbf{D}_{\overline{T}_k} \mathbf{H} \boldsymbol{\beta}_k \right\|^2 \ge 0.
\end{aligned}
\tag{26}
$$

This result shows that $\mathbf{A}$ is positive semi-definite, all its eigenvalues are nonnegative and also proves the existence of a minimizer $\widehat{\boldsymbol{\beta}}$ to the loss function $\mathcal{L}$. However, this is not sufficient to prove the uniqueness of the solution as one cannot prove all eigenvalues are positive.

The solution $\widehat{\boldsymbol{\beta}}$ to Problem (4.3) in the main paper with the minimal norm is given by

$$
\widehat{\boldsymbol{\beta}} = \mathbf{A}^+ \boldsymbol{a},
\tag{27}
$$

where $\mathbf{A}^+$ is the Moore–Penrose inverse of $\mathbf{A}$ and $\boldsymbol{a} = \left(\boldsymbol{a}_1^\top, \ldots, \boldsymbol{a}_K^\top \right)^\top$.

B. Theoretical analysis of the error bounds.

B.1. Error estimation of pseudo-outcome meta-learners.

Step 0. Set-up of the theorem
In the following subsection, we will analyze the error estimation of each two-step meta-learner. Given Assumption 5.1 stating that the observations are generated from a function f respecting the two causal assumptions (3.1-3.2), each unit i has the following observed and potential outcomes

$$
\begin{aligned}
Y_i(t_k) &= f(t_k, \boldsymbol{X}_i) + \varepsilon_i(t_k), \\
Y_i(t_0) &= f(t_0, \boldsymbol{X}_i) + \varepsilon_i(t_0).
\end{aligned}
\tag{28}
$$

where $\varepsilon_i(t)$ are i.i.d. Gaussian $\mathcal{N}(0, \sigma^2)$ and independent of $(T_i, \boldsymbol{X}_i)_{i=1}^n$. As a consequence, the noise $(\epsilon_i)_{i=1}^n = (\varepsilon_i(T_i))_{i=1}^n$ is also Gaussian $\mathcal{N}(0, \sigma^2)$ and is independent of $(T_i, \boldsymbol{X}_i)_{i=1}^n$.

The CATE model τ_k for each $k = 1, \ldots, K$ can be written as:

$$
\begin{aligned}
\tau_k(\boldsymbol{x}) &= \mathbb{E}(Y(t_k) - Y(t_0) \mid \boldsymbol{X} = \boldsymbol{x}) \\
&= \mathbb{E}(f(t_k, \boldsymbol{X}) - f(t_0, \boldsymbol{X}) + \epsilon^* \mid \boldsymbol{X} = \boldsymbol{x}) \\
&= f(t_k, \boldsymbol{x}) - f(t_0, \boldsymbol{x})
\end{aligned}
\tag{29}
$$

where ϵ^* is a noise independent of $\boldsymbol{X}$ (and T) and satisfying $\mathbb{E}(\epsilon^*) = 0$.

Under the assumption 5.2, we write $\tau_k(\boldsymbol{X}) = f(t_k, \boldsymbol{X}) - f(t_0, \boldsymbol{X}) = \mathbf{H} \boldsymbol{\beta}_k^*$ where $\boldsymbol{\beta}_k^* = \boldsymbol{\beta}_{t_k} - \boldsymbol{\beta}_{t_0}$ and $\mathbf{H} = (\mathbf{H}_{ij}) \in \mathbb{R}^{n \times p}$ is the regression matrix, assumed to be full rank matrix, such that $\mathbf{H}_{ij} = f_j(\boldsymbol{X}_i)$ for $i = 1, \ldots, n$ and $j = 0, \ldots, p-1$. With pseudo-outcome meta-learners, we consider a random variable Z_k for a fixed t_k such that

$$
Z_{k,i} = A_{t_k}(T_i, \boldsymbol{X}_i) Y_{\mathrm{obs},i} + B_{t_k}(T_i, \boldsymbol{X}_i), \quad i = 1, \ldots, n,
$$

where the functions $A_{t_k}(T, \boldsymbol{X})$ and $B_{t_k}(T, \boldsymbol{X})$ are given for each pseudo-outcome meta-learner.

Step 1. Identification of $\widehat{\boldsymbol{\beta}}_k$ and $\boldsymbol{\beta}_k^*$
The regression coefficients $\widehat{\boldsymbol{\beta}}_k$ are given by the Ordinary Least Squares (OLS) method

$$
\widehat{\boldsymbol{\beta}}_k = \left(\mathbf{H}^\top \mathbf{H} \right)^{-1} \mathbf{H}^\top \boldsymbol{z}_k,
\tag{30}
$$

where $\boldsymbol{z}_k = (Z_{k,i})_{1 \le i \le n}$. Thus,

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}}_k &= \left(\mathbf{H}^\top \mathbf{H}\right)^{-1} \mathbf{H}^\top \boldsymbol{z}_k \\
&= \left(\mathbf{H}^\top \mathbf{H}\right)^{-1} \mathbf{H}^\top \left(A_{t_k}(T_i, \boldsymbol{X}_i) Y_{\mathrm{obs},i} + B_{t_k}(T_i, \boldsymbol{X}_i)\right)_{i=1}^n \\
&= \left(\mathbf{H}^\top \mathbf{H}\right)^{-1} \mathbf{H}^\top \left(A_{t_k}(T_i, \boldsymbol{X}_i) f(T_i, \boldsymbol{X}_i) + B_{t_k}(T_i, \boldsymbol{X}_i) + A_{t_k}(T_i, \boldsymbol{X}_i)\epsilon_i\right)_{i=1}^n \\
&= \left(\mathbf{H}^\top \mathbf{H}\right)^{-1} \mathbf{H}^\top \left(\tau_k(\boldsymbol{x}) + A_{t_k}(T_i, \boldsymbol{X}_i) f(T_i, \boldsymbol{X}_i) - \tau_k(\boldsymbol{x}) + B_{t_k}(T_i, \boldsymbol{X}_i) + A_{t_k}(T_i, \boldsymbol{X}_i)\epsilon_i\right)_{i=1}^n \\
&= \left(\mathbf{H}^\top \mathbf{H}\right)^{-1} \mathbf{H}^\top \left(\mathbf{H}\boldsymbol{\beta}_k^* + A_{t_k}(T_i, \boldsymbol{X}_i) f(T_i, \boldsymbol{X}_i) - \tau_k(\boldsymbol{x}) + B_{t_k}(T_i, \boldsymbol{X}_i) + A_{t_k}(T_i, \boldsymbol{X}_i)\epsilon_i\right)_{i=1}^n \\
&= \boldsymbol{\beta}_k^* + \left(\mathbf{H}^\top \mathbf{H}\right)^{-1} \mathbf{H}^\top \left(A_{t_k}(T_i, \boldsymbol{X}_i) f(T_i, \boldsymbol{X}_i) - \tau_k(\boldsymbol{x}) + B_{t_k}(T_i, \boldsymbol{X}_i) + A_{t_k}(T_i, \boldsymbol{X}_i)\epsilon_i\right)_{i=1}^n \\
&= \boldsymbol{\beta}_k^* + \left(\mathbf{H}^\top \mathbf{H}\right)^{-1} \mathbf{H}^\top \tilde{\boldsymbol{\epsilon}}_k
\end{aligned}
$$

where $\tilde{\epsilon}_{k,i} = \psi_k(T_i, \boldsymbol{X}_i) + A_{t_k}(T_i, \boldsymbol{X}_i)\epsilon_i$ and $\psi_k(T_i, \boldsymbol{X}_i) = A_{t_k}(T_i, \boldsymbol{X}_i) f(T_i, \boldsymbol{X}_i) - \tau_k(\boldsymbol{X}_i) + B_{t_k}(T_i, \boldsymbol{X}_i)$ to simplify notations.

Let us consider the random vector $\boldsymbol{Z}_k^{(n)}$ such that

$$
\boldsymbol{Z}_k^{(n)} = \left(\frac{1}{n}(\mathbf{H}^\top \tilde{\boldsymbol{\epsilon}}_k)_1, \ldots, \frac{1}{n}(\mathbf{H}^\top \tilde{\boldsymbol{\epsilon}}_k)_p, \frac{1}{n}(\mathbf{H}^\top \mathbf{H})_{11}, \ldots, \frac{1}{n}(\mathbf{H}^\top \mathbf{H})_{pp}\right)^\top \in \mathbb{R}^{p+p^2}, \tag{31}
$$

that allows us to write $\widehat{\boldsymbol{\beta}}_k$ as

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}}_k &= \boldsymbol{\beta}_k^* + \left(\mathbf{H}^\top \mathbf{H}\right)^{-1} \mathbf{H}^\top \tilde{\boldsymbol{\epsilon}}_k \\
&= \boldsymbol{\beta}_k^* + \left(\frac{1}{n}\mathbf{H}^\top \mathbf{H}\right)^{-1} \left(\frac{1}{n}\mathbf{H}^\top \tilde{\boldsymbol{\epsilon}}_k\right) \\
&= \boldsymbol{\beta}_k^* + \phi(\boldsymbol{Z}_k^{(n)}),
\end{aligned} \tag{32}
$$

where $\phi : \mathbb{R}^{p+p^2} \longrightarrow \mathbb{R}^p$ is a $\mathcal{C}^1$-function.

Step 2. The asymptotic behaviour of the OLS estimator's mean and covariance

In order to apply the Central Limit Theorem (CLT) later, we show that the vector $\boldsymbol{Z}_k^{(n)}$ can be written as sum of *i.i.d.* random vectors $\boldsymbol{Z}_{k,i}$.

$$
\begin{aligned}
\boldsymbol{Z}_k^{(n)} &= \left(\frac{1}{n}(\mathbf{H}^\top \tilde{\boldsymbol{\epsilon}}_k)_1, \ldots, \frac{1}{n}(\mathbf{H}^\top \tilde{\boldsymbol{\epsilon}}_k)_p, \frac{1}{n}(\mathbf{H}^\top \mathbf{H})_{11}, \ldots, \frac{1}{n}(\mathbf{H}^\top \mathbf{H})_{pp}\right)^\top \in \mathbb{R}^{p+p^2} \\
&= \left(\frac{1}{n}\sum_{i=1}^n \mathbf{H}_{i1}\tilde{\epsilon}_{k,i}, \ldots, \mathbf{H}_{ip}\tilde{\epsilon}_{k,i}, \frac{1}{n}\sum_{i=1}^n \mathbf{H}_{i1}\mathbf{H}_{i1}, \ldots, \frac{1}{n}\sum_{i=1}^n \mathbf{H}_{ip}\mathbf{H}_{ip}\right)^\top \\
&= \frac{1}{n}\sum_{i=1}^n \left(\mathbf{H}_{i1}\tilde{\epsilon}_{k,i}, \ldots, \mathbf{H}_{ip}\tilde{\epsilon}_{k,i}, \mathbf{H}_{i1}\mathbf{H}_{i1}, \ldots, \mathbf{H}_{ip}\mathbf{H}_{ip}\right)^\top = \frac{1}{n}\sum_{i=1}^n \boldsymbol{Z}_{k,i}.
\end{aligned} \tag{33}
$$

The mean $\boldsymbol{m}$ of the vector $\boldsymbol{Z}_k^{(n)}$ satisfies

$$
\begin{aligned}
\boldsymbol{m} = \mathbb{E}(\boldsymbol{Z}_k^{(n)}) &= \frac{1}{n}\sum_{i=1}^n \mathbb{E}(\boldsymbol{Z}_{k,i}) = \mathbb{E}(\boldsymbol{Z}_{k,i}) \\
&= \left(h_1, \ldots, h_p, F_{11}, \ldots, F_{pp}\right)^\top,
\end{aligned} \tag{34}
$$

where, for $j, j' = 1, \ldots, p$,

$$
\begin{aligned}
h_j &= \mathbb{E}\left[f_{j-1}(\boldsymbol{X})\left(\psi_k(T, \boldsymbol{X}) + A_{t_k}(T, \boldsymbol{X})\epsilon\right)\right] = \mathbb{E}\left(f_{j-1}(\boldsymbol{X})\psi_k(T, \boldsymbol{X})\right) \\
F_{jj'} &= \mathbb{E}\left(f_{j-1}(\boldsymbol{X}) f_{j'-1}(\boldsymbol{X})\right).
\end{aligned} \tag{35}
$$

The covariance matrix $\mathbf{C}$ of $\boldsymbol{Z}_k^{(n)}$ has entries

$$
\mathbf{C}_{jj'} = \mathrm{Cov}\left(\boldsymbol{Z}_j^{(k)}, \boldsymbol{Z}_{j'}^{(k)}\right) = \mathbb{E}(\boldsymbol{Z}_j^{(k)}, \boldsymbol{Z}_{j'}^{(k)}) - \mathbb{E}(\boldsymbol{Z}_j^{(k)})\mathbb{E}(\boldsymbol{Z}_{j'}^{(k)})
$$

$$
= \begin{cases}
\mathbb{E}\left(f_{j-1}(\boldsymbol{X})f_{j'-1}(\boldsymbol{X})\left(\psi_k(T,\boldsymbol{X}) + A_{t_k}(T,\boldsymbol{X})\epsilon\right)^2\right) - h_j h_{j'} & \text{if } j,j' \in \{1,\ldots,p\} \\
\mathbb{E}\left(f_{\tilde{k}-1}(\boldsymbol{X})f_{\tilde{k}'-1}(\boldsymbol{X})f_{l-1}(\boldsymbol{X})f_{l'-1}(\boldsymbol{X})\right) - F_{kk'}F_{ll'} & \text{if } j,j' \in \{p+1,\ldots,p^2\} \\
\mathbb{E}\left(f_{\tilde{k}-1}(\boldsymbol{X})f_{\tilde{k}'-1}(\boldsymbol{X})f_{j-1}(\boldsymbol{X})\left(\psi_k(T,\boldsymbol{X}) + A_{t_k}(T,\boldsymbol{X})\epsilon\right)\right) - h_j F_{kk'} & \text{otherwise.}
\end{cases}
$$

$$
= \begin{cases}
\mathbb{E}\left(f_{j-1}(\boldsymbol{X})f_{j'-1}(\boldsymbol{X})\psi_k^2(T,\boldsymbol{X})\right) + \sigma^2 \mathbb{E}\left(f_{j-1}(\boldsymbol{X})f_{j'-1}(\boldsymbol{X})A_{t_k}^2(T,\boldsymbol{X})\right) - h_j h_{j'} & \text{if } j,j' \in \{1,\ldots,p\} \\
\mathbb{E}\left(f_{\tilde{k}-1}(\boldsymbol{X})f_{\tilde{k}'-1}(\boldsymbol{X})f_{l-1}(\boldsymbol{X})f_{l'-1}(\boldsymbol{X})\right) - F_{\tilde{k}\tilde{k}'}F_{ll'} & \text{if } j,j' \in \{p+1,\ldots,p^2\} \\
\mathbb{E}\left(f_{\tilde{k}-1}(\boldsymbol{X})f_{\tilde{k}'-1}(\boldsymbol{X})f_{j-1}(\boldsymbol{X})\psi_k(T,\boldsymbol{X})\right) - h_j F_{\tilde{k}\tilde{k}'} & \text{otherwise,}
\end{cases}
\tag{36}
$$

where $\tilde{k}, \tilde{k}' = \eta^{-1}(j)$ (respectively, $l, l' = \eta^{-1}(j')$) such that η is the correspondence indexes map between $\boldsymbol{m}$ and F in $\boldsymbol{m}_j = F_{\tilde{k}\tilde{k}'}$ (respectively, $\boldsymbol{m}_{j'} = F_{ll'}$) when $j \geq p+1$ (respectively, when $j' \geq p+1$).

By considering now the vector

$$
\boldsymbol{S}^{(n)} = \sqrt{n}\left(\boldsymbol{Z}_k^{(n)} - \boldsymbol{m}\right) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \left(\boldsymbol{Z}_{k,i} - \boldsymbol{m}\right),
\tag{37}
$$

one can show by the multivariate CLT that

$$
\boldsymbol{S}^{(n)} = \sqrt{n}\left(\boldsymbol{Z}_k^{(n)} - \boldsymbol{m}\right) \xrightarrow{\mathcal{L}} \mathcal{N}(\boldsymbol{0}, \mathbf{C}).
\tag{38}
$$

This allows us to write $\widehat{\boldsymbol{\beta}}_k$ as function of $\boldsymbol{S}^{(n)}$ and $\boldsymbol{m}$. Indeed,

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}}_k &= \boldsymbol{\beta}_k^* + \left(\mathbf{H}^\top\mathbf{H}\right)^{-1}\mathbf{H}^\top\tilde{\boldsymbol{\epsilon}} \\
&= \boldsymbol{\beta}_k^* + \phi(\boldsymbol{Z}^{(n)}) \\
&= \boldsymbol{\beta}_k^* + \phi\left(\boldsymbol{m} + \boldsymbol{S}^{(n)}/\sqrt{n}\right) \\
&= \boldsymbol{\beta}_k^* + \Phi(\boldsymbol{S}^{(n)}, \boldsymbol{m}),
\end{aligned}
\tag{39}
$$

where $\Phi : \mathbb{R}^{p+p^2} \times \mathbb{R}^{p+p^2} \longrightarrow \mathbb{R}^p$ is also $\mathcal{C}^1$-function.

Since $\sqrt{n}\left(\boldsymbol{S}^{(n)} - \boldsymbol{0}\right) \xrightarrow{\mathcal{L}} \mathcal{N}(\boldsymbol{0}, \mathbf{C})$, one obtains by the Delta method

$$
\sqrt{n}\left[\Phi(\boldsymbol{S}^{(n)}, \boldsymbol{m}) - \Phi(\boldsymbol{0}, \boldsymbol{m})\right] \xrightarrow{\mathcal{L}} \mathcal{N}\left(\boldsymbol{0}, J_\Phi^{(1)}(\boldsymbol{0}, \boldsymbol{m})^\top \mathbf{C} J_\Phi^{(1)}(\boldsymbol{0}, \boldsymbol{m})\right),
\tag{40}
$$

where $J_\Phi^{(1)}(\boldsymbol{0}, \boldsymbol{m})$ is the Jacobian matrix at the first $p+p^2$ coordinates of Φ at $(\boldsymbol{0}, \boldsymbol{m})$.

By denoting $\boldsymbol{g}_n$, a Gaussian noise with zero-mean and covariance matrix $\mathbf{C}' = J_\Phi^{(1)}(\boldsymbol{0}, \boldsymbol{m})^\top \mathbf{C} J_\Phi^{(1)}(\boldsymbol{0}, \boldsymbol{m})$, the previous equation is equivalent to

$$
\widehat{\boldsymbol{\beta}}_k = \boldsymbol{\beta}_k^* + \Phi(\boldsymbol{S}_n, \boldsymbol{m}) \approx \boldsymbol{\beta}_k^* + \Phi(\boldsymbol{0}, \boldsymbol{m}) + \boldsymbol{g}_n/\sqrt{n}.
\tag{41}
$$

For n large, the expansions of the first moments is of the form:

$$
\mathbb{E}(\widehat{\boldsymbol{\beta}}_k) \approx \boldsymbol{\beta}_k^* + \Phi(\boldsymbol{0}, \boldsymbol{m}).
\tag{42}
$$

and, the asymptotic variance is also of the form:

$$
\mathbb{V}(\widehat{\boldsymbol{\beta}}_k) \approx \frac{1}{n}\, J_\Phi^{(1)}(\boldsymbol{0}, \boldsymbol{m})^\top \mathbf{C} J_\Phi^{(1)}(\boldsymbol{0}, \boldsymbol{m}).
\tag{43}
$$

This result holds whether the nuisance parameters in A_t and B_t are well-specified or not, so there is no guarantee that $\Phi(\boldsymbol{0}, \boldsymbol{m}) = 0$ and the estimator $\widehat{\boldsymbol{\beta}}_k$ may be biased.

In the following, we assume that the nuisance parameters in A_t and B_t are well-specified i.e. $\mathbb{E}(\psi_k(T, \boldsymbol{X})) \mid \boldsymbol{X} = \boldsymbol{x}) = 0$ in such way that $\mathbb{E}(Z_k \mid \boldsymbol{X} = \boldsymbol{x}) = \tau_k(\boldsymbol{x})$, or equivalently, $\mathbb{E}(\mathbf{H}^\top \tilde{\boldsymbol{\epsilon}}_k) = \mathbf{0}$. Consequently, the estimator of $\widehat{\boldsymbol{\beta}}_k$ is unbiased. In this case, computing the variance $\mathbb{V}(\widehat{\boldsymbol{\beta}}_k)$ becomes much easier and more explicit.

On the one hand, by the multivariate Central Theorem Limit (CTL)

$$\frac{1}{\sqrt{n}} \mathbf{H}^\top \tilde{\boldsymbol{\epsilon}}_k \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}) \tag{44}$$

which is equivalent to

$$\frac{1}{\sqrt{n}} \mathbf{H}^\top \tilde{\boldsymbol{\epsilon}}_k \approx \boldsymbol{g}_n, \tag{45}$$

where $\boldsymbol{g}_n$ is a Gaussian noise with zero-mean and covariance matrix of $\boldsymbol{\Sigma}$ with entries

$$\begin{aligned}
\boldsymbol{\Sigma}_{jj'} &= \mathbb{E}\big[f_j(\boldsymbol{X})f_{j'}(\boldsymbol{X})\big(\psi_k(T, \boldsymbol{X}) + A_{t_k}(T, \boldsymbol{X})\epsilon\big)^2\big] \\
&= \mathbb{E}\big(f_j(\boldsymbol{X})f_{j'}(\boldsymbol{X})\psi_k^2(T, \boldsymbol{X})\big) + \sigma^2 \mathbb{E}\big(f_j(\boldsymbol{X})f_{j'}(\boldsymbol{X})A_{t_k}^2(T, \boldsymbol{X})\big).
\end{aligned} \tag{46}$$

On the other hand, by the law of large numbers (LLN), we have $1/n(\mathbf{H}^\top \mathbf{H}) \xrightarrow{a.s} \mathbf{F}$, thus $1/n(\mathbf{H}^\top \mathbf{H}) \xrightarrow{P} \mathbf{F}$. Since $\mathbf{F}$ is invertible, then

$$n(\mathbf{H}^\top \mathbf{H})^{-1} \xrightarrow{P} \mathbf{F}^{-1}, \tag{47}$$

where $\mathbf{F} = (F_{jj'})_{0 \leq j, j' \leq p-1}$ and $F_{jj'} = \mathbb{E}(f_j(\boldsymbol{X})f_{j'}(\boldsymbol{X}))$.

By Slutsky's theorem,

$$\begin{aligned}
\sqrt{n}(\widehat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k^*) &= n(\mathbf{H}^\top \mathbf{H})^{-1} \cdot 1/\sqrt{n}\, \mathbf{H}^\top \tilde{\boldsymbol{\epsilon}} \\
&\xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{F}^{-1}\boldsymbol{\Sigma}\mathbf{F}^{-1}).
\end{aligned} \tag{48}$$

We can deduce that the asymptotic mean and variance are of the form

$$\begin{aligned}
\mathbb{E}(\widehat{\boldsymbol{\beta}}_k) &= \boldsymbol{\beta}_k^*, \\
\mathbb{V}(\widehat{\boldsymbol{\beta}}_k) &\approx \frac{1}{n} \mathbf{F}^{-1}\boldsymbol{\Sigma}\mathbf{F}^{-1}.
\end{aligned} \tag{49}$$

Step 3. Obtaining the error upper bound

The determinant of the variance matrix, also known as the generalized variance by Wilks (1932; 1967) is usually used as a scalar measure of overall multidimensional scatter and can be useful to compare the variance of each meta-learner.

In our case, comparing the generalized variance is equivalent to comparing $\det\left(\frac{1}{n}\boldsymbol{\Sigma}\right)$ of each pseudo-outcome meta-learner since

$$\det\left(\mathbb{V}(\widehat{\boldsymbol{\beta}}_k)\right) = \left(\det \mathbf{F}^{-1}\right)^2 \det\left(\frac{1}{n}\boldsymbol{\Sigma}\right) = \frac{1}{\left(\det \mathbf{F}\right)^2} \det\left(\frac{1}{n}\boldsymbol{\Sigma}\right), \tag{50}$$

with, obviously, $\det(\boldsymbol{\Sigma}) > 0$ because $\boldsymbol{\Sigma}$ is symmetric positive definite.

In some cases, the polynomials f_j are chosen to be orthonormal with respect to the distribution of $\boldsymbol{X}$ (e.g. Polynomials Chaos (Sudret, 2008)). A consequence of this choice implies that $\mathbf{F}$ is the identity matrix. Therefore, in the following, we focus on computing and bounding $\boldsymbol{\Sigma}$ terms. The assumptions (3.1-5.2) and the following lemma will be used for this purpose.

Lemma B.1. *If $X_1, \ldots, X_m$ is a sequence of random variables and $b > 1$, then*

$$\begin{aligned}
\left|\mathbb{E}\left[\left(\sum_{i=1}^m X_i\right)^2\right]\right| &\leq m \sum_{i=1}^m \mathbb{E}\left[\left|X_i^2\right|\right], \\
\left|\mathbb{E}\left[\left(\sum_{i=1}^m X_i\right)^b\right]\right| &\leq m^{(b-1)} \sum_{i=1}^m \mathbb{E}\left[\left|X_i^b\right|\right].
\end{aligned} \tag{51}$$

Proof. The first inequality is obtained by Cauchy-Schwarz, whereas the second inequality can be proved by Jensen inequality. Indeed, for $b > 1$, the function $x \mapsto x^b$ is convex for $x > 0$ and

$$\left| \frac{\sum_{i=1}^m X_i}{m} \right|^b \leq \frac{\sum_{i=1}^m |X_i|^b}{m}. \tag{52}$$

Therefore,

$$\left| \mathbb{E}\left[\left(\sum_{i=1}^m X_i \right)^b \right] \right| \leq \mathbb{E}\left[\left| \sum_{i=1}^m X_i \right|^b \right] \leq m^{(b-1)} \sum_{i=1}^m \mathbb{E}\left[|X_i^b| \right]. \tag{53}$$

$\square$

In the following and by Assumption 5.1, $f_j(\boldsymbol{X}) \in L^a$ i.e. $f_j(\boldsymbol{X})$ has all possible finite moments for all $j \in \{0, \ldots, p-1\}$. Moreover, there exists $C > 0$ such that:

$$\forall t \in \mathcal{T}, \forall \boldsymbol{x} \in \mathcal{D} : \ |f(t, \boldsymbol{x})| \leq C. \tag{54}$$

B.1.1. ERROR ESTIMATION OF THE M-LEARNER

Let $a, b > 1$ such that $1/a + 1/b = 1$. We denote $\delta_{jj'}^{(a)} = \left| \mathbb{E}\left(f_j^a(\boldsymbol{X}) f_{j'}^a(\boldsymbol{X}) \right) \right|^{1/a}$. By Hölder inequality we show that for the M-learner:

$$
\begin{aligned}
\left| \mathbb{E}\left(f_j(\boldsymbol{X}) f_{j'}(\boldsymbol{X}) \psi_k^2(T, \boldsymbol{X}) \right) \right| &\leq \left| \mathbb{E}\left(f_j^a(\boldsymbol{X}) f_{j'}^a(\boldsymbol{X}) \right) \right|^{1/a} \cdot \left| \mathbb{E}\left(\psi_k^{2b}(T, \boldsymbol{X}) \right) \right|^{1/b} \quad \text{(Hölder)} \\
&\leq \delta_{jj'}^{(a)} \left(2^{2b-1} \mathbb{E}\left[\left(\frac{\mathbf{1}\{T = t_k\}}{r(t_k, \boldsymbol{X})} - 1 \right)^{2b} f^{2b}(t_k, \boldsymbol{X}) \right. \right. \\
&\qquad + \left. \left. \left(\frac{\mathbf{1}\{T = t_k\}}{r(t_k, \boldsymbol{X})} - 1 \right)^{2b} f^{2b}(t_k, \boldsymbol{X}) \right] \right)^{1/b} \\
&\qquad\qquad \text{(Lemma B.1 with } m = 2) \\
&\leq 2^{(2b-1)/b} \delta_{jj'}^{(a)} \left(\mathbb{E}\left[2^{2b-1} \left(\frac{\mathbf{1}\{T = t_k\}}{r^{2b}(t, \boldsymbol{X})} + 1 \right) f^{2b}(t_k, \boldsymbol{X}) \right] \right. \\
&\qquad + \left. \mathbb{E}\left[2^{2b-1} \left(\frac{\mathbf{1}\{T = t_k\}}{r^{2b}(t_k, \boldsymbol{X})} + 1 \right) f^{2b}(t_k, \boldsymbol{X}) \right] \right)^{1/b} \quad \text{(Lemma B.1)} \\
&\leq 2^{2(2b-1)/b} \delta_{jj'}^{(a)} \left(\mathbb{E}\left[\mathbb{E}\left(\frac{\mathbf{1}\{T = t_k\}}{r^{2b}(t, \boldsymbol{X})} + 1 \right) \mid \boldsymbol{X} \right) f^{2b}(t_k, \boldsymbol{X}) \right] \\
&\qquad + \left. \mathbb{E}\left[\mathbb{E}\left(\frac{\mathbf{1}\{T = t_k\}}{r^{2b}(t_k, \boldsymbol{X})} + 1 \right) \mid \boldsymbol{X} \right) f^{2b}(t_k, \boldsymbol{X}) \right] \right)^{1/b} \tag{55} \\
&\leq 2^{2(2b-1)/b} \delta_{jj'}^{(a)} \left(\mathbb{E}\left[\left(\frac{1}{r^{2b-1}(t, \boldsymbol{X})} + 1 \right) f^{2b}(t_k, \boldsymbol{X}) \right] \right. \\
&\qquad + \left. \mathbb{E}\left[\left(\frac{1}{r^{2b-1}(t_k, \boldsymbol{X})} + 1 \right) f^{2b}(t_k, \boldsymbol{X}) \right] \right)^{1/b} \\
&\leq 2^{2(2b-1)/b} \delta_{jj'}^{(a)} \left(\frac{1}{r_{\min}^{2b-1}} + 1 \right)^{1/b} \left(C^{2b} + C^{2b} \right)^{1/b} \quad \text{(Bounding } r \text{ and } f) \\
&\leq 2^{2(2b-1)/b} \delta_{jj'}^{(a)} \left(\frac{1}{r_{\min}^{2b-1}} + \frac{1}{r_{\min}^{2b-1}} \right)^{1/b} 2^{1/b} C^b \\
&\leq 2^{2(2b-1)/b} \delta_{jj'}^{(a)} \frac{2^{1/b}}{r_{\min}^{(2b-1)/b}} 2^{1/b} C^b \\
&\leq 2^4 \, \delta_{jj'}^{(a)} \frac{1}{r_{\min}^{(2b-1)/b}} C^b = \frac{16}{r_{\min}^{(2b-1)/b}} \delta_{jj'}^{(a)} C^b.
\end{aligned}
$$

On the other term, one obtains similarly:

$$
\begin{aligned}
\left|\mathbb{E}\big(f_j(\boldsymbol{X})f_{j'}(\boldsymbol{X})A_{t_k}^2(T,\boldsymbol{X})\big)\right| &\leq \left|\mathbb{E}\big(f_j^a(\boldsymbol{X})f_{j'}^a(\boldsymbol{X})\big)\right|^{1/a}\cdot\left|\mathbb{E}\big(A_{t_k}^{2b}(T,\boldsymbol{X})\big)\right|^{1/b} \quad\text{(Hölder)}\\
&\leq \delta_{jj'}^{(a)}\left|\mathbb{E}\big(A_{t_k}^{2b}(T,\boldsymbol{X})\big)\right|^{1/b}\\
&\leq \delta_{jj'}^{(a)}\left(2^{2b-1}\,\mathbb{E}\Big(\frac{\mathbf{1}\{T=t_k\}}{r(t_k,\boldsymbol{X})}\Big)^{2b}+\mathbb{E}\Big(\frac{\mathbf{1}\{T=t_k\}}{r(t_k,\boldsymbol{X})}\Big)^{2b}\right)^{1/b} \quad\text{(Lemma B.1)}\\
&\leq 2^{(2b-1)/b}\sigma^2\delta_{jj'}^{(a)}\left(\mathbb{E}\Big(\frac{\mathbf{1}\{T=t_k\}}{r^{2b}(t_k,\boldsymbol{X})}\Big)+\mathbb{E}\Big(\frac{\mathbf{1}\{T=t_k\}}{r^{2b}(t_k,\boldsymbol{X})}\Big)\right)^{1/b}\\
&\leq 2^{(2b-1)/b}\sigma^2\delta_{jj'}^{(a)}\Big(\frac{2}{r_{\min}^{2b-1}}\Big)^{1/b}=\frac{4}{r_{\min}^{(2b-1)/b}}\sigma^2\delta_{jj'}^{(a)}.
\end{aligned}
\tag{56}
$$

Thus, by combining the two terms, one gets:

$$
\begin{aligned}
\left|\boldsymbol{\Sigma}_{jj'}^{(\mathrm{M})}\right| &\leq \left|\mathbb{E}\big(f_j(\boldsymbol{X})f_{j'}(\boldsymbol{X})\psi_k^2(T,\boldsymbol{X})\big)\right|+\sigma^2\left|\mathbb{E}\big(f_j(\boldsymbol{X})f_{j'}(\boldsymbol{X})A_{t_k}^2(T,\boldsymbol{X})\big)\right|\\
&\leq \frac{16}{r_{\min}^{(2b-1)/b}}\delta_{jj'}^{(a)}C^b+\frac{4}{r_{\min}^{(2b-1)/b}}\sigma^2\delta_{jj'}^{(a)}\\
&\leq \frac{1}{r_{\min}^{(2b-1)/b}}\big(16\,C^b+4\sigma^2\big)\delta_*^{(b)},
\end{aligned}
\tag{57}
$$

where $\delta_*^{(b)}=\max_{j,j'}\left|\mathbb{E}\big(f_j^{b/(b-1)}(\boldsymbol{X})f_{j'}^{b/(b-1)}(\boldsymbol{X})\big)\right|^{(b-1)/b}=\max_{j,j'}\delta_{jj'}^{(a)}$.

Therefore, for all $\epsilon=b-1>0$, there exists $C_M=4C+\sigma^2$ such that

$$
\left|\boldsymbol{\Sigma}_{jj'}^{(\mathrm{M})}\right|\leq 4r_{\min}^{1/(1+\epsilon)-2}\delta_*^{(1+\epsilon)}C_M.
\tag{58}
$$

In particular, if $\epsilon\ll 1$ then $1/(1+\epsilon)-2\approx-(1+\epsilon)$ and

$$
\left|\boldsymbol{\Sigma}_{jj'}^{(\mathrm{M})}\right|\leq\frac{4}{r_{\min}^{1+\epsilon}}\delta_*^{(1+\epsilon)}C_M
\tag{59}
$$

B.1.2. Error estimation of the DR-learner.
In this case, we have

$$
A_{t_k}(T,\boldsymbol{X})=\frac{\mathbf{1}\{T=t_k\}}{r(t_k,\boldsymbol{X})}-\frac{\mathbf{1}\{T=t_k\}}{r(t_k,\boldsymbol{X})},
\tag{60}
$$

$$
B_{t_k}(T,\boldsymbol{X})=\mu_{t_k}(\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})-\left(\frac{\mathbf{1}\{T=t_k\}}{r(t_k,\boldsymbol{X})}-\frac{\mathbf{1}\{T=t_k\}}{r(t_k,\boldsymbol{X})}\right)\mu_T(\boldsymbol{X}).
\tag{61}
$$

We need just to compute the upper bound of $\mathbb{E}\big(f_j(\boldsymbol{X})f_{j'}(\boldsymbol{X})\psi_k^2(T,\boldsymbol{X})\big)$ such that

$$
\begin{aligned}
\psi_k(T,\boldsymbol{X})&=A_{t_k}(T,\boldsymbol{X})f(T,\boldsymbol{X})-\tau_k(\boldsymbol{x})+B_{t_k}(T,\boldsymbol{X})\\
&=\Big(\frac{\mathbf{1}\{T=t_k\}}{r(t_k,\boldsymbol{X})}-1\Big)f(t_k,\boldsymbol{X})-\Big(\frac{\mathbf{1}\{T=t_k\}}{r(t_k,\boldsymbol{X})}-1\Big)f(t_k,\boldsymbol{X})+\mu_{t_k}(\boldsymbol{X})\Big(1-\frac{\mathbf{1}\{T=t_k\}}{r(t_k,\boldsymbol{X})}\Big)\\
&\qquad-\mu_{t_k}(\boldsymbol{X})\Big(1-\frac{\mathbf{1}\{T=t_k\}}{r(t_k,\boldsymbol{X})}\Big)\\
&=\Big(\frac{\mathbf{1}\{T=t_k\}}{r(t_k,\boldsymbol{X})}-1\Big)\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)-\Big(\frac{\mathbf{1}\{T=t_k\}}{r(t_k,\boldsymbol{X})}-1\Big)\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)
\end{aligned}
\tag{62}
$$

Similarly to the previous calculus, we show that for the DR-learner

$$
\begin{aligned}
\left| \mathbb{E}\big(f_j(\boldsymbol{X})f_{j'}(\boldsymbol{X})\psi_k^2(T,\boldsymbol{X})\big)\right| &\leq \left|\mathbb{E}\big(f_j^a(\boldsymbol{X})f_{j'}^a(\boldsymbol{X})\big)\right|^{1/a} \cdot \left|\mathbb{E}\big(\psi_k^{2b}(T,\boldsymbol{X})\big)\right|^{1/b} \quad \text{(Hölder)}\\
&\leq \delta_{jj'}^{(a)}\Big(2^{2b-1}\,\mathbb{E}\Big[\Big(\frac{\mathbf{1}\{T=t_k\}}{r(t_k,\boldsymbol{X})}-1\Big)^{2b}\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^{2b}\\
&\quad +\Big(\frac{\mathbf{1}\{T=t_k\}}{r(t_k,\boldsymbol{X})}-1\Big)^{2b}\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^{2b}\Big]\Big)^{1/b} \quad \text{(Lemma B.1)}\\
&\leq 2^{(2b-1)/b}\,\delta_{jj'}^{(a)}\Big(\mathbb{E}\Big[\Big(\frac{\mathbf{1}\{T=t_k\}}{r(t_k,\boldsymbol{X})}-1\Big)^{2b}\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^{2b}\Big]\\
&\quad +\mathbb{E}\Big[\Big(\frac{\mathbf{1}\{T=t_k\}}{r(t_k,\boldsymbol{X})}-1\Big)^{2b}\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^{2b}\Big]\Big)^{1/b}\\
&\leq 2^{(2b-1)/b}\,\delta_{jj'}^{(a)}\Big(\mathbb{E}\Big[2^{2b-1}\Big(\frac{\mathbf{1}\{T=t_k\}}{r^{2b}(t,\boldsymbol{X})}+1\Big)\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^{2b}\Big]\\
&\quad +\mathbb{E}\Big[2^{2b-1}\Big(\frac{\mathbf{1}\{T=t_k\}}{r^{2b}(t_k,\boldsymbol{X})}+1\Big)\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^{2b}\Big]\Big)^{1/b} \quad \text{(Lemma B.1)}\\
&\leq 2^{2(2b-1)/b}\,\delta_{jj'}^{(a)}\Big(\mathbb{E}\Big[\Big(\frac{1}{r^{2b-1}(t,\boldsymbol{X})}+1\Big)\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^{2b}\Big]\\
&\quad +\mathbb{E}\Big[\Big(\frac{1}{r^{2b-1}(t_k,\boldsymbol{X})}+1\Big)\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^{2b}\Big]\Big)^{1/b}\\
&\leq 2^{2(2b-1)/b}\,\delta_{jj'}^{(a)}\Big(\frac{1}{r_{\min}^{(2b-1)/b}}+1\Big)\Big(\mathbb{E}\Big[\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^{2b}\Big]\\
&\quad +\mathbb{E}\Big[\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^{2b}\Big]\Big)^{1/b}\\
&\leq 2^{2(2b-1)/b}\,\delta_{jj'}^{(a)}\Big(\frac{1}{r_{\min}^{(2b-1)/b}}+1\Big)\Big[\big(\mathbb{E}\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^{2b}\big)^{1/b}\\
&\quad +\mathbb{E}\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^{2b}\big)^{1/b}\Big] \quad \text{(Subadditivity of }|\boldsymbol{X}|^{1/b})
\end{aligned}
\tag{63}
$$

Hence,

$$
\begin{aligned}
\left|\boldsymbol{\Sigma}_{jj'}^{(\mathrm{DR})}\right| &\leq 2^{2(2b-1)/b}\,\delta_{jj'}^{(a)}\Big(\frac{1}{r_{\min}^{(2b-1)/b}}+1\Big)\Big[\big(\mathbb{E}\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^{2b}\big)^{1/b}\\
&\quad +\big(\mathbb{E}\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^{2b}\big)^{1/b}\Big]+\frac{4}{r_{\min}^{(2b-1)/b}}\sigma^2\delta_{jj'}^{(a)}\\
&\leq 2^{2(2b-1)/b}\,\delta_*^{(b)}\Big(\frac{1}{r_{\min}^{(2b-1)/b}}+1\Big)\Big[\big(\mathbb{E}\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^{2b}\big)^{1/b}\\
&\quad +\big(\mathbb{E}\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^{2b}\big)^{1/b}\Big]+\frac{4}{r_{\min}^{(2b-1)/b}}\sigma^2\delta_*^{(b)}
\end{aligned}
\tag{64}
$$

We consider now $\epsilon = b - 1 > 0$, and we assume that $\epsilon \ll 1$, then

$$
\begin{aligned}
&2^{2(2b-1)/b}\,\delta_*^{(b)}\Big(\frac{1}{r_{\min}^{(2b-1)/b}}+1\Big)\Big[\big(\mathbb{E}\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^{2b}\big)^{1/b}+\big(\mathbb{E}\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^{2b}\big)^{1/b}\Big]\\
&+\frac{4}{r_{\min}^{(2b-1)/b}}\sigma^2\delta_*^{(b)} \approx 4\,\delta_*^{(1+\epsilon)}\Big(\frac{1}{r_{\min}^{1+\epsilon}}+1\Big)\Big(\mathbb{E}\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^2+\mathbb{E}\big(f(t_k,\boldsymbol{X})-\mu_{t_k}(\boldsymbol{X})\big)^2\Big)\\
&\quad +4\,\sigma^2\delta_*^{(1+\epsilon)}\frac{1}{r_{\min}^{1+\epsilon}}.
\end{aligned}
\tag{65}
$$

Consequently,

$$\left| \mathbf{\Sigma}_{jj'}^{(\mathrm{DR})} \right| \leq 4 \Big(\frac{C_{DR}^* + \sigma^2}{r_{\min}^{1+\epsilon}} + C_{DR}^* \Big) \delta_*^{(1+\epsilon)}, \tag{66}$$

where $C_{DR}^* = \mathbb{E}\big(f(t_k, \boldsymbol{X}) - \mu_{t_k}(\boldsymbol{X}) \big)^2 + \mathbb{E}\big(f(t_k, \boldsymbol{X}) - \mu_{t_k}(\boldsymbol{X}) \big)^2 = \mathrm{err}(\mu_{t_k}) + \mathrm{err}(\mu_{t_k})$.

B.1.3. ERROR ESTIMATION OF THE X-LEARNER.
In this case, we have

$$A_{t_k}(T, \boldsymbol{X}) = 2 \times \mathbf{1}\{T = t_k\} - 1, \tag{67}$$

$$B_{t_k}(T, \boldsymbol{X}) = (1 - \mathbf{1}\{T = t_k\})\mu_{t_k}(\boldsymbol{X}) - \mu_{t_k}(\boldsymbol{X}) + \sum_{l \neq k} \mathbf{1}\{T = t_l\}\mu_{t_l}(\boldsymbol{X}). \tag{68}$$

One can write ψ_k as

$$\begin{aligned}
\psi_k(T, \boldsymbol{X}) &= A_{t_k}(T, \boldsymbol{X})f(T, \boldsymbol{X}) - \tau_k(\boldsymbol{x}) + B_{t_k}(T, \boldsymbol{X}) \\
&= \big(2\,\mathbf{1}\{T = t_k\} - 1\big)f(T, \boldsymbol{X}) - (f(t_k, \boldsymbol{X}) - f(t_k, \boldsymbol{X})) + \big(1 - \mathbf{1}\{T = t_k\}\big) \\
&\qquad \mu_{t_k}(\boldsymbol{X}) - \mu_{t_k}(\boldsymbol{X}) + \sum_{l \neq k} \mathbf{1}\{T = t_l\}\mu_{t_l}(\boldsymbol{X}) \\
&= \big(1 - \mathbf{1}\{T = t_k\}\big)(\mu_{t_k}(\boldsymbol{X}) - f(t_k, \boldsymbol{X})) - (\mu_{t_k}(\boldsymbol{X}) - f(t_k, \boldsymbol{X})) \\
&\qquad + \sum_{l \neq k} \mathbf{1}\{T = t_l\}\big(\mu_{t_l}(\boldsymbol{X}) - f(t_l, \boldsymbol{X})\big) = a_k + \sum_{l \neq k} b_l.
\end{aligned} \tag{69}$$

where

$$a_k = \big(1 - \mathbf{1}\{T = t_k\}\big)(\mu_{t_k}(\boldsymbol{X}) - f(t_k, \boldsymbol{X})) - (\mu_{t_k}(\boldsymbol{X}) - f(t_k, \boldsymbol{X})), \tag{70}$$

$$b_l = \mathbf{1}\{T = t_l\}\big(\mu_{t_l}(\boldsymbol{X}) - f(t_l, \boldsymbol{X})\big). \tag{71}$$

Similarly to the M- and DR-learners calculus, and using lemma B.1:

$$
\begin{aligned}
\left|\mathbb{E}\big(f_j(\boldsymbol{X})f_{j'}(\boldsymbol{X})\psi_k^2(T,\boldsymbol{X})\big)\right| &\leq \left|\mathbb{E}\big(f_j^a(\boldsymbol{X})f_{j'}^a(\boldsymbol{X})\big)\right|^{1/a}\cdot\left|\mathbb{E}\big(\psi_k^{2b}(T,\boldsymbol{X})\big)\right|^{1/b}\\
&\leq \delta_{jj'}^{(a)}\left|\mathbb{E}\big(a_t+\sum_{l\neq k}b_l\big)^{2b}\right|^{1/b}\quad\text{(Hölder)}\\
&\leq \delta_{jj'}^{(a)}\Big(2^{2b-1}\big(\mathbb{E}\big(a_t^{2b}\big)+\mathbb{E}\big(\sum_{l\neq k}b_l\big)^{2b}\big)\Big)^{1/b}\quad\text{(Lemma B.1 with }m=2)\\
&\leq 2^{(2b-1)/b}\,\delta_{jj'}^{(a)}\Big(\mathbb{E}\big(a_t^{2b}\big)+\mathbb{E}\big(\sum_{l\neq k}b_l\big)^{2b}\Big)^{1/b}\\
&\leq 2^{(2b-1)/b}\,\delta_{jj'}^{(a)}\Big[2^{2b-1}\Big(\mathbb{E}\big((1-\mathbf{1}\{T=t_k\})^{2b}(\mu_{t_k}(\boldsymbol{X})-f(t_k,\boldsymbol{X}))^{2b}\big)\\
&\quad+\mathbb{E}\big(\mu_{t_k}(\boldsymbol{X})-f(t_k,\boldsymbol{X})\big)^{2b}\Big)+(K-1)^{2b-1}\\
&\quad\times\sum_{l\neq k}\mathbb{E}\big(\mathbf{1}\{T=t_l\}\big(\mu_{t_l}(\boldsymbol{X})-f(t_l,\boldsymbol{X})\big)^{2b}\Big]^{1/b}
\end{aligned}
$$
$$(72)$$

(Lemma B.1 with $m=2$ on the 1$^{\text{st}}$ term, and $m=(K-1)$ on the 2$^{\text{nd}}$ term)

$$
\begin{aligned}
&\leq 2^{(2b-1)/b}\,\delta_{jj'}^{(a)}\Big[2^{2b-1}\Big(\mathbb{E}\big(\mu_{t_k}(\boldsymbol{X})-f(t_k,\boldsymbol{X})\big)^{2b}+\mathbb{E}\big(\mu_{t_k}(\boldsymbol{X})-f(t_k,\boldsymbol{X})\big)^{2b}\Big)\\
&\quad+(K-1)^{2b-1}\sum_{l\neq k}\mathbb{E}\big(\mu_{t_l}(\boldsymbol{X})-f(t_l,\boldsymbol{X})\big)^{2b}\Big]^{1/b}\\
&\leq 2^{(2b-1)/b}\,\delta_{jj'}^{(a)}\Big[2^{(2b-1)/b}\Big(\mathbb{E}\big(\mu_{t_k}(\boldsymbol{X})-f(t_k,\boldsymbol{X})\big)^{2b}\Big)^{1/b}+2^{(2b-1)/b}\Big(\mathbb{E}\big(\mu_{t_k}(\boldsymbol{X})\\
&\quad-f(t_k,\boldsymbol{X})\big)^{2b}\Big)^{1/b}+(K-1)^{(2b-1)/b}\sum_{l\neq k}\Big(\mathbb{E}\big(\mu_{t_l}(\boldsymbol{X})-f(t_l,\boldsymbol{X})\big)^{2b}\Big)^{1/b}\Big]\\
&\leq 2^{2(2b-1)/b}\,\delta_{jj'}^{(a)}\Big[\Big(\mathbb{E}\big(\mu_{t_k}(\boldsymbol{X})-f(t_k,\boldsymbol{X})\big)^{2b}\Big)^{1/b}+\Big(\mathbb{E}\big(\mu_{t_k}(\boldsymbol{X})-f(t_k,\boldsymbol{X})\big)^{2b}\Big)^{1/b}\\
&\quad+\big(\frac{K-1}{2}\big)^{(2b-1)/b}\sum_{l\neq k}\Big(\mathbb{E}\big(\mu_{t_l}(\boldsymbol{X})-f(t_l,\boldsymbol{X})\big)^{2b}\Big)^{1/b}\Big].
\end{aligned}
$$

Given that $\mathbb{E}\big(f_j(\boldsymbol{X})f_{j'}(\boldsymbol{X})A_{t_k}^2(T,\boldsymbol{X})\big)=\mathbb{E}\big(f_j(\boldsymbol{X})f_{j'}(\boldsymbol{X})\big)=\delta_{jj'}^{(1)}$, we deduce finally

$$
\begin{aligned}
\left|\boldsymbol{\Sigma}_{jj'}^{(\mathrm{X})}\right| &\leq \left|\mathbb{E}\big(f_j(\boldsymbol{X})f_{j'}(\boldsymbol{X})\psi_k^2(T,\boldsymbol{X})\big)\right|+\sigma^2\left|\mathbb{E}\big(f_j(\boldsymbol{X})f_{j'}(\boldsymbol{X})A_{t_k}^2(T,\boldsymbol{X})\big)\right|\\
&\leq 2^{2(2b-1)/b}\,\delta_{jj'}^{(a)}\Big[\Big(\mathbb{E}\big(\mu_{t_k}(\boldsymbol{X})-f(t_k,\boldsymbol{X})\big)^{2b}\Big)^{1/b}+\Big(\mathbb{E}\big(\mu_{t_k}(\boldsymbol{X})-f(t_k,\boldsymbol{X})\big)^{2b}\Big)^{1/b}\\
&\quad+\big(\frac{K-1}{2}\big)^{(2b-1)/b}\sum_{l\neq k}\Big(\mathbb{E}\big(\mu_{t_l}(\boldsymbol{X})-f(t_l,\boldsymbol{X})\big)^{2b}\Big)^{1/b}\Big]+\sigma^2\delta_{jj'}^{(1)}\\
&\leq 2^{2(2b-1)/b}\,\delta_*^{(b)}\Big[\Big(\mathbb{E}\big(\mu_{t_k}(\boldsymbol{X})-f(t_k,\boldsymbol{X})\big)^{2b}\Big)^{1/b}+\Big(\mathbb{E}\big(\mu_{t_k}(\boldsymbol{X})-f(t_k,\boldsymbol{X})\big)^{2b}\Big)^{1/b}\\
&\quad+\big(\frac{K-1}{2}\big)^{(2b-1)/b}\sum_{l\neq k}\Big(\mathbb{E}\big(\mu_{t_l}(\boldsymbol{X})-f(t_l,\boldsymbol{X})\big)^{2b}\Big)^{1/b}\Big]+\sigma^2\delta_*^{(1)}
\end{aligned}
$$
$$(73)$$

where $\delta_*^{(1)}=\max_{j,j'}\mathbb{E}\big(f_j(\boldsymbol{X})f_{j'}(\boldsymbol{X})\big)$.

As in the previous cases, we consider now $\epsilon = b - 1 > 0$ with $\epsilon \ll 1$, then

$$
\begin{aligned}
2^{2(2b-1)/b} \, \delta_*^{(b)} \Big[&\Big(\mathbb{E}\big(\mu_{t_k}(\boldsymbol{X}) - f(t_k, \boldsymbol{X})\big)^{2b} \Big)^{1/b} + \Big(\mathbb{E}\big(\mu_{t_k}(\boldsymbol{X}) - f(t_k, \boldsymbol{X})\big)^{2b} \Big)^{1/b} \\
&+ \big(\tfrac{K-1}{2} \big)^{(2b-1)/b} \sum_{l \neq k} \Big(\mathbb{E}\big(\mu_{t_l}(\boldsymbol{X}) - f(t_l, \boldsymbol{X})\big)^{2b} \Big)^{1/b} \Big] + \sigma^2 \delta_*^{(1)} \\
\approx 4 \, \delta_*^{(1+\epsilon)} \Big(&\mathbb{E}\big(f(t_k, \boldsymbol{X}) - \mu_{t_k}(\boldsymbol{X})\big)^2 + \mathbb{E}\big(f(t_k, \boldsymbol{X}) - \mu_{t_k}(\boldsymbol{X})\big)^2 \\
&+ \frac{(K-1)^2}{4} \sum_{l \neq k} \mathbb{E}\big(\mu_{t_l}(\boldsymbol{X}) - f(t_l, \boldsymbol{X})\big)^2 + \sigma^2 \delta_*^{(1)}.
\end{aligned}
\tag{74}
$$

Therefore,

$$
\Big| \boldsymbol{\Sigma}_{jj'}^{(\mathrm{X})} \Big| \leq 4 \delta_*^{(1+\epsilon)} C_X + \sigma^2 \delta_*^{(1)}.
\tag{75}
$$

where $C_X = \mathrm{err}(\mu_{t_k}) + \mathrm{err}(\mu_{t_k}) + \frac{(K-1)^2}{4} \sum_{l \neq k} \mathrm{err}(\mu_{t_l})$.

B.2. Error estimation of the T- and naive X-learners.

In this subsection, we propose to conduct the bias-variance analysis of the T-learner and the naive extension of the X-learner. Some steps of this proof are quite similar to the proof of Appendix B.1.

B.2.1. ERROR ESTIMATION OF THE T-LEARNER.

STEP 0. SET-UP
For all $t \in \mathcal{T}$, we define the set $\mathbf{S}_t = \{i, T_i = t\}$ with $n_t = |\mathbf{S}_t|$. Under Assumptions (3.1-5.2), the T-learner of the CATE can be defined as

$$
\widehat{\tau}_k^{(\mathrm{T})}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})^\top \widehat{\boldsymbol{\beta}}_{t_k} - \boldsymbol{f}(\boldsymbol{x})^\top \widehat{\boldsymbol{\beta}}_{t_0} = \boldsymbol{f}(\boldsymbol{x})^\top (\widehat{\boldsymbol{\beta}}_{t_k} - \widehat{\boldsymbol{\beta}}_{t_0}),
\tag{76}
$$

where $\widehat{\boldsymbol{\beta}}_{t_k}$ and $\widehat{\boldsymbol{\beta}}_{t_0}$ are the OLS estimators of $\boldsymbol{\beta}_{t_k}$ and $\boldsymbol{\beta}_{t_0}$ such that:

$$
\widehat{\boldsymbol{\beta}}_{t_k} = \big(\mathbf{H}_k^\top \mathbf{H}_k \big)^{-1} \mathbf{H}_k^\top \boldsymbol{y}_k,
\tag{77}
$$

$$
\widehat{\boldsymbol{\beta}}_{t_0} = \big(\mathbf{H}_0^\top \mathbf{H}_0 \big)^{-1} \mathbf{H}_0^\top \boldsymbol{y}_0,
\tag{78}
$$

where $\mathbf{H}_k = (f_j(\boldsymbol{X}_i))_{i \in \mathbf{S}_{t_k}, j} \in \mathbb{R}^{n_k \times p}$ (respectively, $\mathbf{H}_0 = (f_j(\boldsymbol{X}_i))_{i \in \mathbf{S}_{t_0}, j} \in \mathbb{R}^{n_0 \times p}$) is the regression matrix and $\boldsymbol{y}_k = (Y_{\mathrm{obs},i})_{i \in \mathbf{S}_{t_k}}$ (respectively, $\boldsymbol{y}_0 = (Y_{\mathrm{obs},i})_{i \in \mathbf{S}_{t_0}}$).

STEP 1. IDENTIFICATION OF $\widehat{\boldsymbol{\beta}}_k$
By similar calculus, we show that:

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}}_{t_k} &= \big(\mathbf{H}_k^\top \mathbf{H}_k \big)^{-1} \mathbf{H}_k^\top \boldsymbol{y}_k \\
&= \big(\mathbf{H}_k^\top \mathbf{H}_k \big)^{-1} \mathbf{H}_k^\top \big(f(t_k, \boldsymbol{X}_i) + \varepsilon_i(t_k) \big)_{i \in \mathbf{S}_{t_k}} \\
&= \boldsymbol{\beta}_{t_k} + \big(\mathbf{H}_k^\top \mathbf{H}_k \big)^{-1} \mathbf{H}_k^\top \boldsymbol{\epsilon}_k \\
&= \boldsymbol{\beta}_{t_k} + \frac{1}{\sqrt{n_k}} \Big(\frac{1}{n_k} \mathbf{H}_k^\top \mathbf{H}_k \Big)^{-1} \Big(\frac{1}{\sqrt{n_k}} \mathbf{H}_k^\top \boldsymbol{\epsilon}_k \Big).
\end{aligned}
\tag{79}
$$

where $\boldsymbol{\epsilon}_k = (\varepsilon_i(t_k))_{i=1}^n$ are i.i.d. Gaussian $\mathcal{N}(0, \sigma^2)$ and independent of $(T_i, \boldsymbol{X}_i)_{i=1}^n$.

Thus,

$$
\sqrt{n}\big(\widehat{\boldsymbol{\beta}}_{t_k} - \boldsymbol{\beta}_{t_k} \big) = \sqrt{\frac{n}{n_k}} \Big(\frac{1}{n_k} \mathbf{H}_k^\top \mathbf{H}_k \Big)^{-1} \Big(\frac{1}{\sqrt{n_k}} \mathbf{H}_k^\top \boldsymbol{\epsilon}_k \Big).
\tag{80}
$$

By similar calculus

$$
\sqrt{n}\big(\widehat{\boldsymbol{\beta}}_{t_0} - \boldsymbol{\beta}_{t_0} \big) = \sqrt{\frac{n}{n_0}} \Big(\frac{1}{n_0} \mathbf{H}_0^\top \mathbf{H}_0 \Big)^{-1} \Big(\frac{1}{\sqrt{n_0}} \mathbf{H}_0^\top \boldsymbol{\epsilon}_0 \Big).
\tag{81}
$$

Therefore, by considering $\widehat{\boldsymbol{\beta}}_k = \widehat{\boldsymbol{\beta}}_{t_k} - \widehat{\boldsymbol{\beta}}_{t_0}$,

$$
\begin{aligned}
\sqrt{n}\big(\widehat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k^*\big) &= \sqrt{n}\big(\widehat{\boldsymbol{\beta}}_{t_k} - \boldsymbol{\beta}_{t_k}\big) + \sqrt{n}\big(\widehat{\boldsymbol{\beta}}_{t_0} - \boldsymbol{\beta}_{t_0}\big) \\
&= \sqrt{\frac{n}{n_k}}\Big(\frac{1}{n_k}\mathbf{H}_k^\top \mathbf{H}_k\Big)^{-1}\Big(\frac{1}{\sqrt{n_k}}\mathbf{H}_k^\top \boldsymbol{\epsilon}_k\Big) + \sqrt{\frac{n}{n_0}}\Big(\frac{1}{n_0}\mathbf{H}_0^\top \mathbf{H}_0\Big)^{-1}\Big(\frac{1}{\sqrt{n_0}}\mathbf{H}_0^\top \boldsymbol{\epsilon}_0\Big).
\end{aligned}
\tag{82}
$$

STEP 2. THE ASYMPTOTIC BEHAVIOUR OF THE OLS ESTIMATOR'S MEAN AND COVARIANCE
Let $\boldsymbol{a} = (\boldsymbol{a}_k, \boldsymbol{a}_0) \in \mathbb{R}^{2p}$ and let ϕ_n denote the characteristic function of the vector $\big(\frac{1}{\sqrt{n_k}}\mathbf{H}_k^\top \boldsymbol{\epsilon}_k, \frac{1}{\sqrt{n_0}}\mathbf{H}_0^\top \boldsymbol{\epsilon}_0\big)$. We have

$$
\begin{aligned}
\phi_n(\boldsymbol{a}) &= \mathbb{E}\Big[\exp i\boldsymbol{a}^\top\big(\frac{1}{\sqrt{n_k}}\mathbf{H}_k^\top \boldsymbol{\epsilon}_k, \frac{1}{\sqrt{n_0}}\mathbf{H}_0^\top \boldsymbol{\epsilon}_0\big)\Big] \\
&= \mathbb{E}\Big[\exp i\big(\boldsymbol{a}_k^\top \frac{1}{\sqrt{n_k}}\mathbf{H}_k^\top \boldsymbol{\epsilon}_k + \boldsymbol{a}_0^\top \frac{1}{\sqrt{n_0}}\mathbf{H}_0^\top \boldsymbol{\epsilon}_0\big)\Big] \\
&= \mathbb{E}\Big[\exp i\Big(\frac{1}{\sqrt{n_k}}\sum_{m=1}^{n_k}\boldsymbol{a}_k^\top \big(H_{mj}\varepsilon_m(t_k)\big)_{j=0}^{p-1} + \frac{1}{\sqrt{n_0}}\sum_{m=1}^{n_0}\boldsymbol{a}_0^\top \big(H_{mj}\varepsilon_m(t_0)\big)_{j=0}^{p-1}\Big)\Big] \\
&= \mathbb{E}\Big[\exp i\Big(\frac{1}{\sqrt{n}}\sum_{m=1}^{n}\boldsymbol{a}_k^\top \big(H_{mj}\big)_{j=0}^{p-1}\varepsilon_m(t_k)\mathbf{1}\{T_m = t_k\} \times \frac{\sqrt{n}}{\sqrt{n_k}} \\
&\qquad\qquad + \frac{1}{\sqrt{n}}\sum_{m=1}^{n}\boldsymbol{a}_0^\top \big(H_{mj}\big)_{j=0}^{p-1}\varepsilon_m(t_0)\mathbf{1}\{T_m = t_0\} \times \frac{\sqrt{n}}{\sqrt{n_0}}\Big)\Big].
\end{aligned}
\tag{83}
$$

Now, let us consider the vector $\boldsymbol{Z}^{(n)} = \big(\boldsymbol{Z}_k^{(n)}, \boldsymbol{Z}_0^{(n)}\big) \in \mathbb{R}^{2p}$ such that

$$
\begin{aligned}
\boldsymbol{Z}^{(n)} &= \Big(\frac{1}{n}\sum_{m=1}^{n}H_{m1}\varepsilon_m(t_k)\mathbf{1}\{T_m = t_k\},\ldots,\frac{1}{n}\sum_{m=1}^{n}H_{mp}\varepsilon_m(t_k)\mathbf{1}\{T_m = t_k\}, \\
&\qquad \frac{1}{n}\sum_{m=1}^{n}\big(H_{m1}\varepsilon_m(t_0)\mathbf{1}\{T_m = t_0\},\ldots,\frac{1}{n}\sum_{m=1}^{n}H_{mp}\varepsilon_m(t_0)\mathbf{1}\{T_m = t_0\}\big) \\
&= \frac{1}{n}\sum_{m=1}^{n}\Big(H_{m1}\varepsilon_m(t_k)\mathbf{1}\{T_m = t_k\},\ldots,H_{mp}\varepsilon_m(t_k)\mathbf{1}\{T_m = t_k\}, \\
&\qquad\qquad H_{m1}\varepsilon_m(t_0)\mathbf{1}\{T_m = t_0\},\ldots,H_{mp}\varepsilon_m(t_0)\mathbf{1}\{T_m = t_0\}\Big) \\
&= \frac{1}{n}\sum_{m=1}^{n}\boldsymbol{Z}_m.
\end{aligned}
\tag{84}
$$

The mean $\boldsymbol{m} = (\boldsymbol{m}_k, \boldsymbol{m}_0)$ of the vector $\boldsymbol{Z}_m$ satisfies, for $j = 0,\ldots,p-1$,

$$
m_{k,j} = \mathbb{E}\big[f_j(\boldsymbol{X})\varepsilon(t_k)\mathbf{1}\{T = t_k\}\big] = 0,
\tag{85}
$$

$$
m_{0,j} = \mathbb{E}\big[f_j(\boldsymbol{X})\varepsilon(t_0)\mathbf{1}\{T = t_0\}\big] = 0,
\tag{86}
$$

and its covariance matrix that satisfies, for $j, j' = 1, \ldots, 2p$,

$$
\operatorname{Cov}\left(\boldsymbol{Z}_{m,j}, \boldsymbol{Z}_{m,j'}\right) = \begin{cases} \mathbb{E}\big[f_{j-1}(\boldsymbol{X})f_{j'-1}(\boldsymbol{X})\varepsilon^2(t_k)\mathbf{1}\{T = t_k\}\big] & \text{if } j, j' \in \{1, \ldots, p\} \\ \mathbb{E}\big[f_{j-1}(\boldsymbol{X})f_{j'-1}(\boldsymbol{X})\varepsilon^2(t_0)\mathbf{1}\{T = t_0\}\big] & \text{if } j, j' \in \{p+1, \ldots, 2p\} \\ \mathbb{E}\big[f_{j-1}(\boldsymbol{X})f_{j'-1}(\boldsymbol{X})\varepsilon(t_k)\varepsilon(t_0)\mathbf{1}\{T = t_k\}\mathbf{1}\{T = t_0\}\big] & \text{otherwise.} \end{cases}
$$

$$
= \begin{cases} \sigma^2\mathbb{E}\big[f_{j-1}(\boldsymbol{X})f_{j'-1}(\boldsymbol{X})\mathbf{1}\{T = t_k\}\big] & \text{if } j, j' \in \{1, \ldots, p\} \\ \sigma^2\mathbb{E}\big[f_{j-1}(\boldsymbol{X})f_{j'-1}(\boldsymbol{X})\mathbf{1}\{T = t_0\}\big] & \text{if } j, j' \in \{p+1, \ldots, 2p\} \\ 0 & \text{otherwise,} \end{cases} \tag{87}
$$

$$
= \begin{cases} \sigma^2\rho(t_k)F_{k,jj'} & \text{if } j, j' \in \{1, \ldots, p\}, \\ \sigma^2\rho(t_0)F_{0,jj'} & \text{if } j, j' \in \{p+1, \ldots, 2p\}, \\ 0 & \text{otherwise,} \end{cases}
$$

where the matrices $\mathbf{F}_k = \left(\mathbb{E}\big[f_{j-1}(\boldsymbol{X})f_{j'-1}(\boldsymbol{X}) \mid T = t_k\big]\right)_{j,j'} \in \mathbb{R}^p$ and $\mathbf{F}_0 = \left(\mathbb{E}\big[f_{j-1}(\boldsymbol{X})f_{j'-1}(\boldsymbol{X}) \mid T = t_0\big]\right)_{j,j'} \in \mathbb{R}^p$ are supposed to be invertible. Note that, for a integrable function h, $\mathbb{E}\big[h(\boldsymbol{X})\mathbf{1}\{T = t\}\big] = \mathbb{P}(T = t)\mathbb{E}\big[h(\boldsymbol{X}) \mid T = t\big]$.

Therefore, using the multivariate CLT on $\boldsymbol{Z}^{(n)}$, we get

$$
\begin{pmatrix} \sqrt{n}\, \boldsymbol{a}_k^\top \boldsymbol{Z}_k^{(n)} \\ \sqrt{n}\, \boldsymbol{a}_0^\top \boldsymbol{Z}_0^{(n)} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{n}}\sum_{m=1}^n \boldsymbol{a}_k^\top \left(H_{mj}\right)_{j=0}^{p-1}\varepsilon_m(t_k)\mathbf{1}\{T_m = t_k\} \\ \frac{1}{\sqrt{n}}\sum_{m=1}^n \boldsymbol{a}_0^\top \left(H_{mj}\right)_{j=0}^{p-1}\varepsilon_m(t_0)\mathbf{1}\{T_m = t_0\} \end{pmatrix}
$$
$$
\xrightarrow{\mathcal{L}} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2\rho(t_k)\boldsymbol{a}_k^\top \mathbf{F}_k\boldsymbol{a}_k & 0 \\ 0 & \sigma^2\rho(t_0)\boldsymbol{a}_0^\top \mathbf{F}_0\boldsymbol{a}_0 \end{pmatrix}\right). \tag{88}
$$

On the other hand,

$$
\left(\frac{\sqrt{n}}{\sqrt{n_k}}, \frac{\sqrt{n}}{\sqrt{n_0}}\right) = \left(\frac{\sqrt{n}}{\sqrt{\sum_{m=1}^n \mathbf{1}\{T_m = t_k\}}}, \frac{\sqrt{n}}{\sqrt{\sum_{m=1}^n \mathbf{1}\{T_m = t_0\}}}\right) \xrightarrow{a.s} \left(\frac{1}{\sqrt{\rho(t_k)}}, \frac{1}{\sqrt{\rho(t_0)}}\right), \tag{89}
$$

where $\rho(t) = \mathbb{P}(T = t)$.

Thus, by the Slutsky theorem:

$$
\frac{1}{\sqrt{n}}\sum_{m=1}^n \left(\boldsymbol{a}_k^\top \left(H_{mj}\right)_{j=0}^{p-1}\varepsilon_m(t_k)\mathbf{1}\{T_m = t_k\}\frac{\sqrt{n}}{\sqrt{n_k}} + \boldsymbol{a}_0^\top \left(H_{mj}\right)_{j=0}^{p-1}\varepsilon_m(t_0)\mathbf{1}\{T_m = t_0\}\frac{\sqrt{n}}{\sqrt{n_0}}\right)
$$
$$
\xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \sigma^2\boldsymbol{a}_k^\top \mathbf{F}_k\boldsymbol{a}_k + \sigma^2\boldsymbol{a}_0^\top \mathbf{F}_0\boldsymbol{a}_0). \tag{90}
$$

Therefore,
$$
\phi_n(\boldsymbol{a}) \xrightarrow{n \to +\infty} \mathbb{E}\left[\exp i\left(\boldsymbol{a}_k^\top \sigma^2\rho(t_k)\mathbf{F}_k\boldsymbol{a}_k + \boldsymbol{a}_0^\top \sigma^2\rho(t_0)\mathbf{F}_0\boldsymbol{a}_0\right)\right] = \phi_{(\boldsymbol{Z}_k, \boldsymbol{Z}_0)}(\boldsymbol{a}), \tag{91}
$$

where $\boldsymbol{Z}_k$ and $\boldsymbol{Z}_0$ are two independent zero-mean random vectors with covariance matrices $\sigma^2\mathbf{F}_k$ and $\sigma^2\mathbf{F}_0$ respectively.

As shown previously in Appendix B.1, we can prove immediately that $\left(1/n_k\, \mathbf{H}_k^\top \mathbf{H}_k\right)^{-1} \xrightarrow{P} \mathbf{F}_k^{-1}$. Moreover, $n_k/n \xrightarrow{a.s} \rho(t_k)$ so $n_k/n \xrightarrow{P} \rho(t_k)$. Thus

$$
\sqrt{\frac{n}{n_k}}\left(\frac{1}{n_k}\mathbf{H}_k^\top \mathbf{H}_k\right)^{-1} \xrightarrow{P} \frac{1}{\sqrt{\rho(t_k)}}\mathbf{F}_k^{-1}. \tag{92}
$$

Finally, given Equation (82) and using the Slutsky theorem, we get

$$
\sqrt{n}\left(\widehat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_k^*\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\mathbf{0}, \frac{1}{\rho(t_k)}\mathbf{F}_k^{-1}\sigma^2\mathbf{F}_k\mathbf{F}_k^{-1} + \frac{1}{\rho(t_0)}\mathbf{F}_0^{-1}\sigma^2\mathbf{F}_0\mathbf{F}_0^{-1}\right)
$$
$$
= \mathcal{N}\left(\mathbf{0}, \frac{\sigma^2}{\rho(t_k)}\mathbf{F}_k^{-1} + \frac{\sigma^2}{\rho(t_0)}\mathbf{F}_0^{-1}\right). \tag{93}
$$

Here also, we can deduce that the asymptotic mean and covariance matrix are of the form

$$
\begin{aligned}
\mathbb{E}(\widehat{\boldsymbol{\beta}}_k) &= \boldsymbol{\beta}_{t_k} - \boldsymbol{\beta}_{t_0} = \boldsymbol{\beta}_k^*, \\
\mathbb{V}(\widehat{\boldsymbol{\beta}}_k) &\approx \frac{1}{n} \Big(\frac{1}{\rho(t_k)} \mathbf{F}_k^{-1} + \frac{1}{\rho(t_0)} \mathbf{F}_0^{-1} \Big) \sigma^2.
\end{aligned}
\tag{94}
$$

STEP 3. OBTAINING THE ERROR UPPER BOUND

The asymptotic covariance matrix is given by the matrices $\mathbf{F}_k^{-1}$ and $\mathbf{F}_0^{-1}$. We assume that the polynomials f_j are chosen to be orthonormal, and that, conditionally to T, their distribution is not significantly different. One can anticipate, therefore, that $\mathbf{F}_k, \mathbf{F}_0 \approx \mathbf{F}$ and easily identify the error's upper bound of the T-learner as:

$$
\frac{1}{\rho(t_k)} + \frac{1}{\rho(t_0)}.
\tag{95}
$$

B.3. Error estimation of the naive X-learner.

Let $\bar{r}$ denote a fixed arbitrary estimator of the GPS (see remark 3.3) and respecting the assumption 3.2, that is, $r_{\min} \leq \bar{r}(t, \boldsymbol{x})$. Let $\widehat{\mu}_{t_k}$ denote the estimator of μ_{t_k}. The model $\widehat{\mu}_{t_k}$ is trained using the sample $\mathbf{S}_{t_k}$, the OLS estimator $\widehat{\boldsymbol{\beta}}_{t_k}$ satisfies

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}}_{t_k} &= \big(\mathbf{H}_k^\top \mathbf{H}_k \big)^{-1} \mathbf{H}_k^\top \boldsymbol{y}_k \\
&= \boldsymbol{\beta}_{t_k} + \big(\mathbf{H}_k^\top \mathbf{H}_k \big)^{-1} \mathbf{H}_k^\top \boldsymbol{\epsilon}_k
\end{aligned}
\tag{96}
$$

where $\boldsymbol{y}_k = (Y_{\mathrm{obs},i})_{i \in \mathbf{S}_{t_k}}$ and $\boldsymbol{\epsilon}_k = (\varepsilon_i(t_k))_{i \in \mathbf{S}_{t_k}}$.

Similarly, the OLS estimator of μ_{t_0} satisfies also

$$
\widehat{\boldsymbol{\beta}}_{t_0} = \boldsymbol{\beta}_{t_0} + \big(\mathbf{H}_0^\top \mathbf{H}_0 \big)^{-1} \mathbf{H}_0^\top \boldsymbol{\epsilon}_0,
\tag{97}
$$

where $\boldsymbol{y}_0 = (Y_{\mathrm{obs},i})_{i \in \mathbf{S}_{t_0}}$ and $\boldsymbol{\epsilon}_0 = (\varepsilon_i(t_0))_{i \in \mathbf{S}_{t_0}}$.

We recall now the definition of the naive extension of the X-learner:

$$
\widehat{\tau}_k^{(\mathrm{X,nv})}(\boldsymbol{x}) = \frac{\bar{r}(t_k, \boldsymbol{x})}{\bar{r}(t_k, \boldsymbol{x}) + \bar{r}(t_0, \boldsymbol{x})} \widehat{\tau}^{(k)}(\boldsymbol{x}) + \frac{\bar{r}(t_0, \boldsymbol{x})}{\bar{r}(t_k, \boldsymbol{x}) + \bar{r}(t_0, \boldsymbol{x})} \widehat{\tau}^{(0)}(\boldsymbol{x}).
\tag{98}
$$

where the estimators $\widehat{\tau}^{(k)}$ and $\widehat{\tau}^{(0)}$ are built respectively on $\mathbf{S}_{t_k}$ and $\mathbf{S}_{t_0}$ by regressing $(D_i^{(k)})_{i \in \mathbf{S}_{t_k}} = (Y_i(t_k) - \widehat{\mu}_{t_0}(\boldsymbol{X}_i))_{i \in \mathbf{S}_{t_k}}$ and $(D_i^{(0)})_{i \in \mathbf{S}_{t_0}} = (\widehat{\mu}_{t_k}(\boldsymbol{X}_i) - Y_i(t_0))_{i \in \mathbf{S}_{t_0}}$ on $\boldsymbol{X}$. In the following, we denote $\widehat{\tau}^{(k)}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})^\top \widehat{\boldsymbol{\beta}}^{(k)}$ and $\widehat{\tau}^{(0)}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})^\top \widehat{\boldsymbol{\beta}}^{(0)}$. Here, $\widehat{\boldsymbol{\beta}}^{(k)}$ denotes the OLS estimator of $\widehat{\tau}^{(k)}$ and is given by:

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}}^{(k)} &= \big(\mathbf{H}_k^\top \mathbf{H}_k \big)^{-1} \mathbf{H}_k^\top \big(Y_{\mathrm{obs,i}} - \widehat{\mu}_{t_0}(\boldsymbol{X}_i) \big)_{i \in \mathbf{S}_{t_k}} \\
&= \big(\mathbf{H}_k^\top \mathbf{H}_k \big)^{-1} \mathbf{H}_k^\top \big(Y_{\mathrm{obs,i}} - \boldsymbol{f}(\boldsymbol{X}_i)^\top \widehat{\boldsymbol{\beta}}_{t_0} \big)_{i \in \mathbf{S}_{t_k}} \\
&= \big(\mathbf{H}_k^\top \mathbf{H}_k \big)^{-1} \mathbf{H}_k^\top \boldsymbol{y}_k - \big(\mathbf{H}_k^\top \mathbf{H}_k \big)^{-1} \mathbf{H}_k^\top \mathbf{H}_k \widehat{\boldsymbol{\beta}}_{t_0} \\
&= \widehat{\boldsymbol{\beta}}_{t_k} - \widehat{\boldsymbol{\beta}}_{t_0} = \widehat{\boldsymbol{\beta}}_k,
\end{aligned}
\tag{99}
$$

where $\widehat{\boldsymbol{\beta}}_k = \widehat{\boldsymbol{\beta}}_{t_k} - \widehat{\boldsymbol{\beta}}_{t_0}$ is the T-learner OLS estimator as given in (93).

By similar calculus, we show that

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}}^{(0)} &= \big(\mathbf{H}_0^\top \mathbf{H}_0 \big)^{-1} \mathbf{H}_0^\top \big(\widehat{\mu}_{t_k}(\boldsymbol{X}_i) - Y_{\mathrm{obs,i}} \big)_{i \in \mathbf{S}_{t_0}} \\
&= \big(\mathbf{H}_0^\top \mathbf{H}_0 \big)^{-1} \mathbf{H}_0^\top \big(\boldsymbol{f}(\boldsymbol{X}_i)^\top \widehat{\boldsymbol{\beta}}_{t_k} - Y_{\mathrm{obs,i}} \big)_{i \in \mathbf{S}_{t_0}} \\
&= \big(\mathbf{H}_0^\top \mathbf{H}_0 \big)^{-1} \mathbf{H}_0^\top \mathbf{H}_0 \widehat{\boldsymbol{\beta}}_{t_k} - \big(\mathbf{H}_0^\top \mathbf{H}_0 \big)^{-1} \mathbf{H}_0^\top \boldsymbol{y}_0 \\
&= \widehat{\boldsymbol{\beta}}_{t_k} - \widehat{\boldsymbol{\beta}}_{t_0} = \widehat{\boldsymbol{\beta}}_k.
\end{aligned}
\tag{100}
$$

It results that

$$
\begin{aligned}
\widehat{\tau}_k^{(\mathrm{X,nv})}(\boldsymbol{x}) &= \frac{\overline{r}(t_k, \boldsymbol{x})}{\overline{r}(t_k, \boldsymbol{x}) + \overline{r}(t_0, \boldsymbol{x})} \boldsymbol{f}(\boldsymbol{x})^{\top} \widehat{\boldsymbol{\beta}}^{(k)} + \frac{\overline{r}(t_0, \boldsymbol{x})}{\overline{r}(t_k, \boldsymbol{x}) + \overline{r}(t_0, \boldsymbol{x})} \boldsymbol{f}(\boldsymbol{x})^{\top} \widehat{\boldsymbol{\beta}}^{(0)} \\
&= \Big(\frac{\overline{r}(t_k, \boldsymbol{x})}{\overline{r}(t_k, \boldsymbol{x}) + \overline{r}(t_0, \boldsymbol{x})} + \frac{\overline{r}(t_0, \boldsymbol{x})}{\overline{r}(t_k, \boldsymbol{x}) + \overline{r}(t_0, \boldsymbol{x})} \Big) \boldsymbol{f}(\boldsymbol{x})^{\top} \widehat{\boldsymbol{\beta}}_k = \boldsymbol{f}(\boldsymbol{x})^{\top} \widehat{\boldsymbol{\beta}}_k
\end{aligned}
\tag{101}
$$

In the end, the naive X-learner is no more than a simple T-learner, the error's upper bound of the naive X-learner is given therefore by:

$$
\sigma^2 \Big(\frac{1}{\rho(t_k)} + \frac{1}{\rho(t_0)} \Big).
\tag{102}
$$

C. Discussion about the binarized R-learner.

Another alternative to R-learning to continuous treatments is proposed by (Kaddour et al., 2021). The approach considers both Assumptions 5.1 and 5.2 on the outcome $Y(t) = \boldsymbol{f}(\boldsymbol{X})^{\top} \boldsymbol{\beta}_t + \varepsilon(t)$, then establishes the binarized (Robinson, 1988) decomposition such that

$$
Y_{\mathrm{obs}} - m(\boldsymbol{X}) = \boldsymbol{f}(\boldsymbol{X})^{\top} (\boldsymbol{\beta}_T - e^{\boldsymbol{\beta}}(\boldsymbol{X})) + \epsilon,
\tag{103}
$$

where $\epsilon = \varepsilon(T)$, $m(\boldsymbol{x}) = \mathbb{E}(Y_{\mathrm{obs}} \mid \boldsymbol{X} = \boldsymbol{x})$ and $e^{\boldsymbol{\beta}}(\boldsymbol{x}) = \mathbb{E}(\boldsymbol{\beta}_T \mid \boldsymbol{X} = \boldsymbol{x})$.

Considering the mean squared error of ϵ as a loss function and minimizing it allows us to identify the optimal $\widehat{\boldsymbol{\beta}}$ and therefore CATEs. Given two nuisance estimators $\widehat{m}$ and $\widehat{e}^{\boldsymbol{\beta}}$ of m and $e^{\boldsymbol{\beta}}$, one needs to solve the following problem:

$$
\widehat{\boldsymbol{\beta}} = \operatorname{argmin}_{\boldsymbol{\beta} \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^{n} \Big[(Y_{\mathrm{obs},i} - \widehat{m}(\boldsymbol{X}_i)) - \boldsymbol{f}(\boldsymbol{X}_i)^{\top} (\boldsymbol{\beta}_{T_i} - \widehat{e}^{\boldsymbol{\beta}}(\boldsymbol{X}_i)) \Big]^2,
\tag{104}
$$

where $\mathcal{F}$ is the space of candidate models $\boldsymbol{\beta}$. The previous problem corresponds to a classical OLS estimator and has, therefore, a unique solution.

If the space of candidate models $\mathcal{F}$ is separable, then the optimization problem can be divided into the following sub-problems:

$$
\widehat{\boldsymbol{\beta}}_{t_0} = \operatorname{argmin} \frac{1}{n_k} \sum_{i \in \mathbf{S}_{t_0}} \Big[(Y_{\mathrm{obs},i} - \widehat{m}(\boldsymbol{X}_i)) - \boldsymbol{f}(\boldsymbol{X}_i)^{\top} (\boldsymbol{\beta}_{t_0} - \widehat{e}^{\boldsymbol{\beta}}(\boldsymbol{X}_i)) \Big]^2
$$

$$
\vdots
$$

$$
\widehat{\boldsymbol{\beta}}_{t_K} = \operatorname{argmin} \frac{1}{n_K} \sum_{i \in \mathbf{S}_{t_K}} \Big[(Y_{\mathrm{obs},i} - \widehat{m}(\boldsymbol{X}_i)) - \boldsymbol{f}(\boldsymbol{X}_i)^{\top} (\boldsymbol{\beta}_{t_K} - \widehat{e}^{\boldsymbol{\beta}}(\boldsymbol{X}_i)) \Big]^2.
$$

However, this approach does not consider the interactions between different $\widehat{\boldsymbol{\beta}}_t$ and is computationally heavy when the number of possible treatments K becomes larger. It also requires specifying the family of models $\mathcal{F}$ and precise the dimension p for Assumption 5.2.

There are two main differences between the generalized R-learner and the binarized: 1) In the binarized R-learner, $(\widehat{\boldsymbol{\beta}}_{t_k})_{k=1}^{K}$ may be solved separately but using a small sample ($\mathbf{S}_{t_k}$ instead of $\mathbf{D}_{\mathrm{obs}}$); 2) The solution $(\widehat{\boldsymbol{\beta}}_{t_k})_{k=1}^{K}$ of the binarized R-learner is unique and is given by the OLS estimator of the binarized R-loss function.

D. Additional details about simulated analytical functions in section 6.1

In this section, we consider a treatment T with $K + 1 = 10$ possible values in $\mathcal{T} = \{t_k := \frac{k}{K}, k \in \{0, \dots, K\}\}$, drawn from a uniform distribution, and the following outcome functions.

The linear model outcome for $X \in \mathbb{R}$:

$$Y(t) \mid X \sim \mathcal{N}\big((1+t)X, \sigma^2\big). \tag{105}$$

The multivariate hazard rate (Imbens, 2000) outcome satisfies for $\boldsymbol{X} \in \mathbb{R}^5$:

$$Y(t) \mid \boldsymbol{X} \sim \mathcal{N}\big(t + \|\boldsymbol{X}\| \exp(-t\|\boldsymbol{X}\|), \sigma^2\big). \tag{106}$$

We compute the exact components of each model in the following subsections: the GPS r, the potential outcome models μ_t and the observed outcome model m.

D.1. Computing nuisance components

THE GENERALIZED PROPENSITY SCORE (GPS).
In the first design (RCT), we sample n units such that T and $\boldsymbol{X}$ are independent. The true propensity score is known

$$r(t, \boldsymbol{X}) = \mathbb{P}(T = t) = 1/(K+1) \text{ for } t \in \mathcal{T}. \tag{107}$$

In the second design (observational studies), we combine $K + 2$ samples in a single sample of n units. The first sample $\mathbf{D}_{K+1}$ contains $n_{K+1} = n/2$ units where the treatment is assigned randomly: $\boldsymbol{X}$ and T are independent, $\mathbb{P}(T = t) = 1/(K+1)$, $\boldsymbol{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_5)$ when the hazard rate model is applied and $X \sim \mathcal{U}(0, 1)$ when the linear model is applied. For $k = 0, \dots, K$, the sample $\mathbf{D}_k$ contains $n_k = n/(2(K+1))$ units and the distribution of $(\boldsymbol{X}, T)$ does not respect a RCT setting. For the linear model, the joint distribution of (X, T) is given by:

$$T = \frac{k}{K} \text{ and } X \text{ follows a uniform distribution } \mathcal{U}(I_k) \text{ with } I_k = \left[\frac{k}{K+1}, \frac{k+1}{K+1}\right). \tag{108}$$

For the hazard rate model, the joint distribution of $(\boldsymbol{X}, T)$ is given by:

$$T = \frac{k}{K}, X_1 \text{ follows a truncated standardized normal distribution on } I_k = \left[q_{\frac{k}{K+1}}, q_{\frac{k+1}{K+1}}\right)$$
$$\text{and } X_j \text{ follow a standardized normal distribution } \mathcal{N}(0,1) \text{ for } 2 \leq j \leq 5, \tag{109}$$

where q_α is the α-quantile of the standardized normal distribution with $q_0 = -\infty$ and $q_1 = +\infty$. This strategy of selecting preferentially only observations with certain characteristics is called *preferential selection* sampling and creates thus a selection bias on observed data.

For all $k \in \{0, \dots, K\}$, the true propensity score satisfies for the linear model:

$$r(t_k, x) = \begin{cases} \frac{K+2}{2(K+1)} & \text{if } x \in I_k, \\ \frac{1}{2(K+1)} & \text{otherwise.} \end{cases} \tag{110}$$

and, for the hazard rate model, it satisfies for $\boldsymbol{x} \in \mathbb{R}^5$:

$$r(t_k, \boldsymbol{x}) = \begin{cases} \frac{K+2}{2(K+1)} & \text{if } x_1 \in I_k, \\ \frac{1}{2(K+1)} & \text{otherwise.} \end{cases} \tag{111}$$

Proof. We show proof for the hazard rate model with normal distribution. The proof remains the same for the linear model in a non-randomized setting.

Let A be a given event, and then

$$\mathbb{P}(A) = \sum_{k=0}^{K+1} \frac{n_k}{n} \mathbb{P}_k(A), \tag{112}$$

where $\mathbb{P}$ is the observed probability distribution of the combined sample and $\mathbb{P}_k$ denotes the probability measure induced by (107), (109) and the unconfoundedness assumption 3.1.

Given the treatment $T = t_j$ and covariate vector $\boldsymbol{x} = (x, x_2, \ldots, x_5)$, we have

$$\begin{aligned}
r(t_j, \boldsymbol{x}) &= \mathbb{P}(T = t_j \mid X_1 = x) \\
&= \lim_{\delta \to 0} \mathbb{P}(T = t_j \mid X_1 \in [x, x + \delta]) \\
&= \lim_{\delta \to 0} \frac{\mathbb{P}\left(T = t_j, X_1 \in [x, x + \delta]\right)}{\mathbb{P}\left(X_1 \in [x, x + \delta]\right)}.
\end{aligned} \tag{113}$$

On the one hand,

$$\begin{aligned}
\mathbb{P}(T = t_j, X_1 \in [x, x + \delta]) &= \sum_{k=0}^{K+1} \frac{n_k}{n} \mathbb{P}_k(T = t_j, X_1 \in [x, x + \delta]) \\
&= \frac{n_j}{n} \mathbb{P}_j(T = t_j, X_1 \in [x, x + \delta]) + \frac{n_{K+1}}{n} \mathbb{P}_{K+1}(T = t_j, X_1 \in [x, x + \delta]) \\
&= \frac{n_j}{n} \mathbb{P}_j(X_1 \in [x, x + \delta]) + \frac{n_{K+1}}{n} \mathbb{P}_{K+1}(T = t_j)\mathbb{P}_{K+1}(X_1 \in [x, x + \delta]) \\
&= \frac{1}{2(K+1)} \mathbb{P}_j(X_1 \in [x, x + \delta]) + \frac{1}{2(K+1)} \mathbb{P}_{K+1}(X_1 \in [x, x + \delta]).
\end{aligned} \tag{114}$$

For $x \in \mathbb{R}$, there exists a unique j_0 such that $x \in I_{j_0}$. For δ small enough, we have $[x, x + \delta] \subset I_{j_0}$ and, consequently, $[x, x + \delta] \cap I_j = \emptyset$ for all $j \neq j_0$. This implies:

$$\mathbb{P}_j(X_1 \in [x, x + \delta]) = \frac{\mathbb{P}_{K+1}\left(X_1 \in [x, x + \delta], X_1 \in I_j\right)}{\mathbb{P}_{K+1}\left(X_1 \in I_j\right)} = \frac{\mathbb{P}_{K+1}\left(X_1 \in [x, x + \delta]\right)}{\mathbb{P}_{K+1}\left(X_1 \in I_j\right)} \mathbf{1}\{j = j_0\}. \tag{115}$$

Therefore,

$$\begin{aligned}
\mathbb{P}(T = t_j, X_1 \in [x, x + \delta]) &= \frac{1}{2(K+1)} \mathbb{P}_{K+1}(X_1 \in [x, x + \delta]) \left(\frac{\mathbf{1}\{j = j_0\}}{\mathbb{P}_{K+1}(X_1 \in I_{j_0})} + 1\right) \\
&= \left(\frac{1}{2}\mathbf{1}\{j = j_0\} + \frac{1}{2(K+1)}\right)\mathbb{P}_{K+1}(X_1 \in [x, x + \delta]).
\end{aligned} \tag{116}$$

On the other hand,

$$\begin{aligned}
\mathbb{P}(X_1 \in [x, x + \delta]) &= \sum_{k=0}^{K+1} \frac{n_k}{n} \mathbb{P}_k(X_1 \in [x, x + \delta]) \\
&= \frac{1}{2(K+1)} \sum_{k=0}^{K} \frac{\mathbb{P}_{K+1}\left(X_1 \in [x, x + \delta], X_1 \in I_k\right)}{\mathbb{P}_{K+1}\left(X_1 \in I_k\right)} + \frac{1}{2}\mathbb{P}_{K+1}(X_1 \in [x, x + \delta]) \\
&= \frac{1}{2(K+1)} \frac{\mathbb{P}_{K+1}(X_1 \in [x, x + \delta])}{\mathbb{P}_{K+1}(X_1 \in I_{j_0})} + \frac{1}{2}\mathbb{P}_{K+1}(X_1 \in [x, x + \delta]) \\
&= \frac{1}{2}\mathbb{P}_{K+1}(X_1 \in [x, x + \delta]) + \frac{1}{2}\mathbb{P}_{K+1}(X_1 \in [x, x + \delta]) \\
&= \mathbb{P}_{K+1}(X_1 \in [x, x + \delta])
\end{aligned} \tag{117}$$

Finally,

$$
\begin{aligned}
r(t_j, \boldsymbol{x}) &= \lim_{\delta \to 0} \frac{\mathbb{P}\left(T = t_j, X_1 \in [x, x + \delta]\right)}{\mathbb{P}\left(X_1 \in [x, x + \delta]\right)} \\
&= \lim_{\delta \to 0} \frac{\left(\frac{1}{2}\mathbf{1}\{j = j_0\} + \frac{1}{2(K+1)}\right)\mathbb{P}_{K+1}(X_1 \in [x, x + \delta])}{\mathbb{P}_{K+1}(X_1 \in [x, x + \delta])} \\
&= \frac{1}{2}\mathbf{1}\{j = j_0\} + \frac{1}{2(K+1)} \\
&= \frac{(K+1)\mathbf{1}\{j = j_0\} + 1}{2(K+1)} \\
&= \begin{cases} \frac{K+2}{2(K+1)} & \text{if } x \in I_j, \\ \frac{1}{2(K+1)} & \text{otherwise.} \end{cases}
\end{aligned}
\tag{118}
$$

$\square$

THE POTENTIAL OUTCOME MODELS.
The potential outcome models are given directly by the conditional mean. For the linear model, μ_t satisfies for all $t \in \mathcal{T}$ and $x \in [0, 1]$:

$$
\mu_t(x) = (1 + t)x.
\tag{119}
$$

For the hazard rate model, μ_t is given by:

$$
\mu_t(\boldsymbol{x}) = t + \|\boldsymbol{x}\| \exp\left(-t\|\boldsymbol{x}\|\right).
\tag{120}
$$

THE OBSERVED OUTCOME MODELS.
For the linear model, the observed outcome model m can be computed as:

$$
\begin{aligned}
m(x) &= \mathbb{E}(Y_{\text{obs}} \mid X = x) \\
&= \mathbb{E}((1 + T)\boldsymbol{X} \mid X = x) \\
&= (1 + \mathbb{E}(T \mid X = x))x \\
&= \left(1 + \sum_{k=0}^{K} r(t_k, x)t_k\right)x,
\end{aligned}
\tag{121}
$$

where r is given by (110).

and, for the hazard rate model, m can be computed as:

$$
\begin{aligned}
m(\boldsymbol{x}) &= \mathbb{E}(\mathbb{E}(Y_{\text{obs}} \mid \boldsymbol{X}, T) \mid \boldsymbol{X} = \boldsymbol{x}) \\
&= \mathbb{E}(T + \|\boldsymbol{X}\| \exp\left(-T\|\boldsymbol{X}\|\right) \mid \boldsymbol{X} = \boldsymbol{x}) \\
&= \mathbb{E}(T \mid \boldsymbol{X} = \boldsymbol{x}) + \|\boldsymbol{x}\| \, \mathbb{E}(\exp\left(-T\|\boldsymbol{X}\|\right) \mid \boldsymbol{X} = \boldsymbol{x}) \\
&= \sum_{k=0}^{K} r(t_k, \boldsymbol{x})t_k + \sum_{k=0}^{K} \|\boldsymbol{x}\| \, r(t_k, \boldsymbol{x}) \exp\left(-t_k\|\boldsymbol{x}\|\right),
\end{aligned}
\tag{122}
$$

where r is given by (111).

D.2. Additional numerical results and plots.

In this section, we present the results of different simulations and scenarios for linear and hazard rate models with $K+1 = 10$, $n = 2000$ for the linear model, and $n = 10000$ for the Hazard rate model. In the randomized setting, the sample $\mathbf{D}_{\text{obs}}$ is sampled randomly, and the propensity score is given by (107). In a non-randomized setting, the sample $\mathbf{D}_{\text{obs}}$ is given by preferential selection as described in Section D and the GPS is given by (110). When we say that *"the models' nuisance components are exact"*, we replace the expression of the estimators $\widehat{\mu}_t, \widehat{m}$ or $\widehat{r}$ with the expressions obtained in Section D.

Linear model in a randomized setting.

Table 5. **mPEHE** and **sdPEHE** for three different ML base-learners; Case where nuisance components are exact.

Meta-learner	XGBoost	RandomForest	Linear Model
M-Learner	2.23 (1.20)	2.09 (1.08)	0.087 (0.096)
DR-Learner	0.165 (0.034)	0.140 (0.033)	$9.65\ (7.84)\ 10^{-3}$
X-Learner	**0.022 (0.004)**	**0.029 (0.004)**	**1.42 (1.46)** 10^{-3}
RLin-Learner		$10.2\ (8.42)\ 10^{-3}$	

Table 6. **mPEHE** and **sdPEHE** for three different ML base-learners; Case where nuisance components are well-specified.

Meta-learner	XGBoost	RandomForest	Linear Model
T-Learner	0.065 (0.019)	0.041 (0.016)	$10.0\ (8.37)\ 10^{-3}$
S-Learner	**0.033 (0.018)**	0.032 (0.028)	**3.03 (2.42)** 10^{-3}
*Nv*X-Learner	0.060 (0.019)	**0.037 (0.016)**	$10.0\ (8.37)\ 10^{-3}$
M-Learner	1.25 (0.610)	1.22 (0.621)	0.201 (0.191)
DR-Learner	0.068 (0.019) – 0.063 (0.020)	0.068 (0.018) – 0.068 (0.018)	$10.0\ (9.14)$ – **5.27 (4.36)** 10^{-3}
X-Learner	0.063 (0.020) – **0.033 (0.017)**	0.045 (0.016) – 0.061 (0.040)	$10.0\ (8.37)$ – $3.28\ (2.98)\ 10^{-3}$
RLin-Learner	0.135 (0.130)	0.137 (0.128)	0.073 (0.063)

For the DR- and X-learners: μ_t are estimated by T-learning (left value) or S-learning (right value).

Table 7. **mPEHE** and **sdPEHE** for three different ML base-learners; Case where the propensity score is misspecified.

Meta-learner	XGBoost	RandomForest	Linear Model
M-Learner	3.86 (2.95)	3.68 (2.80)	1.45 (0.99)
DR-Learner	**0.145 (0.108)**	**0.245 (0.179)**	**0.014 (0.015)**
X-Learner	0.033 (0.017)	0.061 (0.040)	$3.28\ (2.98)\ 10^{-3}$
RLin-Learner	0.336 (0.272)	0.338 (0.277)	0.338 (0.215)

Table 8. **mPEHE** and **sdPEHE** for three different ML base-learners; Case where the outcome models are misspecified.

Meta-learner	XGBoost	RandomForest	Linear Model
M-Learner	1.25 (0.610)	1.22 (0.621)	0.201 (0.191)
DR-Learner	0.811 (0.386)	0.888 (0.378)	0.308 (0.362)
X-Learner	0.304 (0.330)	0.303 (0.330)	0.275 (0.328)
RLin-Learner		**0.073 (0.062)**	

Table 9. **mPEHE** and **sdPEHE** for three different ML base-learners; Case where nuisance components are misspecified.

Meta-learner	XGBoost	RandomForest	Linear Model
M-Learner	3.86 (2.95)	3.68 (2.80)	1.45 (0.99)
DR-Learner	1.87 (1.31)	2.09 (1.48)	0.828 (0.496)
X-Learner	0.304 (0.330)	0.303 (0.330)	0.275 (0.328)
RLin-Learner		**0.277 (0.178)**	

Linear model in non-randomized setting

Table 10. **mPEHE** and **sdPEHE** for three different ML base-learners; Case where nuisance components are exact.

Meta-learner	XGBoost	RandomForest	Linear Model
M-Learner	3.69 (1.80)	2.96 (1.47)	0.153 (0.177)
DR-Learner	0.276 (0.081)	0.206 (0.056)	$10.9\ (8.47)\ 10^{-3}$
X-Learner	**0.022 (0.004)**	**0.028 (0.004)**	**1.69 (1.11) 10^{-3}**
RLin-Learner		$11.0\ (11.1)\ 10^{-3}$	

Table 11. **mPEHE** and **sdPEHE** for three different ML base-learners; Case where nuisance components are well-specified.

Meta-learner	XGBoost	RandomForest	Linear Model
T-Learner	0.067 (0.023)	0.043 (0.016)	$10.5\ (10.0)\ 10^{-3}$
*Reg*T-Learner	0.059 (0.021)	0.042 (0.016)	$13.0\ (11.1)\ 10^{-3}$
S-Learner	**0.033 (0.018)**	0.060 (0.055)	**6.46 (5.11) 10^{-3}**
*Nv*X-Learner	0.062 (0.023)	**0.039 (0.017)**	$9.56\ (10.0)\ 10^{-3}$
M-Learner	1.35 (0.82)	1.14 (0.72)	0.196 (0.153)
DR-Learner	0.065 (0.022) – 0.065 (0.027)	0.069 (0.026) – 0.096 (0.056)	$13.0\ (11.1) - 8.17\ (6.30)\ 10^{-3}$
X-Learner	0.059 (0.021) – **0.034 (0.017)**	0.046 (0.016) – 0.084 (0.058)	$14.7\ (11.6) - 6.49\ (5.23)\ 10^{-3}$
RLin-Learner	0.155 (0.137)	0.124 (0.114)	0.108 (0.097)

For the DR- and X-learners: μ_t are estimated by T-learning (left value) or S-learning (right value).

Hazard rate model in randomized setting

Table 12. **mPEHE** and **sdPEHE** for three different ML base-learners; Case where nuisance components are exact.

Meta-learner	XGBoost	RandomForest	Linear Model
M-Learner	4.27 (1.45)	4.21 (1.28)	0.529 (0.188)
DR-Learner	0.127 (0.022)	0.144 (0.044)	0.106 (0.094)
X-Learner	**0.045 (0.025)**	**0.087 (0.049)**	**0.106 (0.094)**
RLin-Learner		0.107 (0.094)	

Table 13. **mPEHE** and **sdPEHE** for three different ML base-learners; Case where nuisance components are well-specified.

Meta-learner	XGBoost	RandomForest	Linear Model
T-Learner	0.175 (0.046)	0.263 (0.144)	**0.113 (0.091)**
S-Learner	0.159 (0.048)	0.260 (0.130)	0.662 (0.421)
*Nv*X-Learner	0.176 (0.091)	0.313 (0.188)	**0.113 (0.092)**
M-Learner	1.57 (0.471)	1.79 (0.453)	0.824 (0.522)
DR-Learner	0.165 (0.049) – 0.159 (0.047)	0.281 (0.144) – 0.275 (0.137)	0.114 (0.094) – 0.464 (0.286)
X-Learner	0.163 (0.057) – **0.154 (0.051)**	0.279 (0.157) – 0.279 (0.146)	**0.113 (0.092)** – 0.644 (0.380)
RLin-Learner	0.245 (0.136)	0.241 (0.136)	0.717 (0.450)

For the DR- and X-learners: μ_t are estimated by T-learning (left value) or S-learning (right value).

Hazard rate model in non-randomized setting

Table 14. **mPEHE** and **sdPEHE** for three different ML base-learners; Case where nuisance components are exact.

Meta-learner	XGBoost	RandomForest	Linear Model
M-Learner	6.28 (1.88)	5.74 (1.60)	3.72 (1.42)
DR-Learner	0.138 (0.029)	0.139 (0.044)	**0.110 (0.097)**
X-Learner	**0.044 (0.025)**	**0.087 (0.050)**	**0.110 (0.097)**
RLin-Learner		0.299 (0.176)	

Table 15. **mPEHE** and **sdPEHE** for three different ML base-learners; Case where nuisance components are well-specified.

Meta-learner	XGboost	RandomForest	Linear Model
T-Learner	0.183 (0.039)	0.286 (0.155)	0.129 (0.094)
*Reg*T-Learner	0.176 (0.044)	0.286 (0.155)	**0.121 (0.098)**
S-Learner	0.176 (0.056)	0.306 (0.153)	0.671 (0.428)
*Nv*X-Learner	0.190 (0.096)	0.336 (0.200)	0.129 (0.094)
M-Learner	1.61 (0.505)	1.58 (0.472)	0.906 (0.557)
DR-Learner	0.168 (0.045) - 0.178 (0.048)	0.304 (0.158) – 0.322 (0.162)	**0.121 (0.098)** – 0.518 (0.327)
X-Learner	**0.167 (0.053)** – 0.172 (0.057)	0.302 (0.169) – 0.332 (0.167)	0.120 (0.094) – 0.652 (0.388)
RLin-Learner	0.231 (0.081)	**0.186 (0.123)**	1.05 (0.651)

For the DR- and X-learners: μ_t are estimated by T-learning (left value) or S-learning (right value).

D.3. Asymptotic performances when K increases.

In this subsection, we consider the effect of increasing K on the hazard rate function with XGBoost. For each value K, we sample $J = 10$ different non-randomized samples following the *preferential selection* as defined previously in Appendix D.1. The **mPEHE** is then computed by averaging the **mPEHE** over the $J = 10$ samples. The results are drawn in the figure below.
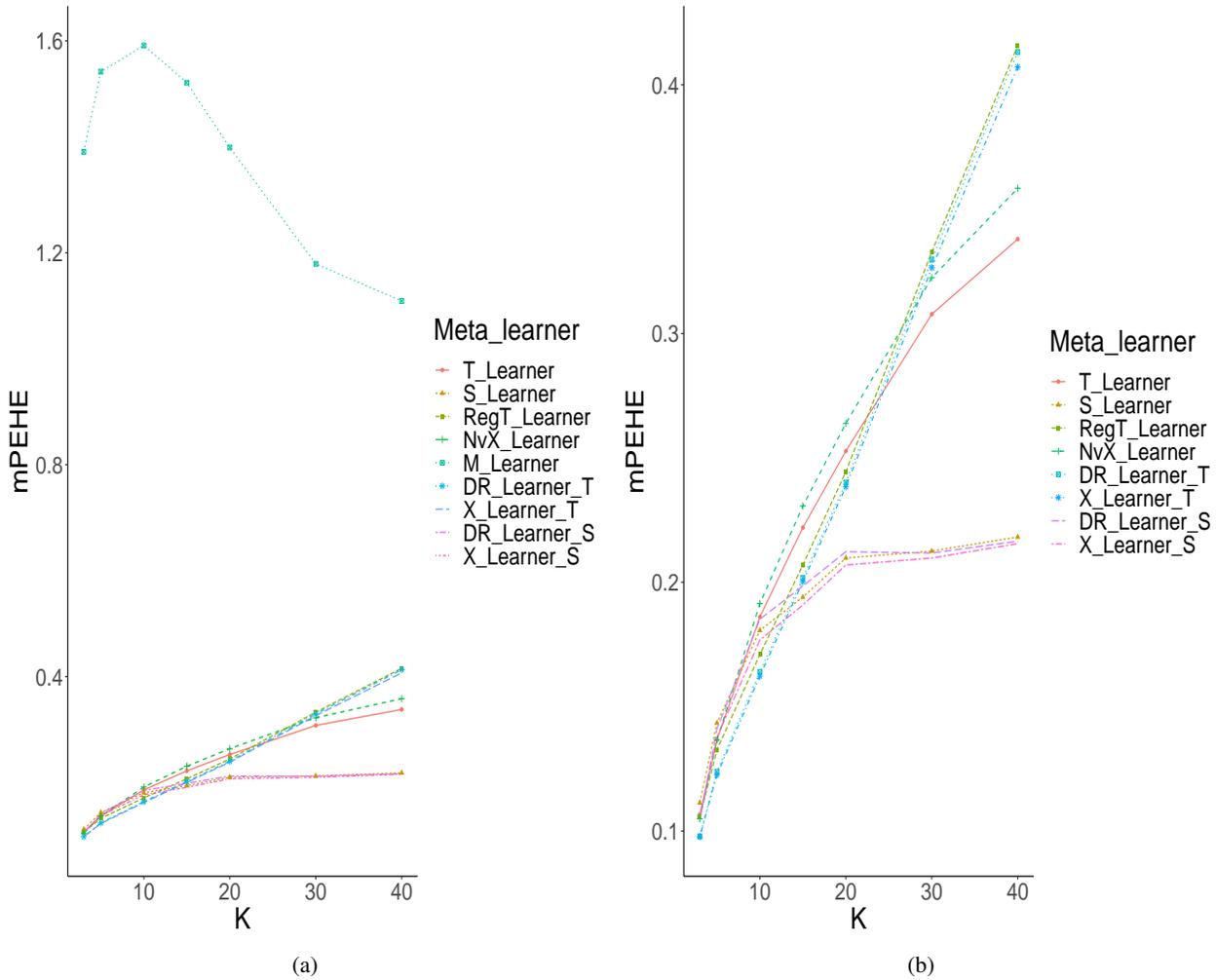

Figure 1. Variation of meta-learners' performances as functions of the number of possible treatment values K for the hazard rate function in an observational design setting. (a): All meta-learners; (b): Without the M-learner

When we consider the effect of increasing K on the hazard rate function with a linear model (with $p = 2$), we notice the expected effect of K on the M-learner: The error bound is increasing with K. This means that the behaviour of the M-learner with XGBoost can be explained by the over-fitting of the base-learner.

(a)   (b)
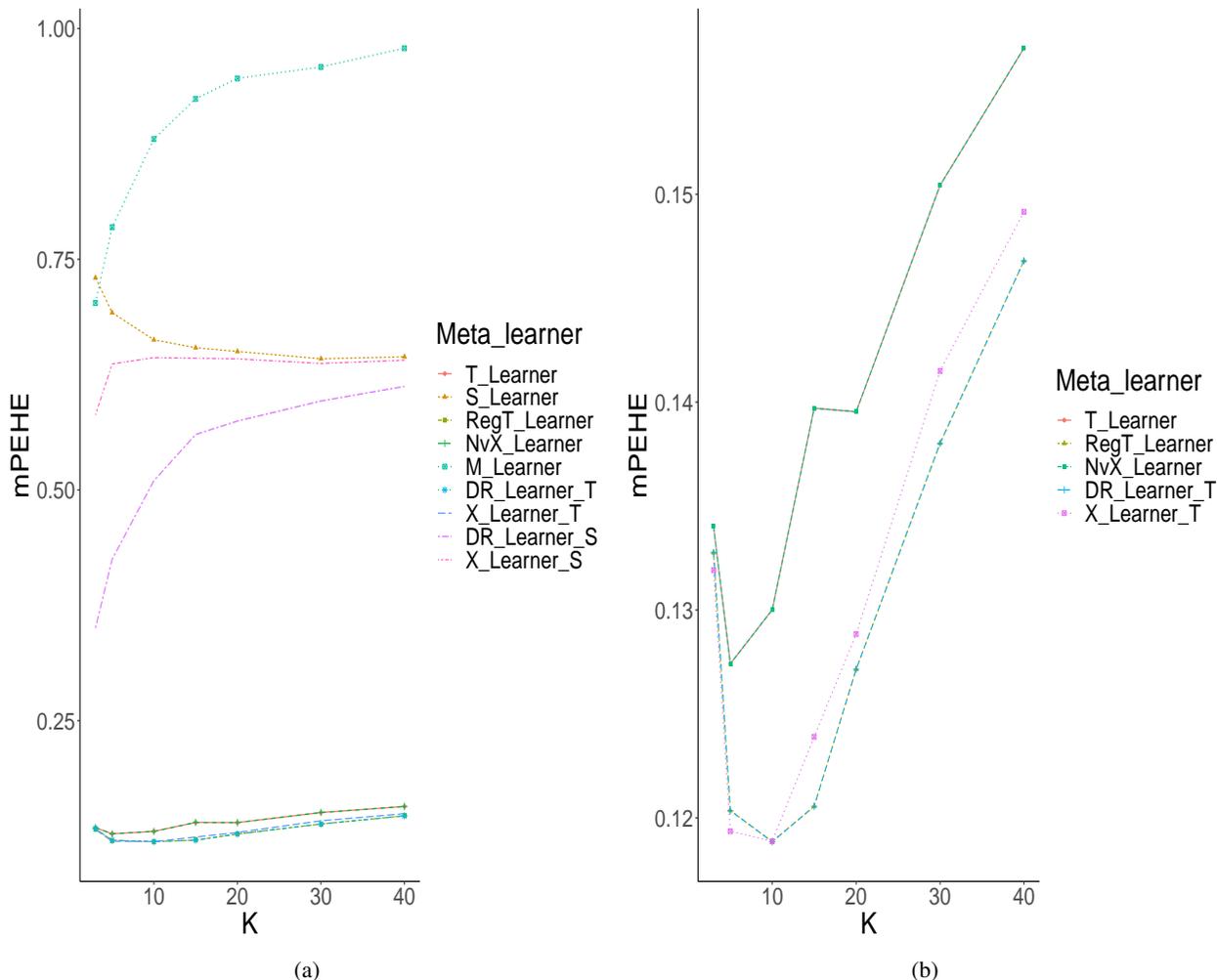
Figure 2. Variation of meta-learner's performances when the number of possible treatment values K for the hazard rate function in observational design setting with a linear model. (a): All meta-learners; (b): Focus on the T-learning methods.

E. Description of the semi-synthetic dataset.

Motivation

The difficulty in evaluating a causal model's performance in real-world applications motivates the need to create a semi-synthetic dataset. In this subsection, we consider a multistage fracturing Enhanced Geothermal System (EGS).

Enhanced Geothermal Systems (EGS) are geothermal wells that generate geothermal energy by creating fluid connectivity in low-permeability conductive rocks through hydraulic, thermal, or chemical stimulation. The EGS concept involves extracting heat by constructing a subsurface fracture system to which water can be added via injection wells. Indeed, rocks are permeable due to slight fractures and pore spaces between mineral grains, and the injected water is heated by contact with the rock and returns to the surface through production wells. Moreover, Enhanced geothermal systems (EGS) have a high potential for developing and supplying renewable energy sources that are more efficient and cheaper than traditional hydrocarbon resources.

For energy companies, the goal is to optimize the design of the geothermal well (fracture spacing, Lateral Length etc.) to generate the maximum geothermal energy. However, some economic and operational problems present challenges: On the one hand, if the fractures are too small or too few, rocks will not be exploited sufficiently. On the other hand, if the number of fractures in a given rock is too high, the fractures may cool down faster. We would have a costly design that will not

maximize the extracted heat.

We assume that the heat extraction performance of the EGS satisfies the following physical model:

$$Q_{well} = Q_{fracture} \times \ell_L/d \times \eta_d, \tag{123}$$

where Q_{well} is the heat extraction performance delivered by the well (output), $Q_{fracture}$ is the *unknown* heat extraction performance from a single fracture that can be generated using a complex seven-parameter model, including reservoir characteristics and fracture design, ℓ_L is the Lateral Length of the well, d is the average spacing between two fractures and η_d is the stage efficiency penalizing the individual contribution when fractures are close to each other. We refer to Figure 3 for a graphical description of the EGS and its inputs/output.

Finally, the model in (123) respects the unconfoundedness assumption 3.1, and we can control all its variables in the simulations. We note that, in practice, all inputs are continuous with a given density. However, we discretize these variables in their input space to create a full factorial design.

Description of the data-set

This section describes the data-generating process of our semi-synthetic dataset simulating the heat delivered by a multistage fracturing EGS. The process involved the creation of a conceptual reservoir model and the modelling of multiple wells' completion scenarios. The output (heat extraction performance) obtained from physics-based simulation experiments was tabulated with inputs in the semi-synthetic dataset.

The input data for the model were fabricated to ease confidentiality and non-disclosure information issues. However, data has been selected from reliable sources such as field observations, journals and books to be within the range of interest. Doing so allowed the building of a plain but representative reservoir model that would provide realistic results of an EGS.

The heat extraction performance from a single fracture ($Q_{fracture}$) is determined using fracture length, fracture height, fracture width, fracture permeability, reservoir porosity, reservoir permeability and pore pressure. Modelling and simulation work were done using preprocessor and reservoir simulation tools PETREL and ECLIPSE.

The four physical parameters of the fracture were investigated, and the list of values used for each parameter can be observed in Table 16. In the end, $10 \times 10 \times 2 \times 3 = 600$ fracture's simulation cases have been realized.

Table 16. Fracture parameters and their range of variation for simulations.

Variable	Range of variation
Fracture length (ft)	$[100, 1000]$ by a step of 100 ft
Fracture height (ft)	$[50, 500]$ by a step of 50 ft
Fracture width (in)	$\{0.1, 0.2\}$
Fracture Permeability (md)	$\{30000, 85000, 19000\}$

To emulate distinct reservoir schemes, it was decided to vary three main parameters; porosity, permeability and pore pressure. For porosity and permeability, the simulator takes the minimum and maximum values and estimates the physical properties across the reservoir. Three different multipliers were applied to define three (Low, Base and High) scenarios. Concerning pore pressure, three specific values were defined to simulate under-normal, normal (base) and overpressure (high) gradient conditions. Therefore, $3 \times 3 \times 3 = 27$ possible scenarios were defined. Table 17 displays the range of minimum and maximum values for the three reservoir parameters to be varied.

Table 17. Reservoir parameters and their range of variation for simulations.

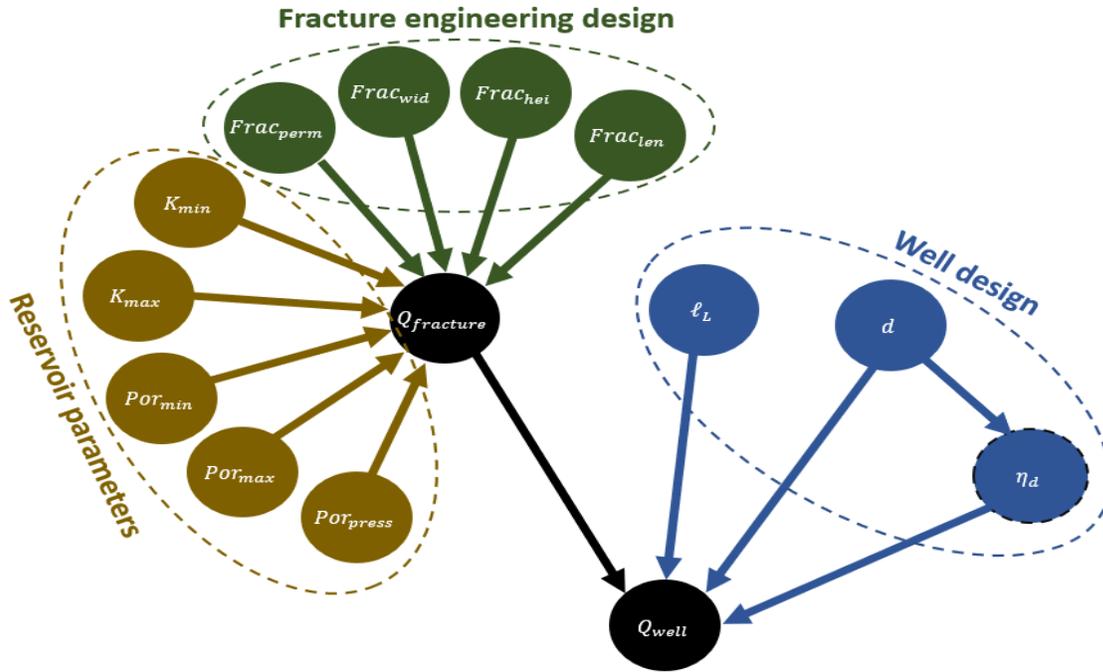| Variable	Range of variation
(K_{min}, K_{max}) (md)	$\{(0.0054, 0.0157), (0.054, 0.157), (0.109, 0.314)\}$
(Por_{min}, Por_{max}) (dec)	$\{(0.0054, 0.0157), (0.054, 0.157), (0.109, 0.314)\}$
Pore pressure (psi)	$\{5000, 7000, 9000\}$

Figure 3. The Causal DAG associated with the multistage EGS. Nodes in yellowish brown represent the reservoir characteristics, they can only be simulated, but in reality, we cannot intervene in these variables. Nodes in Dark green represent the fracture design. Engineers control them, and intervening in them is possible whenever there is a need to make a new fracture in the well. Nodes in blue represent a well's design and can be chosen arbitrarily by engineers or statisticians. Nodes in black denote the outputs. $Q_{fracture}$ is only given by the simulator, whereas Q_{well} is given by the physical model in (123). Note that this graph contains nine nodes, but both $K_{\min}$ and $K_{\max}$ represent the same physical parameter K, and the same remark is valid for $Por_{\min}$ and $Por_{\max}$.

By combining different reservoir scenarios with single fracture simulations, we obtained a single dataset with 16,200 possible cases for a fracture in a reservoir then we simulated the heat extraction performance for each experiment. Simulation's results were tabulated in the dataset "*Single_Fracture_Simulation_Cases_16200.csv*".

The next step is to define well characteristics (lateral lengths and fracture spacing) to evaluate the heat extraction performance of the well when reservoir and fracture properties are not changed.

Table 18. Well parameters and their range of variation.

Variable	Range of variation
Lateral length (ft)	$[2000, 14000]$ by a step of 1000 ft
Fracture spacing (ft)	$[100, 500]$ by a step of 100 ft

Regarding the spacing efficiency coefficient, this coefficient was used to model interactions between fractures and penalize the heat extraction performance of a single fracture in the presence of other close fractures, that is, when the spacing between two fractures is small. Indeed, if the fractures are spaced too close, there may not be enough thermal energy in the rock to heat the water, decreasing the heat extraction efficiency. Modelling this efficiency led to the efficiency table "*Fracture_Efficency.csv*" that describes what would be the well's heat performance behaviour with respect to the fracture spacing selected. Based on this table, one can interpolate the efficiency to draw the curve (see Figure 4) and thus obtain the spacing efficiency coefficient for any desired value fracture spacing.
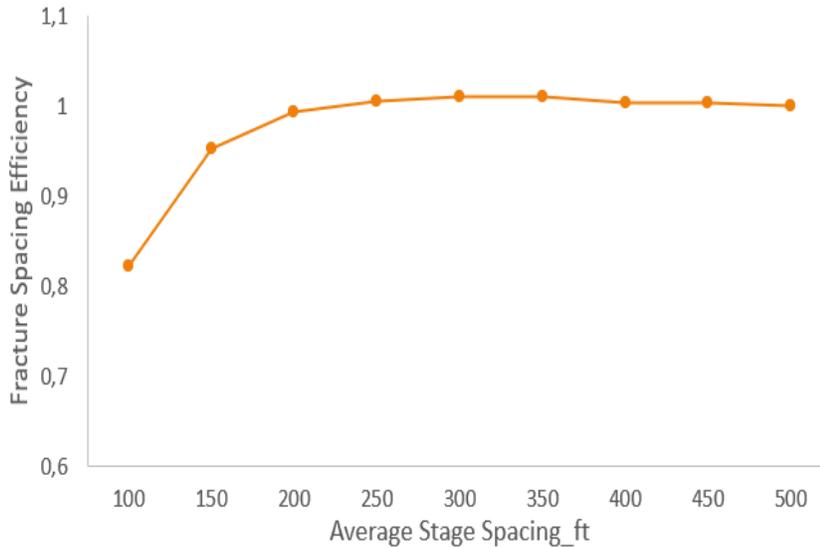
Figure 4. Cross plot between fracture spacing efficiency and average stage spacing.

The final generation of the semi-synthetic dataset "*Main_Dataset.csv*" was achieved by combining two main tables created using the R programming language. This table allows calculating the heat performance of a well for any lateral length and fracture spacing between 500 ft and 100 ft with the associated spacing efficiency coefficient defined in the efficiency table, following the physical model in (123).

The three datasets are available at `https://github.com/nacharki/multipleT-MetaLearners`.

Finally, we emphasize that this study's design methodology focused on generating a semi-synthetic dataset using reservoir numerical simulation and creating a new benchmarking dataset for comparing and validating causal inference methods. Indeed, following the last step of forming the final dataset "*Main_Dataset.csv*", any user can define different distributions (with different values) on lateral lengths in the range $[2000, 14000]$ and fracture spacing in range $[100, 500]$, pick-up the corresponding spacing efficiency coefficients using the curve drawn in Figure 4 and generate a new semi-synthetic dataset by extrapolating them with "*Single_Fracture_Simulation_Cases_16200.csv*" dataset.

The creation of a non-randomized biased dataset.
The idea of this step was to create a collection of biased data from the main semi-synthetic dataset to emulate observational data found in real-world situations. For example, geothermal wells with larger lateral lengths are likely to have more fractures (expensive wells are located in better geological areas). The opposite is seen for smaller wells that tend to be associated with fewer fractures. This situation creates a discrepancy between what engineers expect from physical models and what they observe in the field data. The biased data, with 9,992 observations, was generated by following the *preferential selection* strategy from the main dataset. Figure 5 shows the difference between the *real* heat extraction performance of the EGS and the observed heat extraction performance on the field: low (under-estimated) heat performance for small wells and high (over-estimated) heat performance for large wells.
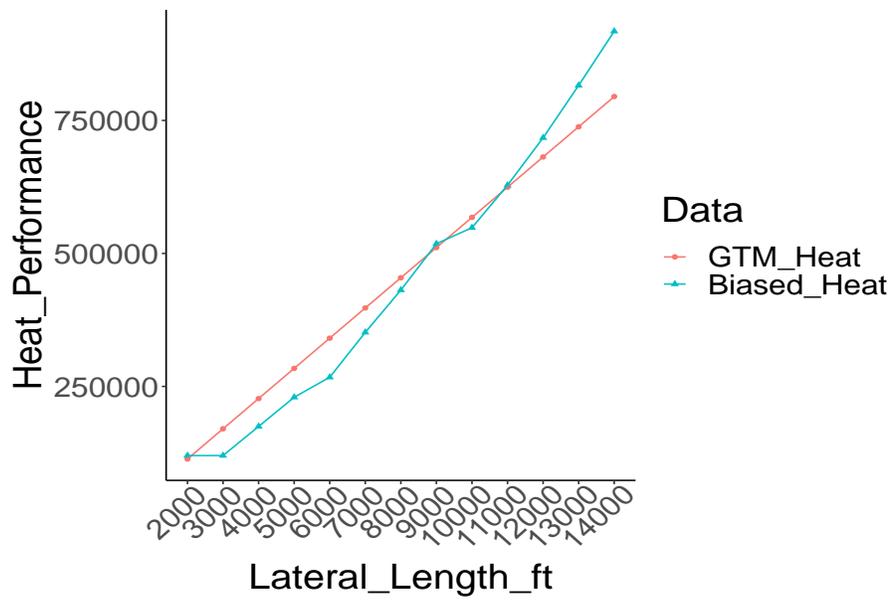
Figure 5. An illustration of selection bias on the heat performance. Red line: The heat extraction performance on the main dataset (i.e. Ground Truth Model). Blue line: The heat performance on the biased dataset (i.e. observed response).